

· 综述 ·

机器学习方法构建健康素养预测模型的范围综述

潘祥, 童莺歌, 李怡萱, 倪珂, 程雯倩, 辛蒙雨, 胡钰滢

杭州师范大学护理学院, 浙江 杭州 311121

摘要: **目的** 对应用机器学习方法构建的健康素养预测模型的种类、构建方法和预测效果进行范围综述, 为该类模型的改进和应用提供参考。**方法** 检索中国知网、万方数据知识服务平台、维普中文科技期刊数据库、PubMed和Web of Science, 收集建库至2024年5月1日发表的应用机器学习方法构建健康素养预测模型研究文献。采用预测模型偏倚风险评估工具进行文献质量评价, 对纳入文献的基本特征、模型构建方法、数据来源、缺失值处理、预测因子和预测效果等进行综述。**结果** 检索获得文献524篇, 最终纳入22篇, 发表时间为2007—2024年。涉及48个健康素养预测模型, 其中25个偏倚风险为高风险, 占52.08%, 主要问题集中在缺失值处理、预测因子选择和模型评价方法。模型构建方法包括回归模型、基于树的机器学习方法、支持向量机和神经网络模型。预测因子主要包括个人、人际关系、组织和社会/政策4个层面的因素, 年龄、文化程度、经济水平、健康状况和互联网使用的出现频率较高。14篇文献进行了模型内部验证, 4篇进行了外部验证。42个模型报告了受试者操作特征曲线下面积, 范围为0.52~0.983, 区分度良好。**结论** 应用机器学习方法构建的健康素养预测模型展现了较好的预测能力, 但研究在偏倚风险、数据处理和验证规范性等方面存在不足。

关键词: 健康素养; 预测模型; 机器学习; 范围综述

中图分类号: R193 **文献标识码:** A **文章编号:** 2096-5087 (2025) 02-0148-06

Health literacy prediction models based on machine learning methods: a scoping review

PAN Xiang, TONG Yingge, LI Yixuan, NI Ke, CHENG Wenqian, XIN Mengyu, HU Yuying

School of Nursing, Hangzhou Normal University, Hangzhou, Zhejiang 311121, China

Abstract: Objective To conduct a scoping review on the types, construction methods and predictive performance of health literacy prediction models based on machine learning methods, so as to provide the reference for the improvement and application of such models. **Methods** Publications on health literacy prediction models conducted using machine learning methods were retrieved from CNKI, Wanfang Data, VIP, PubMed and Web of Science from inception to May 1, 2024. The quality of literature was assessed using the Prediction Model Risk of Bias ASsessment Tool. Basic characteristics, modeling methods, data sources, missing value handling, predictors and predictive performance were reviewed. **Results** A total of 524 publications were retrieved, and 22 publications between 2007 and 2024 were finally enrolled. Totally 48 health literacy prediction models were involved, and 25 had a high risk of bias (52.08%), with major issues focusing on missing value handling, predictor selection and model evaluation methods. Modeling methods included regression models, tree-based machine learning methods, support vector machines and neural network models. Predictors primarily encompassed factors at four aspects: individual, interpersonal, organizational and society/policy aspects, with age, educational level, economic status, health status and internet use appearing frequently. Internal validation was conducted in 14 publications, and external validation was conducted in 4 publications. Forty-two models reported the areas under the receiver operating characteristic curve, which ranged from 0.52 to 0.983, indicating good discrimination. **Conclusion** Health literacy prediction models based on machine learning methods perform well, but have defi-

DOI: 10.19485/j.cnki.issn2096-5087.2025.02.009

基金项目: 教育部规划基金项目 (23YJAZH136)

作者简介: 潘祥, 硕士研究生在读, 护理学专业

通信作者: 童莺歌, E-mail: 1352597965@qq.com

ciencies in risk of bias, data processing and validation.

Keywords: health literacy; prediction model; machine learning; scope review

健康素养指个体获取、理解健康信息并做出健康决策的能力，是评估公众健康水平和卫生服务能力的重要指标。我国高度重视居民健康素养水平，自2012年起持续开展全国居民健康素养动态监测，以识别居民健康教育需求，促进健康公平^[1-2]。但大规模的健康素养普查会增加时间和劳动力成本。为提高居民健康素养评估效率、合理分配医疗资源，使用机器学习方法构建健康素养预测模型识别健康素养不足人群正逐渐成为研究热点^[3-4]。机器学习是一种人工智能方法，利用数据和算法自动构建统计模型，学习、识别并改进模型的性能，实现对未来事件的预测和决策。本研究收集国内外发表的应用机器学习方法构建健康素养预测模型相关文献，了解这些模型的类型、构建方法和预测效果，并评价其方法学质量，提出优化建议，为居民健康素养预测模型的改进和应用提供参考。

1 资料与方法

1.1 研究问题

(1) 健康素养的预测模型有哪些？通常涉及哪些预测因素？(2) 这些预测模型采用什么方法构建？数据来源是什么？(3) 这些预测模型是否进行了内外部验证？性能如何？

1.2 研究报告

本研究遵循范围综述报告的PRISMA扩展版清单(PRISMA extension for scoping reviews, PRISMA-ScR)^[5]进行报告。

1.3 文献检索策略

通过中国知网、万方数据知识服务平台和维普中文科技期刊数据库检索中文文献，检索词为健康素养、预测模型、风险评估、风险预测、决策模型、机器学习和回归模型；通过PubMed和Web of Science数据库检索英文文献，检索词为health literacy、artificial intelligence、machine learning、deep learning、random forest、decision tree、prediction model、logistic regression、predictive model和risk model。结合主题词与自由词检索，并使用逻辑词“OR”和“AND”进行灵活组合。检索时限为建库至2024年5月1日，并对纳入文献的参考文献追踪检索。

1.4 文献纳入和排除标准

纳入标准：(1) 研究主题为健康素养预测模型的构建或验证；(2) 预测模型采用机器学习方法构建；

(3) 描述构建、验证预测模型的具体过程。排除标准：(1) 非中、英文文献；(2) 会议摘要、综述等；(3) 无法获取全文。

1.5 文献的筛选和数据提取

检索到的文献导入EndNote 20软件，删除重复项。由2名护理学研究生根据纳入和排除标准，阅读文献标题和摘要进行初筛，再阅读全文复筛。若出现分歧，通过讨论或咨询第三人解决。由1名具备循证医学知识的健康素养领域专家与2名护理学研究生共同设计文献数据提取表，提取纳入文献的基本特征、模型类型、样本量、缺失值处理方法、机器学习方法类型、模型性能评价指标、模型验证方法和预测因子等。

1.6 方法学质量评价

由2名系统学习过循证护理知识的研究生使用预测模型偏倚风险评估工具(Prediction Model Risk Of Bias ASsessment Tool, PROBAST)进行评估。若出现分歧，通过讨论或咨询第三人解决。PROBAST包括4个领域(研究对象、预测因子、结局和分析)20个问题。每个领域的风险评估等级分为“低风险”“高风险”“不清楚”，若1个领域被评估为高风险，则总体为高风险；若至少1个领域被评估为不清楚，其他领域均为低风险，则总体为不清楚；若所有领域均为低风险但未进行外部验证，则总体为高风险；若模型是基于大样本数据构建且进行了内部验证，则总体为低风险^[6]。

2 结果

2.1 文献筛选结果

检索获得文献524篇，经筛选后最终纳入22篇，见图1。发表于2007—2013年6篇、2014—2019年7篇和2020—2024年9篇；研究地区为中国11篇、美国7篇、英国2篇、欧洲1篇和土耳其1篇。研究主题为普通居民健康素养11篇、老年人健康素养4篇、学生健康素养3篇、传染病健康素养3篇和糖尿病健康素养1篇。

2.2 偏倚风险评估结果

22篇文献涉及健康素养预测模型48个，其中偏倚风险为高风险25个，占52.08%；低风险10个，占20.83%；不清楚13个，占27.08%。偏倚风险主要来自“分析”领域，与变量处理、缺失值处理、预测因子选择方式和模型评价方法有关。见表1。

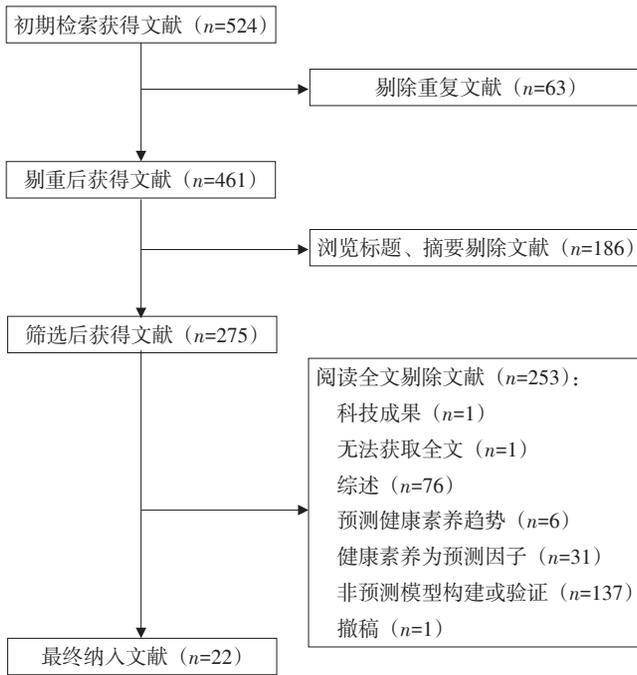


图1 文献检索流程

Figure 1 Flow chart of literature screening

2.3 健康素养预测模型的构建与验证情况

2.3.1 数据来源及缺失值处理

纳入的22篇文献中有10篇数据来源于政府组织的大型居民健康素养调查。研究样本量<1 000人7篇, 1 000~<3 000人7篇, 3 000~<10 000人5篇, ≥10 000人3篇, 平均为16 799人。报告缺失值处理仅有8篇, 其中采用多重插补法4篇, 直接删除缺失数据4篇。报告缺失数据量6篇, 平均缺失值比例为5.75%。见表2。

2.3.2 机器学习方法类型

48个预测模型中, 采用回归模型32个, 占66.67%, 主要为logistic回归模型; 采用基于树的机器学习方法10个, 占20.83%, 主要为随机森林; 采用支持向量机4个, 占8.33%; 采用神经网络模型2个, 占4.17%。见表2。

2.3.3 预测因子

本研究基于社会生态学模型将预测因子分为个人、人际关系、组织和社会/政策4个层面。个人层面包括性别、年龄、文化程度、经济水平、日常生活活动能力、居住地、情绪状态、自我效能感、互联网使用、种族/民族、睡眠问题、心理健康、慢性病共病、健康状况和阅读能力15个因子; 人际关系层面包括职业、婚姻状况、父母文化程度、父母职业状态、父母忽视行为、医务人员建议、居住方式和社会孤立8个因子; 组织层面包括年级、专业、健康教

表1 纳入文献的偏倚风险评价

Table 1 Bias risk assessment of included publications

第一作者	模型数量	研究对象	预测因子	结局	分析	总体
ZHOU ^[7]	6	高风险	低风险	低风险	不清楚	高风险
İNCEOĞLU ^[8]	3	低风险	低风险	低风险	高风险	高风险
XIE ^[9]	2	低风险	低风险	低风险	低风险	低风险
HONG ^[10]	2	低风险	低风险	低风险	低风险	高风险
季娜 ^[11]	2	低风险	低风险	低风险	不清楚	不清楚
李现文 ^[12]	2	不清楚	低风险	低风险	不清楚	不清楚
林丰 ^[13]	2	低风险	低风险	低风险	低风险	低风险
王可 ^[14]	6	不清楚	低风险	高风险	高风险	高风险
HOU ^[15]	1	不清楚	低风险	低风险	低风险	不清楚
ZHANG ^[16]	1	不清楚	低风险	低风险	不清楚	不清楚
VAN DER HEIDE ^[17]	2	低风险	低风险	低风险	低风险	低风险
LUBIMIR ^[18]	4	低风险	低风险	低风险	不清楚	不清楚
MARTIN ^[19]	2	低风险	低风险	高风险	高风险	高风险
MILLER ^[20]	1	低风险	低风险	低风险	高风险	高风险
LAURSEN ^[21]	3	低风险	低风险	低风险	低风险	低风险
HANCHATE ^[22]	1	低风险	低风险	低风险	低风险	低风险
陈思婷 ^[23]	1	低风险	低风险	高风险	不清楚	高风险
LURIE ^[24]	1	低风险	低风险	低风险	高风险	高风险
CAMPBELL ^[25]	3	低风险	低风险	低风险	不清楚	不清楚
ROWLANDS ^[26]	1	低风险	低风险	低风险	高风险	高风险
LEUNG ^[27]	1	低风险	低风险	低风险	高风险	高风险
JEPPESSEN ^[28]	1	低风险	低风险	低风险	高风险	高风险

育课选修情况(学生)、信息获取渠道、就医情况和是否从事医疗相关工作6个因子; 社会/政策层面包括医疗保险、健康支付能力、家庭医生签约情况、自费疫苗接种情况、所在地区贫困水平和移民情况6个因子。出现频率较高的预测因子为年龄、文化程度、经济水平、健康状况和互联网使用。

2.3.4 模型评价

报告受试者操作特征曲线下面积(area under the curve, AUC)的模型42个, 占7.50%, 多数模型区分度良好, AUC值范围0.52~0.983, M(Q_R)为0.78(0.21); 报告准确度的模型19个, 占39.58%; 报告灵敏度31个, 占64.58%; 报告特异度20个, 占41.67%; 报告F1分数15个, 占31.25%; 报告阳性预测值22个, 占45.83%; 报告阴性预测值10个, 占20.83%; 报告Brier分数3个, 占6.25%; 报告校准度指标4个, 占8.33%。见表2。

2.3.5 模型验证

纳入的22篇文献中, 进行模型内部验证14篇, 其中, 使用随机拆分验证法9篇, 占40.91%; 使用K折交叉验证法2篇, 占9.09%; 使用卡方自动交叉验证2篇, 占9.09%; 未报道内部验证方法1篇。

表2 健康素养预测模型构建与验证情况
Table 2 Establishment and validation of health literacy prediction models

第一作者	数据来源	样本量/ 缺失量	缺失值 处理方法	机器学习 方法类型	模型预测 效能	模型验证方法
ZHOU [7]	调查收集	1 874/—	—	logistic回归、随机森林、支持向量机	AUC=0.754、0.817、0.759、0.913、0.925、0.910	内部验证(随机拆分验证)
İNCEOĞLU [8]	调查收集	1 001/—	—	极端梯度提升、logistic回归、随机森林	AUC=0.983、0.976、0.828	内部验证(K折交叉验证)
XIE [9]	调查收集	1 235/65	直接删除	随机森林、Lasso回归	AUC=0.975、0.928	内部验证(随机拆分验证)
HONG [10]	调查收集	780/—	—	极端梯度提升、神经网络模型	R ² =0.909、0.976	—
季娜 [11]	调查收集	1 966/—	—	logistic回归、决策树	AUC=0.677、0.666	内部验证(卡方自动交叉验证)
李现文 [12]	调查收集	334/—	—	logistic回归、决策树	AUC=0.847、0.813	内部验证(卡方自动交叉验证)
林丰 [13]	广东省居民健康素养监测项目	3 855/—	—	logistic回归、神经网络模型	AUC=0.833、0.832	内部验证(随机拆分验证)
王可 [14]	调查收集	838/—	—	logistic回归、随机森林、支持向量机	AUC=0.723、0.815、0.728、0.896、0.944、0.897	内部验证(随机拆分验证)
HOU [15]	调查收集	648/—	—	logistic回归	AUC=0.710	内部验证(随机拆分验证)
ZHANG [16]	调查收集	380/—	—	logistic回归	AUC=0.848	—
VAN DER HEIDE [17]	2008年荷兰成人识字和生活技能调查	3 745/—	多重插补	logistic回归	AUC=0.740、0.580	内部验证(随机拆分验证)、外部验证
LUBIMIR [18]	2008、2010年夏威夷健康调查	5 987/—	—	logistic回归	AUC=0.689、0.624、0.684、0.647	内部验证(随机拆分验证)
MARTIN [19]	2003年美国全国成人识字评估	17 466/—	—	logistic回归、Probit回归	—	—
MILLER [20]	医疗健康素养研究 1992年美国全国成人识字调查	3 209/51	直接删除	logistic回归	AUC=0.733	内部验证(随机拆分验证)、外部验证
LAURSEN [21]	英国2011年生活技能调查	4 773/—	—	logistic回归	AUC=0.78、0.77、0.73	内部验证(随机拆分验证)、外部验证
HANCHATE [22]	美国健康访谈调查	2 824/99	直接删除	logistic回归	AUC=0.81	内部验证、外部验证
陈思婷 [23]	调查收集	1 975/—	—	logistic回归	AUC=0.68	—
LURIE [24]	2003年美国国家成人识字评估	18 541/1 075	多重插补	Probit回归	—	—
CAMPBELL [25]	前瞻性队列研究	1 501/—	—	logistic回归	AUC=0.52、0.69、0.64	—
ROWLANDS [26]	伦敦兰贝斯数据网	228 610/47 097	多重插补	logistic回归	—	—
LEUNG [27]	调查收集	304/6	多重插补	logistic回归	AUC=0.921	—
JEPPESEN [28]	调查收集	225/—	直接删除	logistic回归	AUC=0.820	内部验证(K折交叉验证)

注：“—”表示未报告。

使用独立数据库进行外部验证仅4篇，占18.18%。见表2。

3 讨论

本研究系统检索了国内外关于健康素养预测模型

构建的研究，共纳入22篇文献，涉及普通居民、老年人和学生等多种群体，包括48个预测模型，其中超过一半的模型存在较高的偏倚风险，主要与预测因子的选择、缺失值处理方法及模型评价方式有关。

近年来，健康素养预测模型构建研究逐渐增多，

模型构建方法日趋多样化与复杂化。这些模型通常依赖大样本数据,尤其是大型人口学调查数据。这类数据能增强模型的稳健性和泛化能力,提高预测结果的准确性和可靠性,但也存在数据采集不规范、数据缺失、患者失访和各数据收集系统协调与融合不足等问题^[29]。随着电子医疗记录和电子健康记录的发展,未来研究应尽量选用这些数字化数据资源,并确保数据采集过程的规范性。

预测模型构建方法的选择上,数据科学和人工智能的进步使机器学习在复杂数据处理上展现了巨大的潜力。基于树的算法,如随机森林,因能有效处理复杂的非线性关系和高维数据,且具有较高的模型可解释性,成为研究首选^[7-9]。回归模型因操作简单而被广泛应用,但在处理复杂问题时可能表现不足^[15-18]。支持向量机和神经网络也在部分研究中应用,但是神经网络在数据量要求和可解释性方面有所限制^[7, 13-14]。缺失值处理在模型构建中起关键作用,研究中常见的处理方法包括多重插补和直接删除缺失数据。多重插补法能够最大程度保留数据的信息,提高模型的准确性和稳健性,但其计算复杂性较高^[26-27]。直接删除缺失数据虽简便,但可能影响样本量,降低模型泛化能力^[22, 28]。研究者应根据数据特性和研究目标选择合适的缺失值处理方法。

本研究发现,健康素养预测模型中出现频率较高的预测因子有年龄、文化程度、经济水平、健康状况和互联网使用。模型验证方面,*AUC*值是衡量模型区分度的关键指标,在42个报告*AUC*值的模型中,31个模型的*AUC*值 ≥ 0.7 ,显示出良好的预测性能。然而,许多研究在规范性方面存在不足,如缺乏外部验证或未完整报告模型性能评价指标。

综上所述,目前的健康素养预测模型构建研究仍存在规范性不足、验证不充分等问题。研究应更加注重模型构建的标准化,采用多样化的数据来源,并遵循相关的报告指南,确保研究过程严谨、结果可靠。结合临床与公共卫生实际需求,优化算法选择,加强外部验证,提高预测精度与可解释性。通过不断改进预测模型,进一步推动健康素养的精准干预,促进健康公平,为公共卫生政策的制定提供更有力的支持。

参考文献

[1] 中华人民共和国国家卫生健康委员会. 2023年全国居民健康素养水平提高到29.70% [EB/OL]. [2024-11-01]. <http://www.nhc.gov.cn/xcs/s3582/202404/287e15ca9fd148b5ab9debce59f58c6d.shtml>.

- [2] 闫晓彤, 徐越, 姚丁铭, 等. 2016—2021年浙江省农村居民健康素养分析 [J]. 预防医学, 2022, 34 (10): 1053-1058.
YAN X T, XU Y, YAO D M, et al. Health literacy among rural residents in Zhejiang Province from 2016 to 2021 [J]. China Prev Med J, 2022, 34 (10): 1053-1058. (in Chinese)
- [3] LUO W, NGUYEN T, NICHOLS M, et al. Is demography destiny? Application of machine learning techniques to accurately predict population health outcomes from a minimal demographic dataset [J/OL]. PLoS One, 2015, 10 (5) [2024-11-01]. <https://doi.org/10.1371/journal.pone.0125602>.
- [4] FLAXMAN A D, VOS T. Machine learning in population health: Opportunities and threats [J/OL]. PLoS Med, 2018, 15 (11) [2024-11-01]. <https://doi.org/10.1371/journal.pmed.1002702>.
- [5] TRICCO A C, LILLIE E, ZARIN W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation [J]. Ann Intern Med, 2018, 169 (7): 467-473.
- [6] 陈茹, 王胜锋, 周家琛, 等. 预测模型研究的偏倚风险和适用性评估工具解读 [J]. 中华流行病学杂志, 2020, 41 (5): 776-781.
CHEN R, WANG S F, ZHOU J C, et al. Introduction of the Prediction model Risk Of Bias ASsessment Tool: a tool to assess risk of bias and applicability of prediction model studies [J]. Chin J Epidemiol, 2020, 41 (5): 776-781. (in Chinese)
- [7] ZHOU R S, YIN W H, LI W J, et al. Prediction model for infectious disease health literacy based on synthetic minority oversampling technique algorithm [J/OL]. Comput Math Methods Med, 2022 [2024-11-01]. <https://doi.org/10.1155/2022/8498159>.
- [8] İNCEOĞLU F, DENİZ S, YAGIN F H. Prediction of effective sociodemographic variables in modeling health literacy: a machine learning approach [J/OL]. Int J Med Inform, 2023 [2024-11-01]. <https://doi.org/10.1016/j.ijmedinf.2023.105167>.
- [9] XIE T, ZHANG N, MAO Y, et al. How to predict the electronic health literacy of Chinese primary and secondary school students?: establishment of a model and web nomograms [J/OL]. BMC Public Health, 2022, 22 [2024-11-01]. <https://doi.org/10.1186/s12889-022-13421-4>.
- [10] HONG Y, ZHANG X D, CHEN J X. XGBoost-based prediction modelling and analysis for health literacy assessment [J]. Int J Model Identif Control, 2021, 39 (3): 229-235.
- [11] 季娜, 贾宝琪, 崔莲花. 基于 logistic 回归及决策树模型的大学生健康素养影响因素分析 [J]. 预防医学论坛, 2023, 29 (12): 881-886.
JI N, JIA B Q, CUI L H. Analysis on influencing factors of college students' health literacy based on logistic regression and decision tree model [J]. Prev Med Trib, 2023, 29 (12): 881-886. (in Chinese)
- [12] 李现文, 李春玉, KIM M Y, 等. 决策树与 Logistic 回归在高血压患者健康素养预测中的应用 [J]. 护士进修杂志, 2012, 27 (13): 1157-1159.
LI X W, LI C Y, KIM M Y, et al. Application of decision tree and logistic regression on the health literacy prediction of hypertension patients [J]. J Nurses Train, 2012, 27 (13): 1157-1159.

- (in Chinese)
- [13] 林丰, 林彩虹, 叶少英, 等. 基于逐步 Logistic 回归和神经网络模型的健康素养预测及对比研究 [J]. 中国健康教育, 2018, 34 (8): 699-702.
- LIN F, LIN C H, YE S Y, et al. Comparison on health literacy prediction between stepwise logistic regression and neural network model [J]. *Chin J Health Educ*, 2018, 34 (8): 699-702. (in Chinese)
- [14] 王可, 赵华硕, 张虹, 等. 基于 SMOTE 算法与机器学习的老年人健康素养预测研究 [J]. 中国校医, 2019, 33 (9): 641-643.
- WANG K, ZHAO H S, ZHANG H, et al. Prediction of health literacy of the elderly based on SMOTE algorithm and machine learning [J]. *Chin J School Doctor*, 2019, 33 (9): 641-643. (in Chinese)
- [15] HOU W H, KUO K N, CHEN M J, et al. Simple scoring algorithm to identify community-dwelling older adults with limited health literacy: a cross-sectional study in Taiwan [J/OL]. *BMJ Open*, 2021, 11 [2024-11-01]. <https://doi.org/10.1136/bmjopen-2020-045411>.
- [16] ZHANG Q H, YIN J Y, WANG Y J, et al. A nomogram for predicting the infectious disease-specific health literacy of older adults in China. [J]. *Asian Nurs Res*, 2024, 18 (2): 106-113.
- [17] VAN DER HEIDE I, UITERS E, SØRENSEN K, et al. Health literacy in Europe: the development and validation of health literacy prediction models [J]. *Eur J Public Health*, 2016, 26 (6): 906-911.
- [18] LUBIMIR K T. Developing a predictive model for measuring health literacy in Hawaii's adult populations [D]. Hawaii: University of Hawaii, 2017.
- [19] MARTIN L T, RUDER T, ESCARCE J J, et al. Developing predictive models of health literacy [J]. *J Gen Intern Med*, 2009, 24 (11): 1211-1216.
- [20] MILLER M J, DEGENHOLTZ H B, GAZMARMARIAN J A, et al. Identifying elderly at greatest risk of inadequate health literacy: a predictive model for population-health decision makers [J]. *Res Social Adm Pharm*, 2007, 3 (1): 70-85.
- [21] LAURSEN K R, SEED P T, PROTHEROE J, et al. Developing a method to derive indicative health literacy from routine socio-demographic data [J]. *J Health Commun*, 2015, 1 (1): 1-9.
- [22] HANCHATE A D, ASH A S, GAZMARMARIAN J A, et al. The Demographic Assessment for Health Literacy (DAHL): a new tool for estimating associations between health literacy and outcomes in national surveys [J]. *J Gen Intern Med*, 2008, 23 (10): 1561-1566.
- [23] 陈思婷, 马艺菲, 卓琳, 等. 苏北某市居民传染病健康素养水平与影响因素分析 [J]. 现代预防医学, 2020, 47 (23): 4307-4311.
- CHEN S T, MA Y F, ZHUO L, et al. The level of infectious disease health literacy and its influencing factors among residents in a city in northern Jiangsu province [J]. *Mod Prev Med*, 2020, 47 (23): 4307-4311. (in Chinese)
- [24] LURIE N, MARTIN L T, RUDER T, et al. Estimating and mapping health literacy in the State of Missouri [M]. Santa Monica: RAND Corporation, 2010.
- [25] CAMPBELL P, LEWIS M, CHEN Y, et al. Can patients with low health literacy be identified from routine primary care health records? A cross-sectional and prospective analysis [J]. *BMC Fam Pract*, 2019, 20 (1): 1-11.
- [26] ROWLANDS G, WHITNEY D, MOON G. Developing and applying geographical synthetic estimates of health literacy in GP clinical systems [J/OL]. *Int J Environ Res Public Health*, 2018, 15 (8) [2024-11-01]. <https://doi.org/10.3390/ijerph15081709>.
- [27] LEUNG A Y M, YU E Y T, LUK J K H, et al. Rapid Estimate of Inadequate Health Literacy (REIHL): development and validation of a practitioner-friendly health literacy screening tool for older adults [J]. *Hong Kong Med J*, 2020, 26 (5): 404-412.
- [28] JEPPESEN K M, COYLE J D, MISER W F. Screening questions to predict limited health literacy: a cross-sectional study of patients with diabetes mellitus [J]. *Ann Fam Med*, 2009, 7 (1): 24-31.
- [29] GOLDSTEIN B A, NAVAR A M, PENCINA M J, et al. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review [J]. *J Am Med Inform Assoc*, 2017, 24 (1): 198-208.

收稿日期: 2024-06-27 修回日期: 2024-11-01 本文编辑: 徐文璐