

• 论著—临床研究 •

基于多维度临床数据的肺炎 AI 病原学类型判别模型

王霞¹ 赵玮² 陈征¹ 刘京铭¹ 康波² 郭伟³

[摘要] 目的:基于肺炎患者的临床资料,建立肺炎人工智能(artificial intelligence, AI)病原学类型判别模型,预测肺炎的责任病原体,帮助临床医生选择合适的抗感染治疗方案。方法:回顾性地收集了北京天坛医院急诊科与呼吸科在 2018 年 1 月—2020 年 12 月收治的 197 例肺炎患者的临床资料。选取 158 例(80%)患者资料作为建模组,构建肺炎 AI 病原学类型判别模型,39 例(20%)作为验证组,验证模型的预测效果。同时,将验证组预测结果与 20 名急诊科医师的病原学诊断结果进行对比。结果:基于多维度临床数据构建的肺炎 AI 病原学类型判别模型的病原学验证精度为 94.87%。20 名急诊科医师病原学诊断的准确率分别为 7.69%、15.38%、10.26%、10.26%、15.38%、17.95%、12.82%、10.26%、25.64%、17.95%、7.69%、5.13%、12.80%、20.51%、17.95%、7.69%、28.21%、12.82%、23.08%、15.38%,该模型的验证精度高于临床医师病原学诊断的平均准确率(94.87% vs 14.74%)。结论:借助既往肺炎患者的临床资料,本研究创建了基于多维度临床数据的肺炎 AI 病原学类型判别模型,该模型可用于早期预测肺炎患者的责任病原体,为临床医生早期制定经验性抗感染治疗方案提供参考。受限于样本量,本模型的临床价值有待进一步研究。

[关键词] 肺炎;多维度临床数据;病原学;人工智能

DOI:10.13201/j.issn.1009-5918.2024.07.003

[中图分类号] R563.1 **[文献标志码]** A

AI pathogen type discrimination model for pneumonia based on multi-dimensional clinical data etiology

WANG Xia¹ ZHAO Wei² CHEN Zheng¹ LIU Jingming¹ KANG Bo² GUO Wei³

(¹Department of Emergency, Beijing Tiantan Hospital, Capital Medical University, Beijing, 100070, China; ²National Supercomputing Center in Tianjin; ³Department of Emergency, Beijing Hospital of Traditional Chinese Medicine, Capital Medical University)

Corresponding author: GUO Wei, E-mail: guowei1010@126.com

Abstract Objective: To establish an artificial intelligence(AI) model for pathogen discrimination in pneumonia patients using their medical records. The model aims to predict the causative pathogens of pneumonia and assist clinical physicians in selecting appropriate antimicrobial treatment strategies. **Methods:** Retrospective medical records of 197 pneumonia patients admitted to the Emergency and Respiratory Departments of Beijing Tiantan Hospital from January 2018 to December 2020 were collected. Among these, data from 158 patients(80%) were selected to build the pneumonia AI pathogen discrimination model, while data from 39 patients(20%) were used for model validation. The predictive results of the validation group were also compared with the pathogen diagnoses made by twenty emergency department physicians. **Results:** The AI pathogen discrimination model, based on multi-dimensional clinical data, achieved a pathogen validation accuracy of 94.87%. In contrast, the accuracy of pathogen diagnosis by the twenty emergency department physicians ranged from 5.13% to 28.21%, with an average accuracy of 14.74%. Therefore, the model's validation accuracy outperformed the average accuracy of clinical physician pathogen diagnoses(94.87% vs 14.74%). **Conclusion:** By utilizing historical medical records of pneumonia patients, this study successfully developed an AI pathogen discrimination model based on multi-dimensional clinical data. The model shows promise in early prediction of pneumonia pathogens and provides valuable references for clinical physicians in selecting empirical antimicrobial treatment strategies. However, due to limitations in sample size, further research is warranted to explore the full clinical potential of this model.

¹首都医科大学附属北京天坛医院急诊科(北京,100070)

²国家超级计算天津中心

³首都医科大学附属北京中医医院急诊科

通信作者:郭伟, E-mail: guowei1010@126.com

Key words pneumonia; multi-dimensional clinical data; aetiology; artificial intelligence

肺炎是指肺实质的炎症,主要由感染、理化刺激和免疫损伤等因素引起。其中,感染是最常见的原因,通常被定义为由各种病原微生物引起的肺部炎症。根据患病环境不同,肺炎可分为社区获得性肺炎(community acquired pneumonia, CAP)、医院获得性肺炎(hospital acquired pneumonia, HAP)、呼吸机相关性肺炎(ventilator associated pneumonia, VAP)。肺炎的发生与致病微生物的种类、毒力、数量以及机体免疫反应密切相关^[1]。

根据2015年《全球疾病负担》报道,肺炎仍是美国第八大死亡原因,全球第四大死亡原因,低收入国家的首要死亡原因^[2]。研究表明,衰老过程会导致人体肺功能下降,人体对肺部疾病的易感性增加,因此,随着年龄的增加,肺部感染的发病率逐年增高^[3],相比年轻人,老年人因肺部感染的住院死亡率更高^[4]。随着世界人口的持续老龄化,肺部感染的健康问题变得更加严重^[3,5]。

明确肺炎的致病病原微生物,科学的选择抗感染治疗方案是肺炎治疗的关键^[6]。医生可以根据肺炎患者的病原学检查结果,为患者制定个体化的抗感染治疗方案,提高治疗效果,减少抗感染药物的滥用。然而,传统病原学检查对送检样本要求严格,检测时间较长,无法对混合感染及未知病原检测^[6],新型病原学检测方法操作复杂,价格昂贵,仍无法实现病原微生物的快速诊断^[7]。

一些经验较为丰富的医生能够根据患者的病史及检查结果初步判断出导致肺炎的病原微生物,然而,由于不同医生的年资不同,二级专业不同,判断病原微生物的能力存在差异,错误的治疗方案可能会延长患者住院时间,增加死亡率,并加剧病原微生物的耐药性^[8-9]。

人工智能(artificial intelligence, AI)技术在肺炎领域取得了显著进展。以深度学习为代表的AI算法在区分肺炎患者胸片与健康人胸片以及细菌和病毒性肺炎方面表现出较高的准确性^[10-11]。

本研究旨在利用肺炎患者入院后能快速获取的多维度临床信息,包括病史、实验室检查和影像学检查等结果,探索一种简单、易行的早期病原学诊断方法,拟开发基于多维度临床数据的肺炎AI病原学类型判别模型,辅助临床医师早期判断肺炎感染的责任病原体,为制定抗感染治疗方案提供初步的决策支持。

1 资料与方法

1.1 临床资料

本研究为单中心回顾性研究。研究对象为北京天坛医院2018年1月—2020年12月急诊科与呼吸科入院时诊断为肺炎的患者。本研究经首都

医科大学附属北京天坛医院伦理委员会批准(No: KY2021-136-01)。

入选标准:①入院时第一诊断为肺炎(肺部感染);②入院24h内已完善血常规,C反应蛋白(C-reactive protein, CRP),降钙素原(procalcitonin, PCT),胸部CT检查;③入院后完成两次及以上病原学检测;④出院记录示患者出院转归为“好转”,复查胸部CT报告提示肺炎好转。

排除标准:①有外院就诊经历,且应用过抗感染药物;②入院后病原学检测结果为阴性,或入院后两次及以上病原学检测结果不一致;③患有血液系统疾病,如再生障碍性贫血、白血病等;④患有免疫缺陷性疾病,如艾滋病;⑤肿瘤患者正在放化疗期间;⑥病历资料不完整。

1.2 研究方法

1.2.1 资料收集 本研究使用北京天坛医院在2018年1月—2020年12月急诊科与呼吸科入院时诊断为肺炎患者的电子病历资料,这些资料由首都医科大学附属北京天坛医院数据中心提供。我们收集了患者的年龄、性别、病史(包括症状如咳嗽、乏力、胸闷/憋气/喘憋、呼吸困难、发热等)、入院24h内的血常规指标(白细胞计数、中性粒细胞计数)、CRP、PCT、胸部CT结果(包括肺部感染情况,如双肺感染、单肺多叶感染、单肺单叶感染等),以及入院一周内两次及以上的病原学检查结果。我们将收集到的病历资料分为两组,其中80%的数据作为建模组,用于构建多维度临床数据的肺炎AI病原学判别模型。剩下的20%数据作为验证组,用于验证模型的效果。

1.2.2 模型构架 基于机器学习和集成学习的肺炎文本分类模型包括基学习器和stacking集成学习器两部分。基学习器包括K-近邻算法(K-nearest neighbor, KNN)、随机森林算法(random forest, RF)、逻辑回归算法(logistic regression, LR)、朴素贝叶斯算法(naive bayes algorithm, NB)、支持向量机算法(support vector machine, SVM)等,用于初级文本分类任务。

stacking集成学习器使用双层学习器,第一层为多个基分类器,其预测结果作为第二层学习器的元特征输入。第二层分类器通过元回归器或元分类器进行模型融合,提升了模型的鲁棒性和有效性。相较于其他集成策略,stacking采用分类器进行模型融合,具有更强的非线性表述能力和降低泛化误差的优势。stacking集成学习器的整体流程见图1。

1.2.3 建模框架 为了提升模型预测精度,肺炎文本分类模型采用2种集成策略:基于不同基学习

器的肺炎文本分类模型和基于不同参数的同一基学习器的肺炎文本分类模型。前者使用不同基学习器获得 stacking 转换输出,后者使用不同参数的相同基学习器获取 stacking 输出,例如不同聚类群数的 KNN 近邻学习器、不同预测器个数的随机森

林学习器等;两者得到的 stacking 输出均作为元特征输入用于训练二级元分类器。

整体模型分为 4 部分,其主要包括文本预处理,文本特征挖掘,基学习器分类,元学习器集成,其模型框架图具体如图 2 所示。

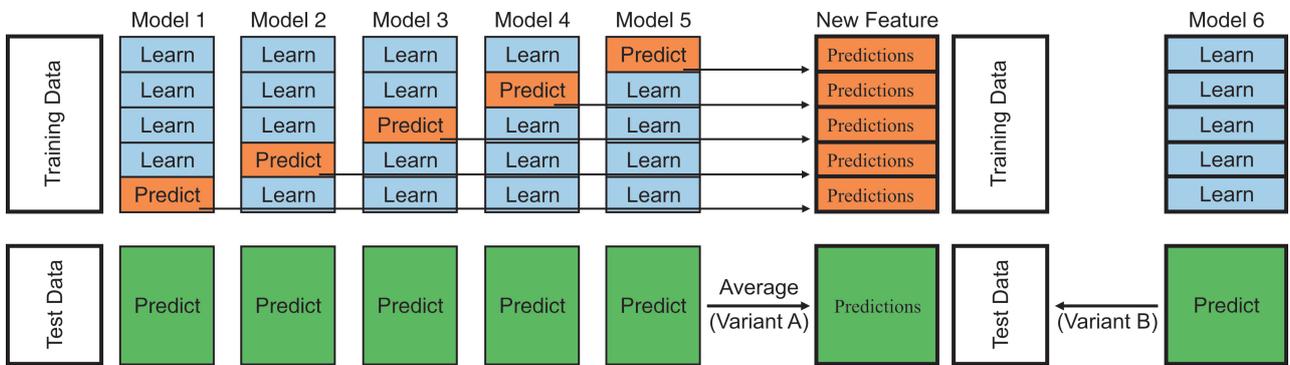


图 1 stacking 集成学习

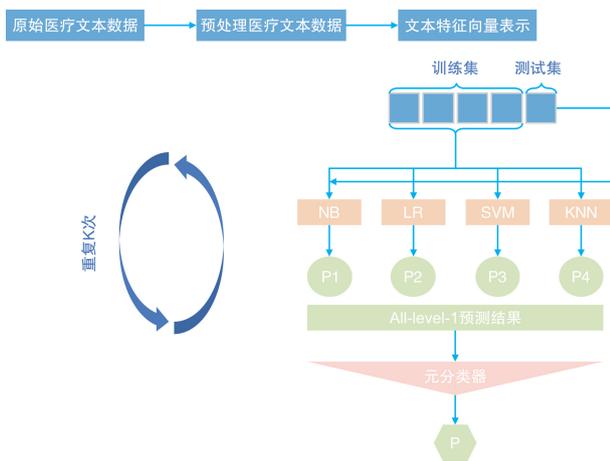


图 2 基于机器学习和集成学习的肺炎文本分类模型

首先,原始医疗文本数据可依据文本长度和文本类型分为 3 类:短文本关键词(咳嗽,发热等),长文本信息(胸部 CT 检查等)和实验室检查信息(白细胞绝对值、中性粒细胞绝对值、PCT、CRP 等)。针对长文本信息,进行文本预处理,即经过去除英文、无用符号和空格、文本去重等处理,随后借助 jieba 分词库将长文本信息切割成短分词信息;针对短文本关键词,原始数据为二进制表示,当对应关键词属性等于 1 时,将该关键词并入到该样例文本信息中,否则则剔除。针对数据信息而言,将原有连续数据离散化处理,其阈值依据医院规定正常浓度范围进行设定,其中白细胞正常值范围为 $3.5 \times 10^9/L \sim 9.5 \times 10^9/L$,中性粒细胞正常值范围为 $1.8 \times 10^9/L \sim 6.3 \times 10^9/L$,PCT 正常值 $< 0.5 \text{ ng/mL}$,CRP 正常值 $< 10 \text{ mg/L}$ 。

其次,通过 TF-IDF 或者 word2vec 将原有的长病症文本信息转换成长度统一且语义丰富的低

维词向量。针对词嵌入提取特征的方式,使用 word2vec 提取文本特征向量补零后维度为 1×3840 ,随后通过 PCA 降维设置保留信息量比例 $n_components=0.95$,最终得到长度统一且语义丰富的低维语义词向量,其维度为 1×35 ;针对文本加权提取特征的方式,使用 TF-IDF 可直接提取相对重要的语义特征向量,其维度为 1×43 。

随后将低维语义词向量输入到 KNN、RF、SVM、NB 在内的多个基学习器中进行逐一训练获得较优参数的预训练器。针对每个基学习器,病症文本数据训练集和测试集按照比例 4 : 1 被随机切分,并通过五重交叉检验得到较优参数,在不同基模型训练的结果分别为 KNN 最优近邻数,RF 最佳估计器个数,其余模型无超参设置。

最后,通过 stacking 集成学习构建并结合多个预训练基学习器完成模型融合和文本预测。其中,本文采用的基学习器可以按照 stacking 策略不同分为两类:不同种类的基学习器集成和相同种类不同参数的基学习器进行集成。其集成方式具体见图 2。

1.3 统计学方法

统计学分析均使用 SPSS 26.0 统计软件处理。连续变量使用 Shapiro-Wilk 法(W 检验)方法进行正态性检验,正态分布连续变量以 $\bar{X} \pm S$ 表示,用 t 检验进行差异性分析,偏态分布的连续变量以 $M (P_{25}, P_{75})$ 表示,用 Mann-Whitney 检验。分类变量以例(%)表示,用 χ^2 检验或 Fisher 精确检验。以病原学培养结果作为金标准,对验证组肺炎患者资料进行病原学判别模型判断,同时将验证组肺炎患者资料交予 10 名临床医师进行病原学诊断,将 20 名医师诊断结果分别与模型判断结果进行对

比,通过 χ^2 检验或Fisher检验确定结果是否存在统计学差异,评价该模型的应用价值,以 $P < 0.05$ 为差异有统计学意义。

2 结果

北京天坛医院879例肺炎患者的病历资料中,171例患者的化验检查结果、胸部CT检查结果以及痰培养结果信息有部分缺失;108例患者在住院前有外院就诊经历,应用过抗感染治疗药物;82例患者入院1周内未行病原学检查或病原学检查仅做过1次,或两次及以上痰培养结果不一致;267例患者入院1周内2次及以上痰培养结果为阴性;41例患者为接受化疗的肿瘤患者;13例患者同时患有血液系统疾病,以上患者均被排除,最终剩余的197例肺炎患者资料纳入本研究,其中男114例(57.9%),女83例(42.1%);平均年龄为53.8岁。支原体肺炎22例(11.2%),军团菌肺炎11例(5.6%),肺炎链球菌肺炎13例(6.6%),大肠埃希菌肺炎10例(5.1%),铜绿假单胞菌肺炎23例(11.7%),肺炎克雷伯菌肺炎30例(15.2%),流感嗜血杆菌肺炎8例(4.1%),鲍曼不动杆菌肺炎38例(19.3%),金黄色葡萄球菌肺炎28例(14.2%),嗜麦芽窄食单胞菌肺炎14例(7.1%)。患者的人口学资料及临床基本特征见表1,影像学资料见表2。

将所有患者资料分为建模组和验证组,建模组肺炎患者病历资料共158例(80%),非典型病原体相关肺炎26例(16.5%),其中支原体肺炎18例(11.4%),军团菌肺炎8例(5.1%);革兰阳性菌(G^+)相关肺炎34例(21.5%),肺炎链球菌肺炎11例(7.0%),金黄色葡萄球菌肺炎23例(14.6%);革兰阴性菌(G^-)相关肺炎98例(62.0%),其中大肠埃希菌肺炎8例(5.1%),铜绿假单胞菌肺炎16例(10.1%),肺炎克雷伯菌肺炎23例(14.6%),流感嗜血杆菌肺炎7例(4.4%),鲍曼不动杆菌肺炎32例(20.3%),嗜麦芽窄食单胞菌肺炎12例(7.6%)。用于多维度临床数据的肺炎AI病原学类型判别模型的建模,建模完成后用建模组肺炎患者的临床数据进行内部验证,非典型病原体组模型输出病原结果未见错误; G^+ 组中,肺炎链球菌肺炎模型输出结果未见错误,金黄色葡萄球菌肺炎验证错误1例,模型错误输出结果为鲍曼不动杆菌; G^- 组中,大肠埃希菌肺炎、流感嗜血杆菌肺炎及嗜麦芽窄食单胞菌肺炎内部验证病原输出结果未见错误,铜绿假单胞菌肺炎验证错误1例,模型错误输出结果为肺炎克雷伯菌,肺炎克雷伯菌肺炎验证错误1例,模型错误输出结果为鲍曼不动杆菌,鲍曼不动杆菌肺炎验证错误2例,模型错误输出结果均为肺炎克雷伯菌。建模组模型内部验证训练精度为96.83%,具体输出结果见表3。

表1 建模组与验证组肺炎患者的临床特征比较

临床资料	建模组 (158例)	验证组 (39例)	P
年龄/岁	52(42.39~ 56.42)	66(54.89~ 68.34)	0.060
性别/例(%)			0.820
男	90(56.96)	23(58.87)	
女	68(43.04)	16(41.03)	
病史/例(%)			
受凉/淋雨	14(8.86)	6(15.38)	0.227
咳嗽	120(75.95)	37(94.87)	0.109
咳痰	103(65.19)	36(96.31)	0.101
乏力	4(2.53)	10(25.64)	0.071
胸闷/憋气	44(27.85)	25(64.10)	0.063
呼吸困难	12(7.59)	4(10.26)	0.393
发热	65(41.14)	22(56.41)	0.085
寒战	11(6.96)	3(7.69)	0.550
头痛	4(2.53)	5(12.82)	0.016
咽痛	6(3.80)	2(5.13)	0.494
肌痛	8(5.06)	0	0.165
化验检查			
白细胞/ $(\times 10^9/L)$	12.36(4.38)	12.09(5.00)	0.015
中性粒细胞/ $(\times 10^9/L)$	8.1(3.74)	9.66(4.89)	0.216
PCT/(ng/mL)	0.32(0.62~ 2.13)	0.31(0.51~ 3.63)	0.489
CRP/(mg/L)	90.44(79.14)	110.18(89.24)	0.392
痰细菌学检测结果/例(%)			
支原体	18(11.39)	4(10.26)	0.551
军团菌	8(5.06)	3(7.69)	0.376
肺炎链球菌	11(6.96)	2(5.13)	0.506
大肠埃希菌	8(5.06)	2(5.13)	0.624
铜绿假单胞菌	16(10.13)	7(17.94)	0.140
肺炎克雷伯菌	23(14.56)	7(17.94)	0.598
流感嗜血杆菌	7(4.43)	1(2.56)	0.506
鲍曼不动杆菌	32(20.25)	6(15.38)	0.490
金黄色葡萄球菌	23(14.56)	5(12.82)	0.781
嗜麦芽窄食单胞菌	12(7.59)	2(5.13)	0.450

验证组肺炎患者病历资料共39例(20.0%),非典型病原体相关肺炎7例(18.0%),其中支原体肺炎4例(10.3%),军团菌肺炎3例(7.7%); G^+ 相关肺炎7例(18.0%),其中肺炎链球菌肺炎2例(5.1%),金黄色葡萄球菌肺炎5例(12.8%); G^- 相关肺炎25例(64.0%),其中大肠埃希菌肺炎2例(5.1%),铜绿假单胞菌肺炎7例(17.9%),肺炎克雷伯菌肺炎7例(17.9%),流感嗜血杆菌肺炎1例(2.6%),鲍曼不动杆菌肺炎6例(15.4%),嗜麦

芽窄食单胞菌肺炎 2 例(5.1%),用于模型的再次验证。同时将验证组患者资料交予 20 名临床医师进行病原学诊断,最后将多维度临床数据的肺炎 AI 病原学类型判别模型的判别结果分别与 20 名医师病原学诊断结果进行对比。经验证组验证可见,非典型病原体组及 G⁺ 组模型输出病原结果未见错误,仅 G⁻ 组中流感嗜血杆菌肺炎及鲍曼不动杆菌肺炎各验证错误 1 例,流感嗜血杆菌肺炎模型错误输出结果为大肠埃希菌,鲍曼不动杆菌肺炎模型错误输出结果为肺炎克雷伯菌。模型的病原学验证精度为 94.87%,具体输出结果见表 4。

表 2 建模组与验证组肺炎患者的影像学特征比较例(%)

影像学特征	建模组 (158 例)	验证组 (39 例)	P
双肺感染	68(43.04)	21(53.85)	0.225
磨玻璃影	16(23.53)	3(14.29)	0.282
斑片影	8(11.76)	5(23.81)	0.172
炎性改变	44(64.71)	13(61.90)	0.815
合并支扩	36(52.94)	3(14.29)	0.002
合并肺气肿	24(79.41)	11(52.38)	0.161
合并肺大疱	19(27.94)	4(19.05)	0.306
合并气胸	5(7.35)	1(0)	0.565
合并胸腔积液	32(47.06)	8(38.10)	0.470
合并支气管壁增厚	18(26.47)	6(28.57)	0.850
单肺感染	90(56.96)	18(46.15)	0.225
单肺多叶	54(60.00)	11(61.11)	0.930
磨玻璃影	10(18.52)	1(9.09)	0.401
斑片影	6(11.11)	1(9.09)	0.662
炎性改变	38(70.37)	9(81.82)	0.357
合并支扩	12(22.22)	2(18.18)	0.562
合并肺气肿	17(31.48)	1(9.09)	0.124
合并肺大疱	7(12.96)	0	0.254
合并气胸	3(5.56)	0	0.568
合并胸腔积液	19(35.19)	3(27.27)	0.449
合并支气管壁增厚	24(44.44)	4(36.36)	0.441
单肺单叶	36(40.00)	7(38.89)	0.930
磨玻璃影	8(22.22)	2(28.57)	0.524
斑片影	2(5.56)	0	0.698
炎性改变	26(72.22)	5(71.43)	0.644
合并支扩	3(8.33)	0	0.579
合并肺气肿	8(22.22)	0	0.209
合并肺大疱	2(5.56)	0	0.698
合并气胸	1(0)	0	0.837
合并胸腔积液	2(5.56)	0	0.698
合并支气管壁增厚	15(41.67)	4(57.14)	0.365

该 AI 模型肺炎病原学判别结果的验证精度为 94.87%,20 名医师诊断的准确率分别为 7.69%、15.38%、10.26%、10.26%、15.38%、17.95%、12.82%、10.26%、25.64%、17.95%、7.69%、5.13%、12.80%、20.51%、17.95%、7.69%、28.21%、12.82%、23.08%、15.38%,平均准确率为 14.74%,与临床医师的准确率相比,模型诊断的准确率明显较高。将 AI 模型病原学判别结果分别与 20 名医师病原学诊断结果进行 Fisher 检验,结果具有统计学意义(均 P<0.01),模型的判断准确性高于医师的判断准确性,见表 5。

表 3 建模组内部验证结果

肺炎病原体	金标准 检测 个数	模型输出 结果/个		准确率 /%	错误输出 病原类型
		正确	错误		
支原体	18	18	0	100.00	—
军团菌	8	8	0	100.00	—
肺炎链球菌	11	11	0	100.00	—
金黄色葡萄球菌	23	22	1	95.65	鲍曼不动杆菌
大肠埃希菌	8	8	0	100.00	—
铜绿假单胞菌	16	15	1	93.75	肺炎克雷伯菌
肺炎克雷伯菌	23	22	1	95.65	鲍曼不动杆菌
流感嗜血杆菌	7	7	0	100.00	—
鲍曼不动杆菌	32	30	2	93.75	肺炎克雷伯菌×2
嗜麦芽窄食单胞菌	12	12	0	100.00	—
总体	158	153	5	96.83	—

注:肺炎克雷伯菌×2 表示 2 例错误输出病原类型均为肺炎克雷伯菌。

表 4 验证组验证结果

肺炎病原体	金标准 检测 个数	模型输出 结果/个		准确率 /%	错误输出 病原类型
		正确	错误		
支原体	4	4	0	100.00	—
军团菌	3	3	0	100.00	—
肺炎链球菌	2	2	0	100.00	—
金黄色葡萄球菌	5	5	0	100.00	—
大肠埃希菌	2	2	0	100.00	—
铜绿假单胞菌	7	7	0	100.00	—
肺炎克雷伯菌	7	7	0	100.00	—
流感嗜血杆菌	1	0	1	0	大肠埃希菌
鲍曼不动杆菌	6	5	1	83.33	肺炎克雷伯菌
嗜麦芽窄食单胞菌	2	2	0	100.00	—
总体	39	37	2	94.87	—

表 5 AI 病原学判别模型判别结果与 20 名医师诊断结果

分组	诊断 正确/个	诊断 错误/个	正确率/%	P
AI 模型组	37	2	94.87	
医师 1	3	36	7.69	<0.01
医师 2	6	33	15.38	<0.01
医师 3	4	35	10.26	<0.01
医师 4	4	35	10.26	<0.01
医师 5	6	33	15.38	<0.01
医师 6	7	32	17.95	<0.01
医师 7	5	34	12.82	<0.01
医师 8	4	35	10.26	<0.01
医师 9	10	29	25.64	<0.01
医师 10	7	32	17.95	<0.01
医师 11	3	36	7.69	<0.01
医师 12	2	37	5.13	<0.01
医师 13	5	34	12.82	<0.01
医师 14	8	31	20.51	<0.01
医师 15	7	32	17.95	<0.01
医师 16	3	36	7.69	<0.01
医师 17	11	28	28.21	<0.01
医师 18	5	34	12.82	<0.01
医师 19	9	30	23.08	<0.01
医师 20	6	33	15.38	<0.01

3 讨论

本研究旨在利用多维度临床数据构建肺炎 AI 病原学类型判别模型,通过分析患者的病历资料、实验室检查和影像学检查结果等信息,实现对肺炎病原体的早期预测和诊断,为临床医生提供治疗方案的参考。

既往的研究表明,AI 技术在肺炎诊治方面已经取得了可观的成果。早在 2003 年,Heckerling 等^[12]利用神经网络算法预测肺炎的诊断,显示出较高的区分精度。2018 年,Kermany 等^[13]使用深度学习算法成功预测小儿肺炎的诊断,并实现了较高的准确率和灵敏度。此外,2019 年 Hwang 等^[14]开发了对影像进行自动分类的系统,用于肺部病灶性质的判断和肺炎的预测,经过外部验证,按图像分类的 ROC 曲线下面积达 0.965(95%CI:0.955~0.975)。在肺炎病原学研究领域,2019 年潘丽艳等^[10]应用 AI 系统基于肺炎影像学表现构建儿童肺炎病原学类型自动判别模型,基于深度卷积神经网络的肺炎病原学类型二分类判别模型的准确率达 80.48%,特异度为 82.07%,灵敏度为 77.55%,AUC 达 0.82,成果显著。

然而,现有的 AI 研究大多聚焦于影像学检

查^[10,15],本研究在前期学者研究的基础上,利用多维度的患者临床资料,深入探索了肺炎致病病原体与 AI 的联系,成功地建立了基于多维度临床数据的肺炎 AI 病原学类型判别模型,并对其在建模组和验证组的预测效果进行了评估。在建模组中,我们的模型展现出了 96.83% 的训练精度,而在验证组中,预测模型的病原学类型验证精度也达到了 94.87%。这些结果表明,我们的 AI 模型在肺炎病原体的预测方面表现出了很高的准确性和稳定性。此外,本研究所用病历数据避免了抗感染药物的应用对于痰培养结果的影响。原始数据的准确性保证了模型判断肺炎致病菌的敏感性,排除标准保证了本研究数据的可靠性,降低了数据本身对建模精度的影响,提高了临床数据的应用价值,增强了模型预测结果的可信性。

目前尚没有将肺炎致病菌种类 AI 识别模型与临床医师病原学种类判断结果进行对比的报道。本研究将模型的预测结果与 20 名急诊科医师的病原学诊断结果进行对比。结果显示,模型的验证精度高于医师的平均准确率(94.87% vs 14.74%)。这表明该 AI 模型在病原学类型判别方面具备潜在的优势,有望成为临床医生的有力辅助工具,提高肺炎病原体早期预测的准确率,从而改善患者的治疗效果和预后。

对于验证组 39 例肺炎患者的病历资料,20 名临床医师判断具体致病病原体的平均准确率为 14.74%,临床医师判断致病菌的准确率较低,我们考虑原因如下:对于新入院的肺炎患者,根据患者的症状及表现,医师首先会考虑该患者是否为感染,结合检查结果,进一步考虑致病病原微生物的种类,如病毒、非典型病原体、G⁺ 菌及 G⁻ 菌,经初步判断,选择覆盖该病原种类的抗感染药物,鲜有临床医师会直接判断具体致病病原微生物,缩小抗感染药物的选择范围。

利用患者多维度的临床信息,我们成功研发了基于多维度临床数据的肺炎 AI 病原学类型判别模型,模型输出结果可以具体到某一种病原微生物,相较于既往的研究,输出结果更加精确。另外,我们的模型基于北京天坛医院的病历数据,应用时也仅涉及肺炎患者已完成的常规检查结果,在后期应用于临床的过程中,避免了患者额外的支出,对患者不会造成任何创伤,具有很大的经济适用性。

本研究也存在一定的局限性。首先,由于样本量的限制,我们的模型可能对某些病原体类型的判别效果较为有限。因此,未来需要进一步扩大样本规模,包含更多不同类型的肺炎患者数据,以验证和改进模型的稳定性和准确性。其次,本研究仅利用了单中心的数据,这可能受到地域和人群特性的影响。为使模型具有更广泛的适用性,需要考虑跨

多个医院和不同地区的数据进行验证。

总体而言,本研究的肺炎 AI 病原学类型判别模型显示出了潜在的临床应用价值。通过利用多维度临床数据进行肺炎病原体的早期预测和诊断,我们的模型有望成为临床决策中重要的辅助工具,为制定个体化的抗感染治疗方案提供有力支持。随着 AI 技术的不断进步和应用,我们对于肺炎病原学的理解和诊断能力也将逐步得到提升,从而为患者的健康和生命质量带来积极的影响。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] Scheld WM. Developments in the pathogenesis, diagnosis and treatment of nosocomial pneumonia [J]. Surg Gynecol Obstet, 1991, 172(Suppl): 42-53.
- [2] Heron M. Deaths; leading causes for 2015 [J]. Natl Vital Stat Rep, 2017, 66(5): 1-76.
- [3] Chongthanadon B, Thirawattanasoot N, Ruangsomboon O. Clinical factors associated with in-hospital mortality in elderly versus non-elderly pneumonia patients in the emergency department [J]. BMC Pulm Med, 2023, 23(1): 330.
- [4] Hespanhol V, Bárbara C. Pneumonia mortality, comorbidities matter? [J]. Pulmonology, 2020, 26(3): 123-129.
- [5] Schöll N, Rohde GGU. Ambulant erworbene Pneumonie bei älteren Menschen [Community-acquired Pneumonia in the Elderly] [J]. Pneumologie, 2019, 73(10): 605-616.
- [6] 邹晓辉, 曹彬. 呼吸道感染病原学诊断年度进展 2021 [J]. 中华结核和呼吸杂志, 2022, 45(1): 78-82.
- [7] 中华医学会检验医学分会. 高通量宏基因组测序技术检测病原微生物的临床应用规范化专家共识 [J]. 中华检验医学杂志, 2020, 43(12): 1181-1195.
- [8] Modi AR, Kovacs CS. Hospital-acquired and ventilator-associated pneumonia: Diagnosis, management, and prevention [J]. Cleve Clin J Med, 2020, 87(10): 633-639.
- [9] Fleuren LM, Klausch TLT, Zwager CL, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy [J]. Intensive Care Med, 2020, 46(3): 383-400.
- [10] 潘丽艳, 梁会营. 基于深度学习的儿童肺炎病原学类型判别模型 [J]. 中国数字医学, 2019, 14(3): 59-61, 110.
- [11] Li YY, Zhang ZY, Dai C, et al. Accuracy of deep learning for automated detection of pneumonia using chest X-Ray images: a systematic review and meta-analysis [J]. Comput Biol Med, 2020, 123: 103898.
- [12] Heckerling PS, Gerber BS, Tape TG, et al. Prediction of community-acquired pneumonia using artificial neural networks [J]. Med Decis Making, 2003, 23(2): 112-121.
- [13] Kermany DS, Goldbaum M, Cai WJ, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning [J]. Cell, 2018, 172(5): 1122-1131. e9.
- [14] Hwang EJ, Park S, Jin KN, et al. Development and validation of a deep learning-based automated detection algorithm for major thoracic diseases on chest radiographs [J]. JAMA Netw Open, 2019, 2(3): e191095.
- [15] Stephen O, Sain M, Maduh UJ, et al. An efficient deep learning approach to pneumonia classification in healthcare [J]. J Healthc Eng, 2019, 2019: 4180949.

(收稿日期: 2024-01-08)