·论著·

多模态模型对内镜逆行胰胆管造影术后胰腺炎的预测价值

杨婷　许祐铭　于红刚
武汉大学人民医院消化内科，武汉 430060
杨婷现在岳阳市中心医院消化内科，岳阳 414000
通信作者：于红刚，Email：yuhonggang@whu.edu.cn

【摘要】 目的　通过多步骤开发和验证多模态数据模型，探究模型对内镜逆行胰胆管造影术（endoscopic retrograde cholangiopancreatography，ERCP）后胰腺炎（post-ERCP pancreatitis）的预测价值。**方法**　回顾性收集武汉大学人民医院2016年1月至2022年9月行ERCP的860份患者病历作为模型训练集，收集该医院同时段行ERCP的367份患者病历作为内部测试集，回顾性收集2021年2月至2022年1月南京大学医学院附属鼓楼医院行ERCP的86份患者病历和同济大学第十人民医院行ERCP的40份患者病历共126份病历作为外部测试集进行模型验证。通过文献检索，识别了与PEP相关的49个电子病历特征和1个十二指肠乳头宏观形态（macroscopic shape of major duodenal papilla，MDP）图像特征。首先，49个电子病历特征被分为基线、诊断、操作和预防策略四个类别。随后，四类特征被逐个添加到训练集中，构建了单模态模型1~4。然后，将MDP特征添加到模型1~4的训练集中以构建多模态模型5~8。最后，计算8个模型在内部和外部测试集中预测ERCP后胰腺炎的准确率、灵敏度、特异度、受试者工作特征曲线下面积（area under the receiver operating characteristic curve，AUC）、阳性预测值（positive predictive value，PPV）和阴性预测值（negative predictive value，NPV），比较多模态模型与相应的单模特模型的表现，比较相邻多模态模型的表现。**结果**　在内部测试集中，随着特征类别的添加，多模态模型的准确率、灵敏度、特异度、AUC、PPV、和NPV均稳步提升。与模型2相比，添加了MDP特征的模型6将准确率从73.57%（95%*CI*：69.04%~78.10%）显著提高至86.92%（95%*CI*：83.46%~90.39%）（χ^2 =20.64，*P*<0.001），特异度从73.45%（95%*CI*：68.73%~78.18%）显著提高至86.43%（95%*CI*：82.77%~90.09%）（χ^2 =17.81，*P*<0.001）。在外部测试集中有类似的趋势。在内、外部测试集中，模型8的表现均是最优的，其AUC值分别为0.91（95%*CI*：0.86~0.96）、0.94（95%*CI*：0.90~0.98）。在模型8中，MDP图像特征、困难插管的权重分别位列第1、第2。**结论**　多模态模型预测PEP的效能较好。MDP图像特征和困难插管对PEP预测均具有重要意义。

【关键词】 胰胆管造影术，经内镜逆行；　术后胰腺炎；　机器学习；　预测模型
基金项目：湖北省消化疾病微创诊治临床医学研究中心项目（2022DCC004）

Predictive value of multimodal models for post-endoscopic retrograde cholangiopancreatography pancreatitis

Yang Ting, Xu Youming, Yu Honggang
Department of Gastroenterology, Renmin Hospital of Wuhan University, Wuhan 430060, China
Yang Ting is working at the Department of Gastroenterology, Yueyang Central Hospital, Yueyang 414000, China
Corresponding author: Yu Honggang, Email: yuhonggang@whu.edu.cn

DOI：10.3760/cma.j.cn321463-20250410-00038
收稿日期　2025-04-10　　**本文编辑**　钱程
引用本文：杨婷，许祐铭，于红刚. 多模态模型对内镜逆行胰胆管造影术后胰腺炎的预测价值[J]. 中华消化内镜杂志，XXXX，XX(XX)：1-8. DOI：10.3760/cma.j.cn321463-20250410-00038.

优先出版

【Abstract】 **Objective** To develop and validate multimodal models through multiple steps and investigate the prediction value of models for post-endoscopic retrograde cholangiopancreatography (ERCP) pancreatitis (PEP). **Methods** A total of 860 medical records of patients who received ERCP in Renmin Hospital of Wuhan University from January 2016 to September 2022 were retrospectively collected as the training set, 367 medical records of patients who received ERCP in the same hospital during the same time period were retrospectively collected as internal test set. Additionally, a total of 126 medical records (86 from Nanjing Drum Tower Hospital, the Affiliated Hospital of Nanjing University Medical School, and 40 from the Tenth People's Hospital of Tongji University) were retrospectively collected from February 2021 to January 2022 as the external test set for model validation. Through literature review, 49 electronic medical record features and the macroscopic shape of major duodenal papilla (MDP) feature associated with PEP were identified. Initially, the electronic medical record features were categorized into baseline, diagnosis, technique and prevention strategies. Subsequently, single-modal models 1-4 were incrementally constructed by incorporating these four categories of features into the training set. Then, the MDP feature was added to the training set of models 1-4 to develop multimodal models 5-8. All models underwent testing using both internal and external test sets and their accuracy, sensitivity, specificity, area under the receiver operating characteristic curve (AUC), positive predictive value (PPV) and negative predictive value (NPV) were compared. **Results** In the internal test set, the accuracy, sensitivity, specificity, AUC, PPV and NPV of multimodal models improved steadily with increasing numbers of feature categories. Adding the MDP feature to model 2 (resulting in model 6) , the accuracy was significantly improved from 73.57% (95%*CI*: 69.04% -78.10%) to 86.92% (95%*CI*: 83.46%-90.39%) (χ^2=20.64, *P*<0.001), the specificity from 73.45% (95%*CI*: 68.73%~78.18%) to 86.43% (95%*CI*:82.77%-90.09%) (χ^2=17.81, *P*<0.001). A similar trend was observed in the external test set. Notably, the performance of model 8 were the best, whose AUC were 0.91 (95%*CI*: 0.86-0.96) in the internal test and 0.94 (95%*CI*:0.90-0.98) in the external test, respectively. In model 8, the image feature of MDP had the highest weight, followed by difficult cannulation. **Conclusion** The multimodal model for predicting PEP demonstrates good performance. The image feature of MDP and difficult cannulation are crucial predictors for PEP.

【Key words】 Cholangiopancreatography, endoscopic retrograde; Pancreatitis; Machine learning; Prediction model

Fund program: Project of Hubei Provincial Clinical Medical Research Center for Minimally Invasive Diagnosis and Treatment of Digestive Diseases (2022DCC004)

优先出版

内镜逆行胰胆管造影术(endoscopic retrograde cholangiopancreatography, ERCP)是治疗胆胰管疾病的主要手段,具有侵入性,可导致多种并发症,ERCP术后胰腺炎(post-ERCP pancreatitis, PEP)为其特有的并发症[1],在高风险人群中发生率高达15%[2]。因此,诸多学者致力于探究PEP相关因素以降低其发生率、减轻严重程度。一般将已确定的危险因素分为患者相关因素和手术相关因素[3]。此外,十二指肠乳头的宏观形态(macroscopic shape of major duodenal papilla, MDP)[4]、内镜医师操作经验[5]均为困难插管的危险因素,从而导致PEP的发生。已发现的保护因素则包括非甾体类消炎药(nonsteroidal anti-inflammatory drugs, NSAIDs)[6]、乳酸林格氏液静脉补液[7]、胰管支架置入[8]等。基于上述部分因素,已有多个研究团队开发了包含电子病历数据的PEP风险预测模型[9-15],但将MDP图像特征纳入人工智能模型来进行PEP风险预测的研究不多。

本研究通过人工智能技术将MDP图像特征和操作特征与其他电子病历数据一起整合进风险预测过程以构建多模态模型,并通过多步骤验证过程评估各因素相对贡献。

资料与方法

一、研究对象

回顾性收集武汉大学人民医院2016年1月至2022年9月行ERCP的860份患者病历作为构建模型的训练集,收集该医院同时段行ERCP的367份患者病历作为内部测试集,回顾性收集2021年2月至2022年1月南京大学医学院附属鼓楼医院行ERCP的86份患者病历和同济大学第十人民医院行ERCP的40份患者病历,共126份病历作为外部测试集进行模型验证,同时收集每份病历多张MDP图像[4]。

病历纳入标准:(1)患者年龄>18岁;(2)十二

指肠主乳头仍保留原始形态；(3)在插管前，使用十二指肠镜对焦主乳头或主乳头附近采图。病历排除标准：(1)患者数据缺失；(2)急性胰腺炎活动期或ERCP前2周内有急性胰腺炎(排除急性胆源性胰腺炎)；(3)ERCP失败。

PEP诊断标准：ERCP后新发的腹痛或原有腹痛加重，且ERCP后24 h血淀粉酶或脂肪酶是正常值的3倍以上[16]。

本研究为基于既往多中心临床研究队列数据的二次分析，原研究已获得各参与中心伦理委员会审批。由于本研究为回顾性数据分析，伦理委员会豁免了患者知情同意。本研究分析经武汉大学人民医院伦理委员会批准，伦理号：WDRY2019K073(X05)。

二、研究方法

1.PEP相关因素的确定：通过文献研究并参考国际疾病分类第11次修订版(international classification of diseases-11th revision，ICD-11)[17]，本研究确定了49个电子病历特征和MDP特征，随后将其分为五个不同的部分：(1)基线特征，包括人口学信息和ERCP前的实验室检查结果，共8个；(2)诊断特征，在治疗开始之前内镜医师参考ICD-11做出的诊断，共22个；(3)操作特征，如内镜医师经验、困难插管和是否行经内镜乳头括约肌切开术(endoscopic sphincterotomy，EST)，共16个；(4)预防策略，由内镜专家根据ERCP相关治疗方案采取的措施，包括放置胰腺支架、使用NSAIDs和术后静脉补液，共3个；(5)MDP特征。

2.MDP图像的特征提取：在训练深层神经网络之前对图像进行预处理，提取出位于图像中心的十二指肠镜白光成像区域，对图像采取对比度归一化处理。随后使用预处理的MDP图像和预训练的ResNet-50深度神经网络模型[18]。利用该模型提取MDP特征，为每份病历生成仅基于图像的胰腺炎风险特征参数。

3.模型的开发和评估：按先后顺序添加49个电子病历特征中的4个部分至训练集，构建出单模态模型1~4，即模型1仅含基线特征，添加诊断特征至模型1构建出模型2，添加操作特征至模型2构建出模型3，添加预防策略至模型3构建出模型4。然后将MDP的特征参数分别添加到单模态模型1~4中构建出相应的多模态模型5~8。在内、外部测试集中评估各模型的表现。模型1~4、5~8表现的逐步提升为各类别特征对PEP预测的贡献，模型5~8相对于模型1~4表现的提升为MDP对PEP预测的贡献。

三、统计分析

应用SPSS 27.0软件进行统计学分析。符合正态分布的定量数据用$\bar{x}\pm s$描述，组间比较采用t检验。分类数据用例(%)描述，组件比较使用卡方检验。计算各模型预测PEP的准确率、灵敏度、特异度、受试者工作特征曲线下面积(area under the receiver operating characteristic curve，AUC)、阳性预测值(positive predictive value，PPV)和阴性预测值(negative predictive value，NPV)。使用卡方检验比较不同模型的准确率、灵敏度、特异度、PPV和NPV。P值均为双侧，P<0.05为差异有统计学意义。

结　果

一、内、外部测试集患者的基线特征和PEP发病率

在内部测试集和外部测试集患者的总体年龄、女性占比和PEP发生率差异均无统计学意义(表1)。

二、在内部测试集中多模态模型与相应单模态模型的比较

在内部测试集中，多模态模型与相应单模态模型相比，准确率、灵敏度、特异度、AUC、PPV、NPV均有所提高(表2~5)，模型8在所有模型中表现最佳，其AUC为0.91 (95%*CI*：0.86~0.96)。

三、在内部测试集中相邻多模态模型的比较

在内部测试集中，随着特征类别的增加，多模

表1　内部和外部测试集患者的基线特征和PEP发生率比较

项目	病历数	十二指肠图像(张)	年龄(岁，$\bar{x}\pm s$)	女性[例(%)]	术后胰腺炎[例(%)]
内部测试集	367	1 541	61.05±15.28	192(52.32)	28(7.63)
外部测试集	126	480	62.83±14.54	66(52.38)	13(10.32)
统计量			t=−1.14	χ^2=0.00	χ^2=0.89
P值			0.254	0.990	0.346

注：PEP指内镜逆行胰胆管造影术后胰腺炎

先行出版

表2　在内部测试集中模型1和模型5的预测效能比较

模型(特征)	准确率	灵敏度	特异度	AUC (95%*CI*)	阳性预测值	阴性预测值
模型1(基线)	61.58%(226/367)	64.29%(18/28)	61.36%(208/339)	0.67(0.58~0.76)	12.08%(18/149)	95.41%(208/218)
模型5(MDP+基线)	84.74%(311/367)	89.29%(25/28)	84.37%(286/339)	0.79(0.72~0.86)	32.05%(25/78)	98.96%(286/289)
χ^2值	63.94	4.91	45.38		13.30	6.27
*P*值	<0.001	0.027	<0.001		<0.001	0.012

注:MDP指十二指肠乳头的宏观形态;AUC受试者工作特征曲线下面积

表3　在内部测试集中模型2和模型6的预测效能比较

模型(特征)	准确率	灵敏度	特异度	AUC (95%*CI*)	阳性预测值	阴性预测值
模型2(基线+诊断)	73.57%(270/367)	75.00%(21/28)	73.45%(249/339)	0.76 (0.71~0.81)	18.92%(21/111)	97.27%(249/256)
模型6(MDP+基线+诊断)	86.92%(319/367)	92.86%(26/28)	86.43%(293/339)	0.84 (0.80~0.88)	36.11%(26/72)	99.32%(293/295)
χ^2值	20.64	3.31	17.81		6.76	3.61
*P*值	<0.001	0.069	<0.001		0.009	0.058

注:MDP指十二指肠乳头的宏观形态;AUC受试者工作特征曲线下面积

表4　在内部测试集中模型3和模型7的预测效能比较

模型(特征)	准确率	灵敏度	特异度	AUC (95%*CI*)	阳性预测值	阴性预测值
模型3(基线+诊断+操作)	75.48%(277/367)	85.71%(24/28)	74.63%(253/339)	0.82 (0.73~0.91)	21.82%(24/110)	98.44%(253/257)
模型7(MDP+基线+诊断+操作)	89.92%(330/367)	92.86%(26/28)	89.68%(304/339)	0.87 (0.82~0.92)	42.62%(26/61)	99.35%(304/306)
χ^2值	26.75	0.75	26.17		8.21	1.08
*P*值	<0.001	0.388	<0.001		0.004	0.299

注:MDP指十二指肠乳头的宏观形态;AUC受试者工作特征曲线下面积

表5　在内部测试集中模型4和模型8的预测效能比较

模型(特征)	准确率	灵敏度	特异度	AUC (95%*CI*)	阳性预测值	阴性预测值
模型4(基线+诊断+操作+预防)	85.83%(315/367)	89.29%(25/28)	85.55%(290/339)	0.88 (0.80~0.96)	33.78%(25/74)	98.98%(290/293)
模型8(MDP+基线+诊断+操作+预防)	92.64%(340/367)	96.43%(27/28)	92.33%(313/339)	0.91 (0.86~0.96)	50.94%(27/53)	99.68%(313/314)
χ^2值	8.87	1.08	7.93		3.76	1.15
*P*值	0.003	0.299	0.005		0.052	0.283

注:MDP指十二指肠乳头的宏观形态;AUC受试者工作特征曲线下面积

优先出版

态模型的准确率、灵敏度、特异度、PPV、NPV均逐步提升,但相邻多模态模型之间的比较差异无统计学意义。模型5与模型6的准确率(χ^2=0.72, *P*=0.397)、灵敏度(χ^2=0.00, *P*>0.999)、特异度(χ^2=0.58, *P*=0.447)、PPV(χ^2=0.28, *P*=0.600)和NPV(χ^2=0.00, *P*=0.982)差异无统计学意义。模型6与模型7的准确率(χ^2=1.61, *P*=0.205)、灵敏度(χ^2=0.27, *P*=0.604)、特异度(χ^2=1.70, *P*=0.193)、PPV(χ^2=0.59, *P*=0.443)和NPV(χ^2=0.00, *P*=0.982)差异无统计学意义。模型7与模型8的准确率(χ^2=1.71, *P*=0.191)、灵敏度(χ^2=0.00, *P*>0.999)、特异度(χ^2=1.46, *P*=0.227)、PPV(χ^2=0.79, *P*=0.374)和NPV(χ^2=0.37, *P*=0.546)差异无统计学意义。

四、在外部测试集中多模态模型与相应单模态模型的比较

在外部测试集中,多模态模型与相应单模态模型相比,准确率、灵敏度、特异度、AUC、PPV、NPV均呈上升趋势(表6~9),模型8在所有模型中表现最佳,其AUC为0.94(95%*CI*:0.90~0.98)。

五、在外部测试集中相邻多模态模型的比较

在外部测试集中,随着特征类别的增加,多模态模型的准确率、灵敏度、特异度、PPV、NPV均逐步提升,但相邻多模态模型之间的比较差异均无统计学意义。模型5与模型6的准确率(χ^2=0.1.24, *P*=0.265)、灵敏度(χ^2=0.00, *P*>0.999)、特异度(χ^2=1.46, *P*=0.227)、PPV(χ^2=0.65, *P*=0.419)和NPV(χ^2=0.01, *P*=0.933)差异无统计学意义。模型6与模型7的准确率(χ^2=2.82, *P*=0.093)、灵敏度(χ^2=0.00, *P*>0.999)、特异度(χ^2=2.06, *P*=0.151)、PPV(χ^2=1.13, *P*=0.288)和NPV(χ^2=0.00, *P*=0.947)差异无统计学意义。模型7与模型8的准确率(χ^2=0.87, *P*=0.351)、灵敏度(χ^2=0.30, *P*>0.999)、特异度(χ^2=

表6　在外部测试集中模型1和模型5的预测效能比较

模型(特征)	准确率	灵敏度	特异度	AUC (95% *CI*)	阳性预测值	阴性预测值
模型1(基线)	68.25%(86/126)	53.85%(7/13)	69.91%(79/113)	0.71(0.64~0.78)	9.33%(7/75)	92.94%(79/85)
模型5(MDP+基线)	77.78%(98/126)	69.23%(9/13)	78.76%(89/113)	0.85(0.80~0.90)	27.27%(9/33)	95.70%(89/93)
χ^2值	2.90	0.65	0.34	2.32	4.51	0.22
*P*值	0.089	0.688	0.559	0.128	0.034	0.637

注:MDP指十二指肠乳头的宏观形态;AUC受试者工作特征曲线下面积

表7　在外部测试集中模型2和模型6的预测效能比较

模型(特征)	准确率	灵敏度	特异度	AUC(95%*CI*)	阳性预测值	阴性预测值
模型2(基线+诊断)	70.63%(89/126)	61.54%(8/13)	71.68%(81/113)	0.76(0.70~0.82)	20.00%(8/40)	94.19%(81/86)
模型6(MDP+基线+诊断)	83.33%(105/126)	76.92%(10/13)	84.96%(96/113)	0.87(0.83~0.91)	37.04%(10/27)	96.97%(96/99)
χ^2值	5.73	0.18	5.86		2.38	0.32
*P*值	0.017	0.673	0.016		0.123	0.571

注:MDP指十二指肠乳头的宏观形态;AUC受试者工作特征曲线下面积

表8　在外部测试集中模型3和模型7的预测效能比较

模型(特征)	准确率	灵敏度	特异度	AUC(95%*CI*)	阳性预测值	阴性预测值
模型3(基线+诊断+操作)	76.98%(97/126)	69.23%(9/13)	77.88%(88/113)	0.80(0.75~0.85)	26.47%(9/34)	95.65%(88/92)
模型7(MDP+基线+诊断+操作)	90.48%(114/126)	84.62%(11/13)	91.15%(103/113)	0.92(0.89~0.95)	52.38%(11/21)	98.10%(103/105)
χ^2值	8.42	0.22	7.61		3.77	0.34
*P*值	0.004	0.645	0.006		0.052	0.562

注:MDP指十二指肠乳头的宏观形态;AUC受试者工作特征曲线下面积

表9　在外部测试集中模型4和模型8的预测效能比较

模型(特征)	准确率	灵敏度	特异度	AUC(95%*CI*)	阳性预测值	阴性预测值
模型4(基线+诊断+操作+预防)	87.30%(110/126)	76.92%(10/13)	88.50%(100/113)	0.85(0.82~0.88)	43.48%(10/23)	97.09%(100/103)
模型8(MDP+基线+诊断+操作+预防)	93.65%(118/126)	84.62%(11/13)	94.69%(107/113)	0.94(0.90~0.98)	64.71%(11/17)	98.17%(107/109)
χ^2值	2.95	0.00	2.82		1.77	0.27
*P*值	0.086	>0.999	0.093		0.184	0.605

注:MDP指十二指肠乳头的宏观形态;AUC受试者工作特征曲线下面积

1.08,*P*=0.300)、PPV(χ^2=0.59,*P*=0.444)和NPV(χ^2=0.00,*P*>0.999)差异无统计学意义。

六、特征贡献分析

对模型8而言,权重最高的前5个特征是MDP、插管困难、使用NSAIDs、既往胰腺炎和针刀括约肌预切开术(图1)。此外,在所有多模态模型中,MDP的权重均是最高的。

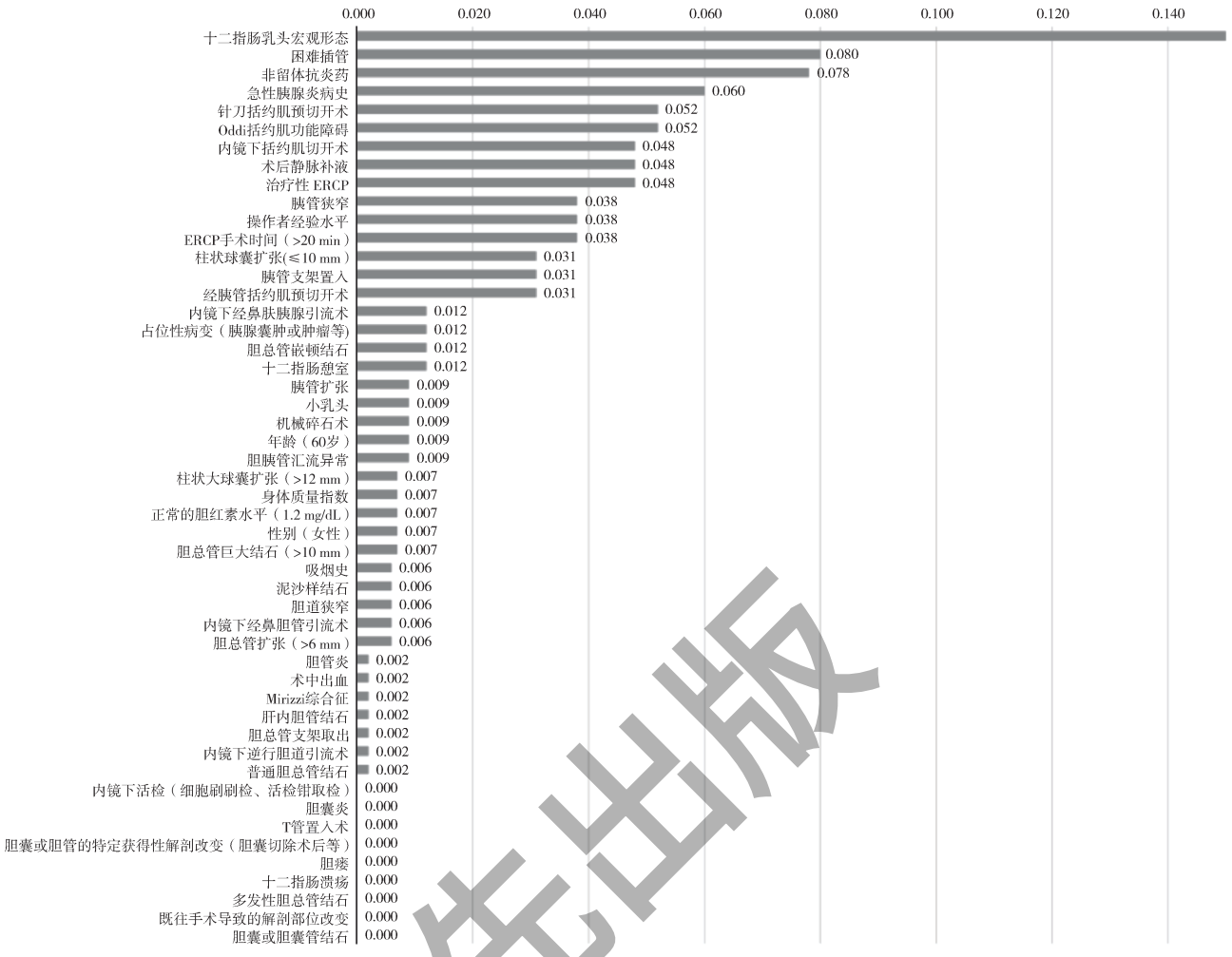
讨　论

本研究通过按步骤添加基线、诊断、操作和预防策略4个不同类别的风险因素构建单模态模型1~4,再将MDP图像的特征输入单模态模型以构建多模态模型5~8,在内、外部测试集中验证了模型表现。多模态模型与相应单模态模型的比较以及模型8的特征权重排名表明ERCP手术过程中的MDP特征对PEP的预测可能具有重要作用。

ERCP是胆胰管疾病的主要治疗方法,PEP是其最常见和最严重的并发症[2]。在精准医疗时代,医师需及时获得准确的PEP风险评估以掌握病情进展方向,规避不合理的诊疗,加强医患沟通。不少团队利用已知的风险因素构建了PEP预测模型,但其表现有待提高,潜在原因可能是模型训练时纳入的特征不够全面,如未考虑MDP特征[9-10, 14]。鉴于此,本研究尝试纳入该特征以开发更准确的多模态模型,同时探索了模型8中不同特征的贡献排名。

训练模型时纳入足够多的特征,尤其是纳入与PEP有显著相关性的特征是模型表现满意的先决条件。本研究确定了50个特征,先后构建了8个模型,随着特征数量的增加,模型在内、外部测试集中的表现均逐步提升,多模态模型8在内外部测试集

优先出版



注：ERCP指经内镜逆行胰胆管造影术

图1　模型8预测内镜逆行胰胆管造影术后胰腺炎50个特征的权重分析

中的准确率高达92.64%和93.65%。此外，模型8的表现和特征权重排名表明，本研究在构建预测PEP的人工智能模型时未遗漏具有重大临床意义的特征，未来开发PEP模型时亦需考虑MDP图像和包含困难插管在内的内镜医师的操作因素。

ERCP术者关注到MDP的不同，并猜测某些类型的MDP胆管插管更具挑战性，更易发生PEP[19]。为此，多个MDP分型方法被提出，Haraldsso是认可度、使用率最高的乳头分型方法，他将乳头分为规则型（Ⅰ型）、小乳头型（Ⅱ型）、突出或下垂型（Ⅲ型）和皱折或脊状型（Ⅳ型）[20]。在此基础上，Mohamed等[21]对纳入壶腹部憩室有无这一变量提出了修订版。研究表明，Ⅱ型或Ⅳ型乳头PEP发生率更高[22-23]。有团队通过内镜图像进行深度学习分析，开发了一种基于乳头形态的放射组学特征，由ResNet50模型生成的放射评分预测PEP的AUC达0.821，6个被机器学习算法识别的关键预测因子（年龄、白细胞计数、插管次数、小乳头、导管内超声检查、放射评分）拟合逻辑回归模型发现，放射评分的优势比最大，为7.79（95%*CI*：4.48~13.55）[24]，值得注意的是，放射评分是基于乳头形态生成的，但在拟合逻辑回归模型时同时纳入了放射评分和小乳头变量，两个变量间是否存在共线性，导致模型范化能力变差尚不得而知。在本研究中，MDP图像特征的重要性得到了更充分的证实。首先，4个多模态模型与相应单模态模型相比预测效能均提升，这表明在ERCP的任何阶段MDP特征的添加都是极具临床意义的。其次，在所有多模态模型中，MDP的权重都是最高的，在未过滤掉任何变量的全特征模型8中，MDP的权重为0.15。

有学者认为，MDP对PEP的影响可能是因为某些类型的乳头更有可能需要球囊扩张，损伤乳头、引起水肿，阻碍胰液的排出。Beran等[25]认为，MDP可影响胆管插管，Ⅲ型乳头更容易导致困难插管，而困难插管是PEP的危险因素。在另外两项研究中，Ⅱ型乳头需要多次尝试胆管插管，这导致

优先出版

了更高的PEP发生率[22, 26]。总之，MDP与PEP的发生高度相关。插管前，内镜医师可参考术前决策模型6的PEP风险评估，如风险较大，应尽量避免或降低出现困难插管的可能，如减少初学者的插管尝试次数或直接由经验丰富的内镜医师插管，及时行经内镜乳头括约肌切开术[27]，同时尽早制定预防方案如术前NSAIDs纳肛、胰管支架置入及积极静脉补液[28-29]。

模型7的临床价值在于对ERCP后患者进行风险分层管理，优化医疗资源配置，如低风险患者可回家观察、门诊随诊，降低院内感染，缩短住院时间，减轻医疗护理成本；高风险患者则收住院，监测生命体征、血淀粉酶，医护人员可分配更多的精力监测他们的病情变化，及时发现PEP，干预病变进展，减轻病情程度。作为最优模型，模型8能为ERCP后医师对患者的PEP风险告知提供及时有效的数据支持，使患者或家属对PEP风险有更直观的感知，帮助高风险患者对PEP的发生做好心理准备，降低医患纠纷的可能性。总之，这些模型可在ERCP的不同阶段为内镜医师的临床决策提供关键信息，从而改善患者预后，并节约医疗资源。

既往多用单因素或多因素逻辑回归来研究PEP[11, 13]，但各因素对PEP的风险影响可能是非线性的，传统的逻辑回归模型难以有效捕捉这些深层关联。因此，有学者尝试引入机器学习算法来揭示特征与PEP之间的关系。Archibugi等[9]比较了梯度提升和逻辑回归模型预测PEP的表现，发现前者显著优于后者（AUC 0.700比0.585，P=0.012）。类似地，Takahashi等[30]的研究也证实了随机森林模型优于逻辑回归模型（AUC 0.770比0.663，P=0.002），然而仅纳入10个特征构建随机森林模型，可能存在欠拟合。相比之下，本研究纳入了50个危险因素，共构建了8个预测模型，ERCP前即可利用已知的特征来预测PEP风险，并在外部数据集中证实了模型的泛化性，对临床决策、病情掌控有重大的意义，此外，MDP图像的特征权重非常大，而上述模型均未纳入该特征。

本研究的局限性有以下几点。第一，除了PEP，本研究未分析ERCP的其他并发症如出血、穿孔和感染等，这可能会影响模型在临床上的实际应用。第二，轻度和重度PEP对患者的健康影响大不相同，少数情况甚至能造成患者死亡。鉴于二者的危险因素、严重程度、治疗方案和转归存在差异，增加一个子模型对PEP高危患者的严重程度分类能提高模型的临床价值。最后，这是一项回顾性研究，需计划前瞻性多中心临床研究收集更多样本进一步分析，根据结果对模型的参数进行调整，同时借助计算机技术、人工智能算法的发展来提高模型的临床价值。

总之，本研究整合了50个特征开发了4个多模态模型，可在不同阶段评估PEP风险，这些模型能辅助临床决策、优化医疗资源配置，有望降低PEP发生率。

利益冲突　所有作者声明不存在利益冲突

作者贡献声明　杨婷：数据统计、论文撰写；许祐铭：研究设计、数据统计、数据分析、论文修改；于红刚：对文章的知识性内容做批评性审阅、经费支持

参 考 文 献

[1] Pal P, Ramchandani M. Management of ERCP complications [J]. Best Pract Res Clin Gastroenterol, 2024,69:101897. DOI: 10.1016/j.bpg.2024.101897.

[2] Akshintala VS, Kanthasamy K, Bhullar FA, et al. Incidence, severity, and mortality of post-ERCP pancreatitis: an updated systematic review and meta-analysis of 145 randomized controlled trials[J]. Gastrointest Endosc, 2023,98(1):1-6.e12. DOI: 10.1016/j.gie.2023.03.023.

[3] 吴碧芳, 刘赟鹏, 谢韵, 等. ERCP术后胰腺炎相关因素及防治进展[J]. 中华胰腺病杂志,2021,21(1):75-80. DOI: 10.3760/cma.j.cn115667-20200316-00039.

[4] Haraldsson E, Kylänpää L, Grönroos J, et al. Macroscopic appearance of the major duodenal papilla influences bile duct cannulation: a prospective multicenter study by the Scandinavian Association for Digestive Endoscopy Study Group for ERCP[J]. Gastrointest Endosc, 2019,90(6):957-963. DOI: 10.1016/j.gie.2019.07.014.

[5] Lee HJ, Cho CM, Heo J, et al. Impact of hospital volume and the experience of endoscopist on adverse events related to endoscopic retrograde cholangiopancreatography: a prospective observational study[J]. Gut Liver, 2020, 14(2): 257-264. DOI: 10.5009/gnl18537.

[6] Lyu Y, Cheng Y, Wang B, et al. What is impact of nonsteroidal anti-inflammatory drugs in the prevention of post-endoscopic retrograde cholangiopancreatography pancreatitis: a meta-analysis of randomized controlled trials[J]. BMC Gastroenterol, 2018, 18(1): 106. DOI: 10.1186/s12876-018-0837-4.

[7] Radadiya D, Devani K, Arora S, et al. Peri-procedural aggressive hydration for post endoscopic retrograde cholangiopancreatography (ERCP) pancreatitis prophylaxsis: meta-analysis of randomized controlled trials[J]. Pancreatology, 2019, 19(6): 819-827. DOI: 10.1016/j.pan.2019.07.046.

[8] Elmunzer BJ, Foster LD, Serrano J, et al. Indomethacin with or without prophylactic pancreatic stent placement to prevent pancreatitis after ERCP: a randomised non-inferiority trial[J]. Lancet, 2024, 403(10425): 450-458. DOI: 10.1016/S0140-6736(23)02356-5.

[9] Archibugi L, Ciarfaglia G, Cárdenas-Jaén K, et al. Machine

优先出版

learning for the prediction of post-ERCP pancreatitis risk: a proof-of-concept study[J]. Dig Liver Dis, 2023,55(3):387-393. DOI: 10.1016/j.dld.2022.10.005.

[10] Chiba M, Kato M, Kinoshita Y, et al. The milestone for preventing post-ERCP pancreatitis using novel simplified predictive scoring system: a propensity score analysis[J]. Surg Endosc, 2021, 35(12): 6696-6707. DOI: 10.1007/s00464-020-08173-4.

[11] Fu Z, Song J, Pi Y, et al. A risk prediction model for post-endoscopic retrograde cholangiopancreatography pancreatitis after stent insertion for malignant biliary obstruction: development and validation[J]. Dig Dis Sci, 2023, 68(4):1574-1584. DOI: 10.1007/s10620-022-07649-8.

[12] Park CH, Park SW, Yang MJ, et al. Pre- and post-procedure risk prediction models for post-endoscopic retrograde cholangiopancreatography pancreatitis[J]. Surg Endosc, 2022, 36(3):2052-2061. DOI: 10.1007/s00464-021-08491-1.

[13] Zheng R, Chen M, Wang X, et al. Development and validation of a risk prediction model and scoring system for post-endoscopic retrograde cholangiopancreatography pancreatitis[J]. Ann Transl Med, 2020, 8(20): 1299. DOI: 10.21037/atm-20-5769.

[14] Coté GA, Lynch S, Easler JJ, et al. Development and validation of a prediction model for admission after endoscopic retrograde cholangiopancreatography[J]. Clin Gastroenterol Hepatol, 2015, 13(13): 2323-2332. e1-9. DOI: 10.1016/j.cgh.2015.06.023.

[15] DiMagno MJ, Spaete JP, Ballard DD, et al. Risk models for post-endoscopic retrograde cholangiopancreatography pancreatitis (PEP): smoking and chronic liver disease are predictors of protection against PEP[J]. Pancreas, 2013,42(6): 996-1003. DOI: 10.1097/MPA.0b013e31827e95e9.

[16] ASGE Standards of Practice Committee, Chandrasekhara V, Khashab MA, et al. Adverse events associated with ERCP[J]. Gastrointest Endosc, 2017, 85(1): 32-47. DOI: 10.1016/j.gie.2016.06.051.

[17] World Health Organization. International Classification of Diseases (ICD) [Z]. [2026-05-02]. https://www. who. int/standards/classifications/classification-of-diseases.

[18] Wu Z, Yao L, Liu W, et al. Development and validation of a deep learning-based histologic diagnosis system for diagnosing colorectal sessile serrated lesions[J]. Am J Clin Pathol, 2023, 160(4):394-403. DOI: 10.1093/ajcp/aqad058.

[19] Tari E, Gagyi EB, Rancz A, et al. Morphology of the papilla can predict procedural safety and efficacy of ERCP: a systematic review and meta-analysis[J]. Sci Rep, 2024, 14(1): 7341. DOI: 10.1038/s41598-024-57758-9.

[20] Haraldsson E, Lundell L, Swahn F, et al. Endoscopic classification of the papilla of Vater. Results of an inter- and intraobserver agreement study[J]. United European Gastroenterol J, 2017, 5(4): 504-510. DOI: 10.1177/2050640616674837.

[21] Mohamed R, Lethebe BC, Gonzalez-Moreno E, et al. Morphology of the major papilla predicts ERCP procedural outcomes and adverse events[J]. Surg Endosc, 2021, 35(12): 6455-6465. DOI: 10.1007/s00464-020-08136-9.

[22] Chen PH, Tung CF, Peng YC, et al. Duodenal major papilla morphology can affect biliary cannulation and complications during ERCP, an observational study[J]. BMC Gastroenterol, 2020,20(1):310. DOI: 10.1186/s12876-020-01455-0.

[23] Balan GG, Arya M, Catinean A, et al. Anatomy of major duodenal papilla influences ercp outcomes and complication rates: a single center prospective study[J]. J Clin Med, 2020, 9(6)DOI: 10.3390/jcm9061637.

[24] Chen K, Lin H, Zhang F, et al. Duodenal papilla radiomics-based prediction model for post-ERCP pancreatitis using machine learning: a retrospective multicohort study[J]. Gastrointest Endosc, 2024,100(4):691-702.e9. DOI: 10.1016/j.gie.2024.03.031.

[25] Beran A, Abourshеid T, Ali AH, et al. Predictors of post-endoscopic retrograde cholangiopancreatography pancreatitis: a comprehensive systematic review and meta-analysis[J]. Clin Gastroenterol Hepatol, 2025, 23(11): 1905-1916. DOI: 10.1016/j.cgh.2024.11.014.

[26] 杨杰, 张锦华, 丁向萍, 等. 十二指肠乳头形态对选择性胆管插管的影响[J]. 中华消化内镜杂志,2022,39(6):453-458. DOI: 10.3760/cma.j.cn321463-20210813-00270.

[27] Lou L, Wang X, Zhang Y, et al. Prolonged cannulation time is an independent risk factor for moderate-to-severe post-endoscopic retrograde cholangiopancreatography (ERCP) pancreatitis: a large cohort study[J]. Ann Transl Med, 2023, 11(5):188. DOI: 10.21037/atm-22-4697.

[28] Buxbaum JL, Freeman M, Amateau SK, et al. American Society for Gastrointestinal Endoscopy guideline on post-ERCP pancreatitis prevention strategies: summary and recommendations[J]. Gastrointest Endosc, 2023, 97(2): 153-162. DOI: 10.1016/j.gie.2022.10.005.

[29] Luo H, Zhao L, Leung J, et al. Routine pre-procedural rectal indometacin versus selective post-procedural rectal indometacin to prevent pancreatitis in patients undergoing endoscopic retrograde cholangiopancreatography: a multicentre, single-blinded, randomised controlled trial[J]. Lancet, 2016, 387(10035): 2293-2301. DOI: 10.1016/S0140-6736(16)30310-5.

[30] Takahashi H, Ohno E, Furukawa T, et al. Artificial intelligence in a prediction model for postendoscopic retrograde cholangiopancreatography pancreatitis[J]. Dig Endosc, 2024,36(4):463-472. DOI: 10.1111/den.14622.

优先出版