

基于随机森林模型构建住院老年患者轻度认知功能障碍的临床预测模型

吴瑞凯¹, 黄思莹², 张媛^{2*}, 韩正风^{2*}

¹新疆医科大学公共卫生学院, 新疆 乌鲁木齐

²新疆医科大学第一附属医院老年医学科, 新疆 乌鲁木齐

收稿日期: 2024年3月27日; 录用日期: 2024年4月21日; 发布日期: 2024年4月30日

摘要

目的: 构建住院老年患者轻度认知功能障碍的临床预测模型, 对早期轻度认知功能障碍患者识别提供依据。方法: 便利抽样法选取2022年1月~2023年4月在新疆医科大学第一附属医院住院的老年患者1019例, 将其按7:3比例随机分为建模队列($n = 713$)和验证队列($n = 306$), 采用一般资料、体格检查、量表评定, MCI相关指标诊断对其进行调查。随机森林算法结果进行重要性变量排序, 构建LASSO回归分析建模队列, 基于logistic回归结果构建住院老年患者MCI的列线图, 同时对模型进行校准, 同时验证模型效益。结果: 本研究MCI检出率21.6% (220/1019), 男性22.5% (106/472)、女性20.8% (114/547), 单因素分析得到17个差异变量, LASSO回归分析当lambda.min值为0.005607618时误差最小, 对应的影响因素数目为8个, 重要性排序居前8位的自变量为年龄、日常生活能力评分、文化程度、Morse跌倒风险评分高风险、营养风险、用药种数、吸烟、职业, 多因素逐步logistics回归, 最终得到3个影响因素: 年龄、日常生活能力评分、文化程度($P < 0.05$)。构建住院老年MCI患者的临床预测模型并绘制列线图。构建训练组和验证组列线图的ROC曲线AUC大小分别为0.933 (95% CI: 0.913~0.953)和0.929 (95% CI: 0.896~0.961), 通过Hosmer-Lemeshow检验显示训练组 $P = 0.919$, 验证组 $P = 0.726$ (均 > 0.05), 两组拟合优度较好, 该列线图模型具有良好的校准度。结论: 基于住院老年患者MCI的影响因素构建临床预测模型, 年龄增加、日常生活能力下降、受教育年限较少(3项指标)明显增加老年人发生MCI的风险。经过系列验证提示该模型的训练组和验证组均具有净收益范围, 一致性和预测效能较好, 为临床医务人员早期筛查MCI提供依据。

关键词

轻度认知功能障碍, 住院老年患者, 随机森林, LASSO回归, 多因素Logistic回归, 临床预测模型, 列线图

To Construct a Clinical Prediction Model of Mild Cognitive Impairment in Hospitalized Elderly Patients Based on Random Forest Model

*通讯作者。

文章引用: 吴瑞凯, 黄思莹, 张媛, 韩正风. 基于随机森林模型构建住院老年患者轻度认知功能障碍的临床预测模型[J]. 临床医学进展, 2024, 14(4): 2801-2816. DOI: 10.12677/acm.2024.1441359

Ruikai Wu¹, Siying Huang², Yuan Zhang^{2*}, Zhengfeng Han^{2*}

¹School of Public Health, Xinjiang Medical University, Urumqi Xinjiang

²Department of Geriatrics, The First Affiliated Hospital of Xinjiang Medical University, Urumqi Xinjiang

Received: Mar. 27th, 2024; accepted: Apr. 21st, 2024; published: Apr. 30th, 2024

Abstract

Objective: To construct a clinical prediction model of mild cognitive dysfunction in hospitalized elderly patients, and to provide evidence for early identification of patients with mild cognitive dysfunction. **Methods:** A total of 1019 elderly patients hospitalized in the First Affiliated Hospital of Xinjiang Medical University from January 2022 to April 2023 were randomly divided into a modeling cohort (n = 713) and a validation cohort (n = 306) according to a ratio of 7:3 by convenient sampling method. General data, physical examination, scale evaluation and MCI related index diagnosis were used to investigate the patients. Rank the importance variables of the results of random forest algorithm, build a LASSO regression analysis modeling cohort, and build a Nomogram of MCI for hospitalized elderly patients based on the logistic regression results. At the same time, calibrate the model and verify the benefit of the model. **Results:** In this study, the detection rate of MCI was 21.6% (220/1019), male 22.5% (106/472), female 20.8% (114/547). Univariate analysis obtained 17 difference variables, and LASSO regression analysis had the smallest error when lambda.min value was 0.005607618. The corresponding number of influencing factors is 8, and the top 8 independent variables in the order of importance are age, daily living ability score, education level, Morse fall risk score, high risk, nutritional risk, number of drugs, smoking and occupation. Multi-factor logistics regression is carried out, and finally 3 influencing factors are obtained: Age, daily living ability score, education level (P < 0.05). The clinical prediction model of hospitalized elderly patients with MCI was constructed and the nomogram was drawn. The ROC curve AUC sizes of the training group and the verification group were 0.933 (95% CI: 0.913~0.953) and 0.929 (95% CI: 0.896~0.961), respectively. Hosmer-Lemeshow test showed that P = 0.919 in the training group. Verification group P = 0.726 (all > 0.05), the two groups of goodness of fit is good, the column-line model has a good calibration degree. **Conclusion:** A clinical prediction model was built based on the influencing factors of MCI in hospitalized elderly patients. The risk of MCI in elderly patients was significantly increased by increasing age, decreasing ability of daily living and less years of education (3 indicators). A series of validations indicated that both the training group and the verification group of the model had a net benefit range, good consistency and prediction efficiency, which provided a basis for early screening of MCI by clinical staff.

Keywords

MCI, Hospitalized Elderly Patients, Random Forest, LASSO Regression, Multivariate Logistic Regression, Clinical Prediction Model, Nomograph

Copyright © 2024 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

随着人口老龄化一步加重，老年健康问题受到国内外专家学者的重点关注。我国痴呆患病人数占全

球 1/4, 其中阿尔兹海默病(Alzheimer's Disease, AD)占痴呆的 60%~80%, 是老年人失能和死亡的主要原因。最新专家共识报告[1], 我国 60 岁以上人群中 有 983 万 AD 患者、3877 万轻度认知功能障碍(Mild Cognitive Impairment, MCI)患者, 且随着人口老龄化, 我国已成为全球 AD 患者数增速最快的国家之一, AD 防治是一个世界性难题, 首要原因在于难以早期发现早期诊断, 我国综合医院门诊早期痴呆诊断率仅为 0.1%。MCI 患者是 AD 的高危人群, 发展为 AD 的概率是未患 MCI 人群的 10 倍。

MCI 是正常老化到痴呆之间的过渡状态, 同时也是痴呆防治的重要“干预窗口期”[2] [3]。目前针对住院老年患者 MCI 的风险预测模型研究较少, 为了使 MCI 的早期预防成为可能, 本研究调查住院老年患者 MCI 现状并分析其影响因素, 构建发生 MCI 的风险预测模型, 为医护人员进行早期筛查和干预提供参考。

2. 对象与方法

2.1. 研究对象

2022 年 1 月~2023 年 4 月在新疆医科大学第一附属医院老年病科住院的老年患者采用便利抽样法进行调查。纳入标准: 1) 年龄 ≥ 60 岁; 2) 一般情况良好; 3) 能够配合完成问卷调查。排除标准: 1) 严重认知障碍无法交流; 2) 有严重的心、脑血管疾病或晚期肿瘤; 3) 由于视力及听力障碍无法完成调查者; 4) 长期卧床者。参照《中国老年人认知障碍诊治流程专家建议》中 MCI 的诊断标准[4], 研究对象分为非 MCI 组($n = 799$)和 MCI 组($n = 220$), 本研究通过了新疆医科大学第一附属医院伦理委员会批准, 所有调查对象均知情同意。

2.2. 研究方法

2.2.1. 病史资料

通过查阅病历和询问病史, 记录所有受试者生化检查、住院疾病诊断资料。

2.2.2. 问卷调查

问卷包括以下内容: 1) 一般人口学资料调查表: 性别、年龄、文化程度、职业、婚姻、居住情况、生活方式(是否吸烟、饮酒)。2) 简易智能量表(Mini-Mental State Examination, MMSE) [5], 评估研究对象的时间定向、地点定向、即刻记忆、注意与计算能力、语言能力及视空间能力, 共 30 个条目, 总分(范围为 0~30 分), 得分越低表明认知功能越差。3) 蒙特利尔认知评估量表(Montreal Cognition Assessment, MoCA), 共评估 8 个认知领域(视空间与执行功能、命名、记忆、注意、语言、抽象、延迟回忆、定向), 用以评估患者的整体认知功能。总分(范围为 0~30 分), 得分越低, 认知障碍越严重。4) 日常生活能力评定(Activity of Daily Living, ADL)量表[6]: 使用 Barthel 指数进行日常生活能力进行了评价, 共包括进食、洗澡等 10 项评估, 总分为 100 分, 超过 60 分的人群被认为基本生活可自理, 低于 60 分认为生活需要帮助。5) Morse 跌倒风险评估量表[7]: 表该量表测试内容包括跌倒史、超过 1 个医学诊断、行走辅助、静脉注射治疗或留置套管针、步态、认知状态。总分 125 分, 评分 > 45 分为跌倒高风险。6) Fried 衰弱量表[8]: 包括 5 项内容: ① 体质量下降: 近半年内体质量下降 $> 5\%$ (非节食或运动); ② 步速减慢: 步行 4.5 m, ≥ 7 s; ③ 握力降低: 男性 < 28 kg, 女性 < 18 kg; ④ 躯体活动下降: 男性每周散步 < 2.5 h, 女性每周散步 < 2 h; ⑤ 疲乏: 过去 1 周内 > 3 d 觉得自己做任何事情都费劲或缺乏干劲。总分 0~5 分, 0 分为无衰弱, 1~2 分为衰弱前期, 3~5 分为衰弱。

2.2.3. 统计分析

SPSS25.0 和 R4.3.1 软件进行统计分析, 符合正态分布的计量资料采用($\bar{x} \pm s$)表示, 两组间比较方差

齐采用独立样本 t 检验, 方差不齐采用独立样本 t' 检验; 计数资料以相对数表示, 组间比较采用 χ^2 检验。R 软件进行随机森林模型分析, 单因素分析中差异有统计学意义的变量纳入随机森林模型, 得出变量重要性评分并排序, 采用套索(The Least Absolute Shrinkage and Selection Operator, LASSO)回归分析进行变量选择。筛选出的变量运用多因素 logistics 回归进行多因素分析, 采用逐步法确定住院老年患者发生 MCI 的危险因素, 利用一致性指数(Concordance-Index, C-Index), 受试者工作特征(Receiver Operator Characteristic, ROC)曲线, ROC 曲线下面积(Area Under the Curve, AUC)、Calibration 校正曲线及决策曲线(Decision Curve Analysis, DCA)对模型的鉴别、校准以及临床应用进行评估, 并利用验证队列对模型进行验证。

3. 结果

3.1. 研究对象的一般情况

共纳入研究对象 1019 例, 年龄 60~99 岁, 平均年龄(73.3 ± 8.2)岁, 男性 472 例、女性 547 例。住院老年患者 MCI 检出率为 21.6% (220/1019), 男性 22.5% (106/472)、女性 20.8% (114/547), 依据是否患有 MCI, 将患者分为 MCI 组($n = 220$)和非 MCI 组($n = 799$), 见表 1。

3.2. MCI 与非 MCI 组患者一般资料及问卷调查、量表评估、慢性病、生化指标比较

两组患者一般资料及问卷调查中年龄 MCI 组 82.98 ± 5.34 高于正常组 70.62 ± 6.75 , 差异有统计学意义($P < 0.001$); BMI 组 MCI 组 24 ± 4.07 低于正常组 25.56 ± 3.95 , 差异有统计学意义($P < 0.001$); 文化程度文盲的 MCI 检出率(51.3%)高于小学(32.5%)和中学及以上(15.7%), 差异有统计学意义($P < 0.001$); 职业体力劳动的 MCI 检出率(24.5%)高于脑力劳动(14.2%), 差异有统计学意义($P < 0.001$); 婚姻未婚/离婚/丧偶组的 MCI 检出率(39.1%)高于已婚组(17.1%), 差异有统计学意义($P < 0.001$); 居住情况的养老院机构 MCI 检出率(45.9%)高于独居(20.7%)和家人同住(20.7%), 差异有统计学意义($P < 0.001$); 吸烟的 MCI 检出率(27.2%)低于不吸烟(20.1%), 差异有统计学意义($P < 0.05$), 见表 1。量表评估中日常生活能力评分中度依赖的 MCI 检出率(67.3%)高于重度依赖(65.7%)、轻度依赖(30.5%)、无依赖(6.0%), 差异有统计学意义($P < 0.001$); Morse 跌倒风险评分高风险 MCI 检出率(27.2%)高于无高风险(5.9%), 差异有统计学意义($P < 0.001$); Fried 衰弱量表评分衰弱综合症的 MCI 检出率(39.6%)高于正常组(17.3%)、衰弱前期(16.2%), 差异有统计学意义($P < 0.001$), 见表 2。用药种数(种) MCI 组 6.58 ± 2.74 高于正常组 5.94 ± 2.42 , 差异有统计学意义($P < 0.001$); 高血脂的 MCI 检出率(15.6%)低于未患高血脂(23.4%), 差异有统计学意义($P < 0.05$); 冠心病的 MCI 检出率(24.9%)高于未患冠心病(16.4%), 差异有统计学意义($P < 0.001$); 脑血管病的 MCI 检出率(25.3%)高于未患脑血管病(18.7%), 差异有统计学意义($P < 0.05$); 慢性阻塞性肺疾病的 MCI 检出率(30.2%)低于未患慢性阻塞性肺疾病(18.6%), 差异有统计学意义($P < 0.001$); 慢性肾功能不全的 MCI 检出率(36.3%)低于未患慢性肾功能不全(20.3%), 差异有统计学意义($P < 0.001$); 营养风险的 MCI 检出率(34.7%)低于未患营养风险(17.0%), 差异有统计学意义($P < 0.001$), 见表 3。

Table 1. Comparison of general data of elderly hospitalized patients in MCI and non-MCI groups ($n = 1019$)

表 1. MCI 与非 MCI 组老年住院患者一般资料的比较($n = 1019$)

变量	正常组($n = 799$)	MCI 组($n = 220$)	t/χ^2	P
年龄(岁)	70.62 ± 6.75	82.98 ± 5.34	-25.073	<0.001
BMI (kg/m^2)	25.56 ± 3.95	24 ± 4.07	5.146	<0.001
性别[n (%)]				
男	366 (77.5%)	106 (22.5%)	0.391	0.532
女	433 (79.2%)	114 (20.8%)		

续表

文化程度[n (%)]				
文盲	37 (48.7)	39 (51.3)	68.862	<0.001
小学	133 (67.5)	64 (32.5)		
中学及以上	629 (84.3)	117 (15.7)		
职业[n (%)]				
体力劳动	551 (75.5%)	179 (24.5%)	13.06	<0.001
脑力劳动	248 (85.8%)	41 (14.2%)		
婚姻[n (%)]				
已婚	673 (82.9%)	139 (17.1%)	47.213	<0.001
未婚/离婚/丧偶	126 (60.9%)	81 (39.1%)		
居住情况[n (%)]				
家人合住	710 (79.3%)	185 (20.7%)	13.454	<0.001
独居	69 (79.3%)	18 (20.7%)		
养老院机构	20 (54.1%)	17 (45.9%)		
饮酒[n (%)]				
不饮酒	698 (78.6%)	190 (21.4%)	0.153	0.696
饮酒	101 (77.1%)	30 (22.9%)		
吸烟[n (%)]				
不吸烟	641 (79.9%)	161 (20.1%)	5.106	<0.05
吸烟	158 (72.8%)	59 (27.2%)		

Table 2. Comparison of scale assessment results of elderly hospitalized patients in MCI and non-MCI groups (n = 1019)
表 2. MCI 与非 MCI 组老年住院患者量表评估结果的比较(n = 1019)

变量	正常组(n = 799)	MCI 组(n = 220)	χ^2	P
日常生活能力评分[n (%)]				
无	469 (94.0%)	30 (6.0%)	196.225	<0.001
轻度依赖	301 (69.5%)	132 (30.5%)		
中度依赖	17 (32.7%)	35 (67.3%)		
重度依赖	12 (34.3%)	23 (65.7%)		
Morse 跌倒风险评分高风险[n (%)]				
否	254 (94.1%)	16 (5.9%)	53.239	<0.001
是	545 (72.8%)	204 (27.2%)		
Fried 衰弱量表评分[n (%)]				
正常	81 (82.7%)	17 (17.3%)	56.763	<0.001
衰弱前期	579 (83.8%)	112 (16.2%)		
衰弱综合征	139 (60.4%)	91 (39.6%)		

Table 3. Comparison of chronic diseases among hospitalized elderly patients in MCI and non-MCI groups (n = 1019)
表 3. MCI 与非 MCI 组老年住院患者慢性病比较(n = 1019)

变量	正常组(n = 799)	MCI 组(n = 220)	t'/ χ^2	P
用药种数(种)	5.94 ± 2.42	6.58 ± 2.74	-3.357	<0.001
高血压[n (%)]				
无	206 (79.8%)	52 (20.2%)	0.42	0.517
有	593 (77.9%)	168 (22.1%)		
高血脂[n (%)]				
无	599 (76.6%)	183 (23.4%)	6.519	<0.05
有	200 (84.4%)	37 (15.6%)		
糖尿病[n (%)]				
无	530 (79.6%)	136 (20.4%)	1.553	0.213
有	269 (76.2%)	84 (23.8%)		
冠心病[n (%)]				
无	331 (83.6%)	65 (16.4%)	10.249	<0.001
有	468 (75.1%)	155 (24.9%)		
脑血管病[n (%)]				
无	466 (81.3%)	107 (18.7%)	6.576	<0.05
有	333 (74.7%)	113 (25.3%)		
慢性阻塞性肺疾病[n (%)]				
无	616 (81.4%)	141 (18.6%)	15.276	<0.001
有	183 (69.8%)	79 (30.2%)		
慢性肾功能不全[n (%)]				
无	748 (79.7%)	191 (20.3%)	11.022	<0.001
有	51 (63.7%)	29 (36.3%)		
营养风险[n (%)]				
无	626 (83.0%)	128 (17.0%)	36.456	<0.001
有	173 (65.3%)	92 (34.7%)		
睡眠障碍[n (%)]				
无	722 (78.1%)	203 (21.9%)	0.751	0.386
有	77 (81.9%)	17 (18.1%)		

3.3. MCI 患者影响因素的筛选

3.3.1. MCI 变量的重要性排序

以 MCI 为因变量, 将单因素分析中差异有统计学意义的变量纳入随机森林模型。各个变量相关性见图 1, 方差膨胀因子(Variance Inflation Factor, VIF)均 < 2, 容忍度均远 > 0.1, 纳入的自变量不存在多重共线性, 见图 2。利用 R 中“Random Forest”程序包输出结果[9], %IncMSE (Increase in Mean Squared Error) 为精度平均减小值, %IncMSE 越大, 说明变量在影响因素中的重要性也越高[10]。随机森林模型结果显

示，变量重要性从高到低依次为：年龄、日常生活能力评分、文化程度、Morse 跌倒风险评分高风险、营养风险、用药种数、吸烟、职业、脑血管病、高血脂、冠心病、居住情况、BMI、婚姻、慢性阻塞性肺疾病、慢性肾功能不全、Fried 衰弱量表评分，见图 3。

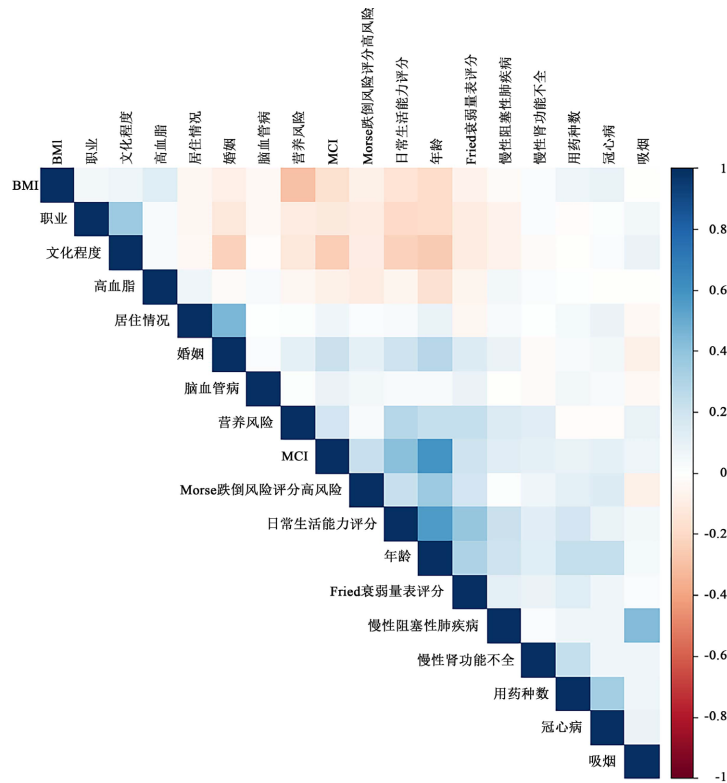


Figure 1. Correlation diagram of each variable
图 1. 各个变量相关图

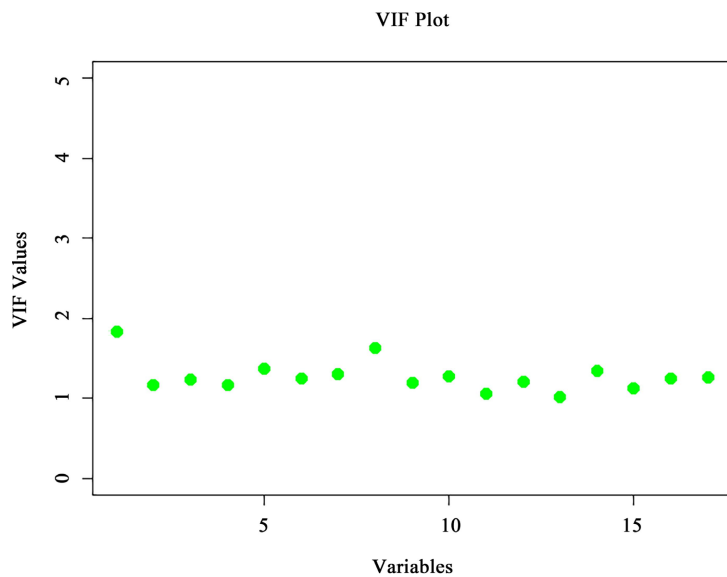


Figure 2. Variance inflation factor diagram
图 2. 方差膨胀因子图

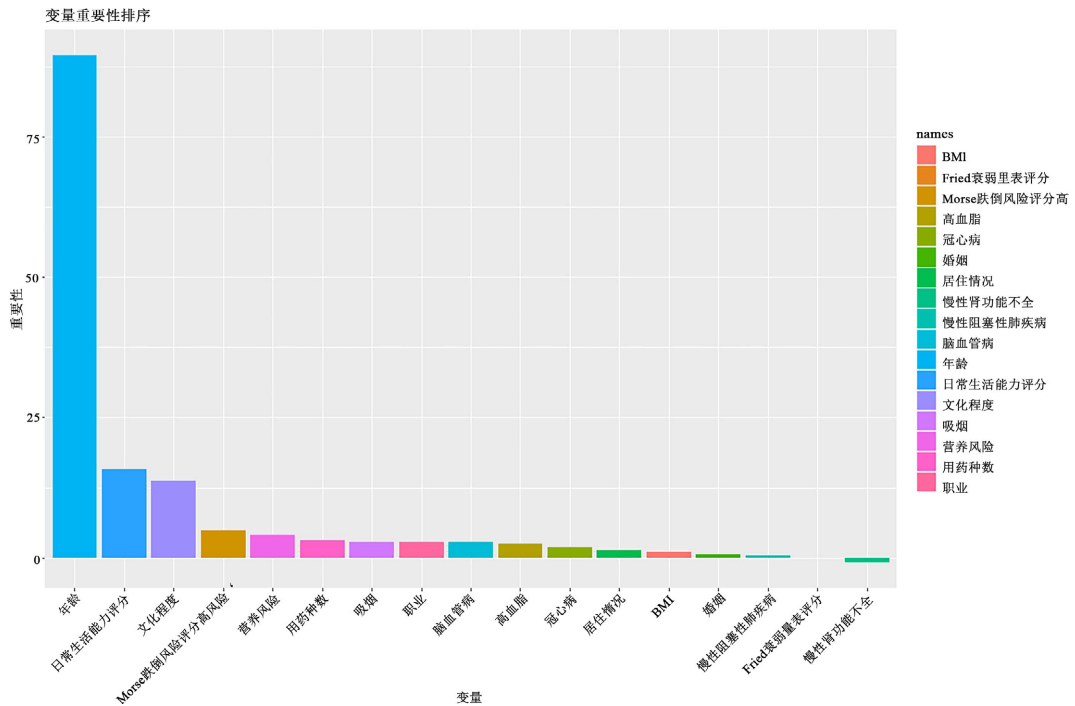
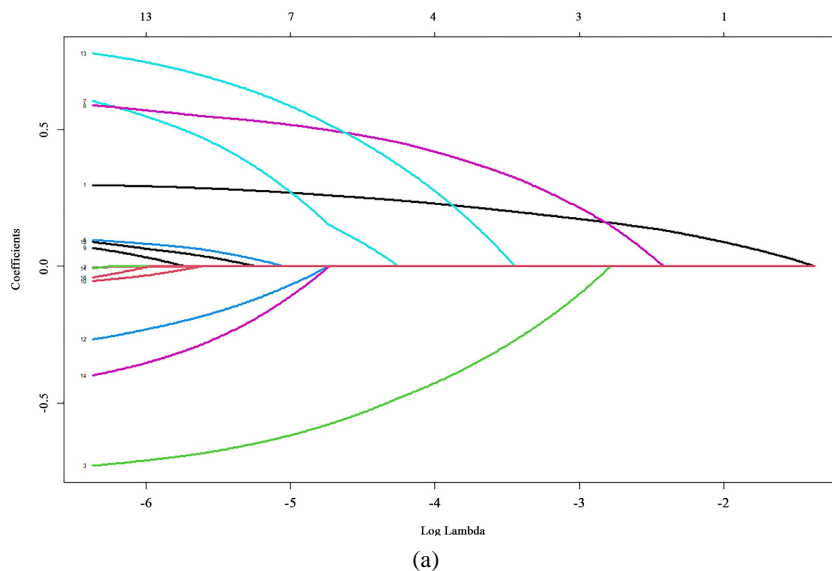


Figure 3. Ranking the importance of MCI affecting hospitalized elderly patients
图 3. 影响住院老年患者 MCI 的重要性排序

3.3.2. 变量筛选

根据变量重要性排序结果，R 软件中 glmnet 函数包对单因素分析中差异的 17 个变量进行套索(The Least Absolute Shrinkage and Selection Operator, LASSO)方法结合既往文献证据筛选关键的影响变量[11]，详见图 4。图中右图左侧垂直的灰线代表 lambda.min，右侧垂直的灰线代表 lambda.1se。在[lambda.min, lambda.1se]区间内模型的偏差变动幅度最小[12]。当 lambda.min 值为 0.005607618 时，误差最小所对应的影响因素数目为 8 个。因此，将排名居前 8 位的年龄、日常生活能力评分、文化程度、Morse 跌倒风险评分高风险、营养风险、用药种数、吸烟、职业纳入多因素 logistics 回归分析。



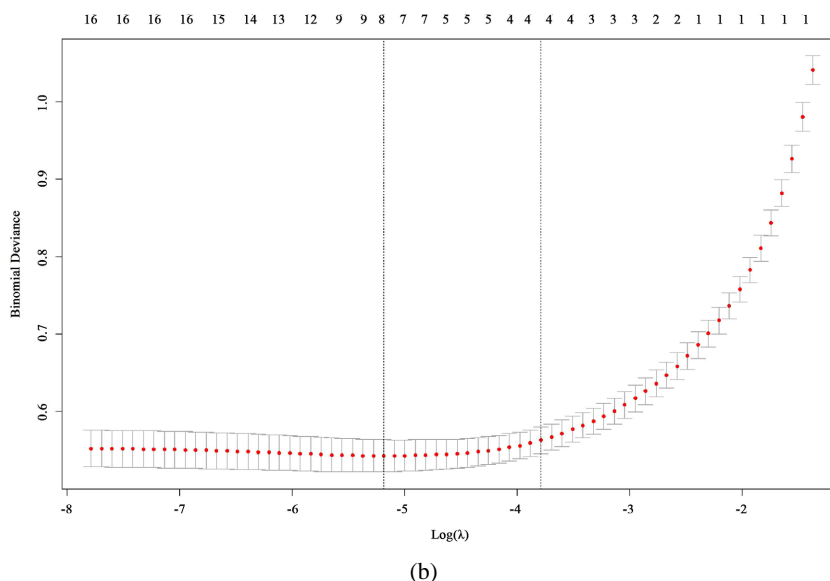


Figure 4. Variable filtering by LASSO regression
图 4. LASSO 回归进行变量筛选

3.4. MCI 影响因素的多因素 Logistics 回归分析

以 MCI 作为因变量，以随机森林模型筛选出的重要性居前 8 位为自变量，进行多因素 logistics 回归分析。结果显示，年龄、日常生活能力评分、文化程度是住院老年 MCI 患者的主要影响因素($P < 0.05$)，详见图 5。

Variable	N	Odds ratio	p
年龄	1019	1.34 (1.28, 1.41)	<0.001
文化程度		Reference	
文盲	76		
小学	197	0.38 (0.18, 0.80)	0.011
中学大学及以上	746	0.20 (0.10, 0.41)	<0.001
职业		Reference	
体力劳动	730		
脑力劳动	289	0.99 (0.55, 1.77)	0.978
吸烟		Reference	
不吸烟	802		
吸烟	217	1.51 (0.89, 2.56)	0.126
日常生活能力评分		Reference	
无需依赖	499		
轻度依赖	433	1.33 (0.77, 2.30)	0.308
中度依赖	52	4.84(1.91, 12.56)	<0.001
重度依赖	35	5.79 (2.06, 16.94)	0.001
Morse跌倒风险评分高风险		Reference	
否	270		
是	749	1.26 (0.65, 2.56)	0.501
营养风险		Reference	
无	754		
有	265	0.87 (0.54, 1.40)	0.573
用药种数	1019	0.98 (0.90, 1.07)	0.648

Figure 5. Forest diagram of logistics regression analysis
图 5. Logistics 回归分析森林图

3.5. MCI 风险临床预测模型列线图

列线图顶端评分标尺(默认为 0~100 分), 每一变量通过垂直线获得对应的积分, 然后所有变量的积分累积得到总分, 总分线在列线图底部的预测线上得到相应的预测风险值。见图 6。

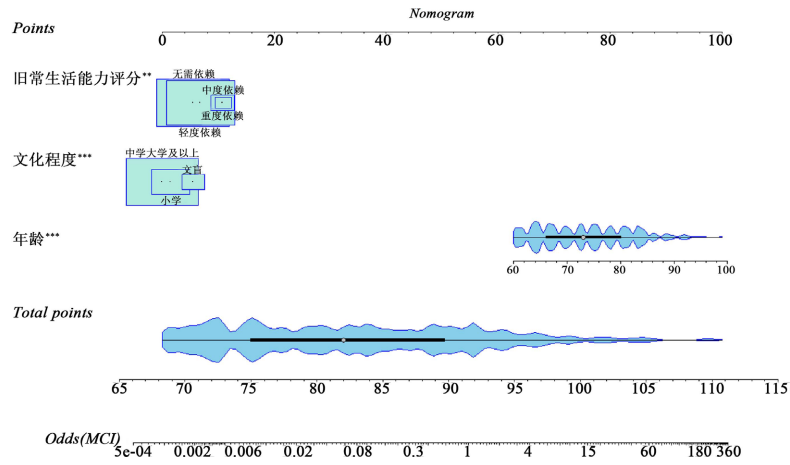


Figure 6. Nomogram of clinical prediction model of MCI in elderly hospitalized patients
图 6. 老年住院患者 MCI 的临床预测模型列线图

3.6. 列线图区分度

训练组列线图模型的 AUC 大小及 95% 置信区间为 0.933 (95% CI: 0.913~0.953), C-Index = 0.929, 有较高准确性。其最佳临界点为 0.300, 对应的特异性为 0.873, 灵敏度为 0.856。验证组构建列线图模型的 AUC 大小及 95% 置信区间为 0.929 (95% CI: 0.896~0.961), C-Index = 0.922, 有较高准确性。其最佳临界点为 0.339, 对应的特异性为 0.875, 灵敏度为 0.878 (见图 7)。

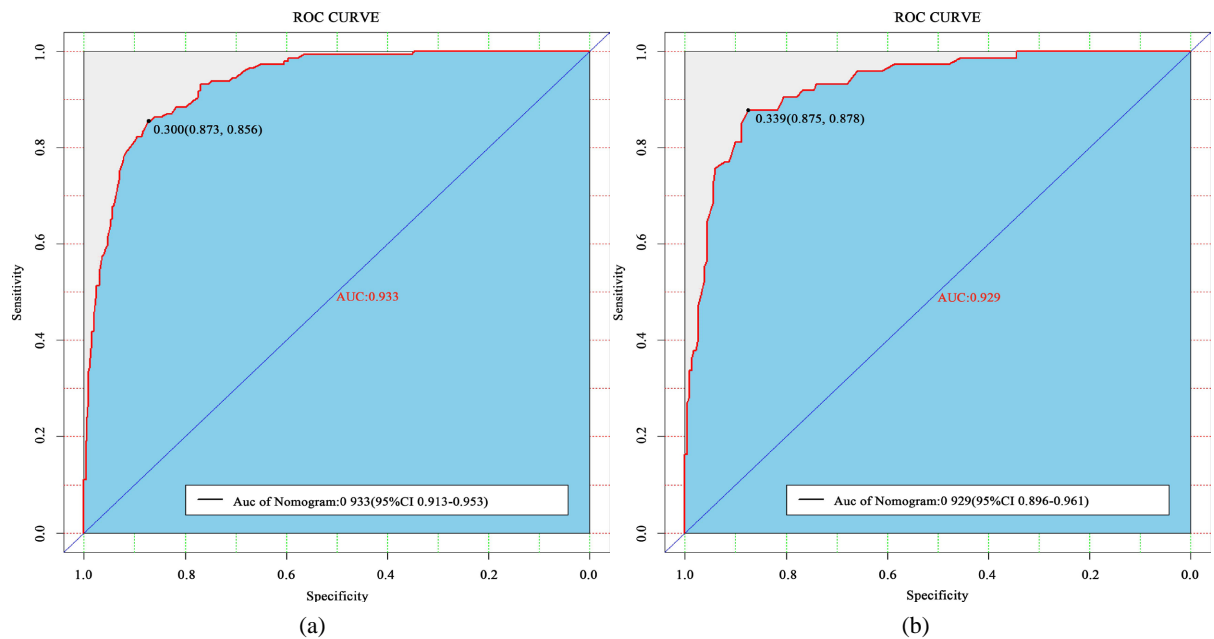


Figure 7. ROC curves constructed by training group and validation group models
图 7. 训练组和验证组模型构建的 ROC 曲线

3.7. 列线图校准度

住院老年患者 MCI 的列线图实际曲线、Bootstrapping 校正曲线与理想曲线非常接近, 说明模型的预测能力优良。通过 Hosmer-Lemeshow 检验显示训练组 $P = 0.919$ 优于验证组 $P = 0.726$, 均 > 0.05 , 两组拟合优度较好, 该临床预测模型有良好的校准度(见图 8, X 轴表示列线图预测 MCI 的概率, Y 轴表示 MCI 发生的实际概率)。

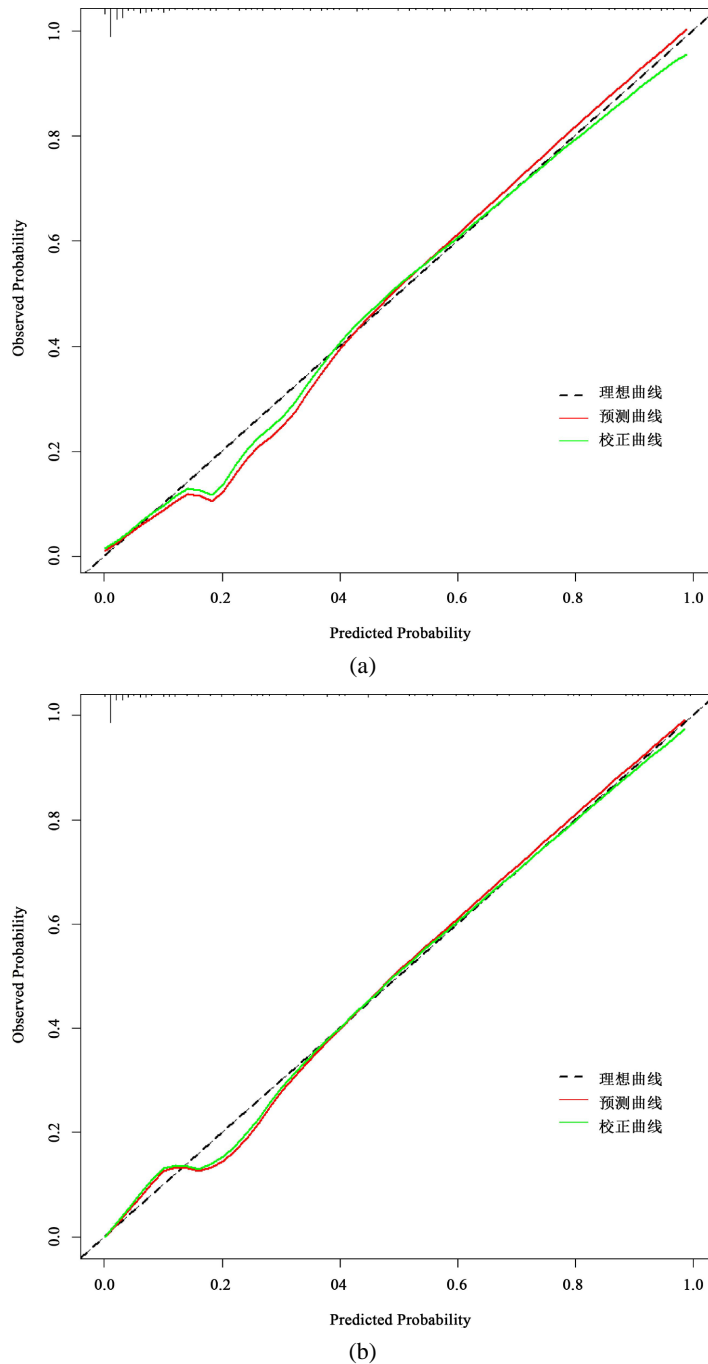


Figure 8. Calibration curves of the training group and verification group models
图 8. 训练组和验证组模型构建的校准曲线

3.8. 临床决策曲线

图 9 中的上图显示当训练组患者列线图模型阈概率值设定为 30%~100%，净获益范围为 0%~95%。下图则显示验证组患者列线图模型阈概率值为 30%~100%时，净获益范围为 0%~95%。两种类型的住院老年 MCI 模型均有净收益范围，模型具有临床价值。这一结论显示了列线图模型在住院老年 MCI 患者预测方面的临床应用价值(决策曲线的 X 轴为阈概率，Y 轴为每个阈概率对应的净获益)。

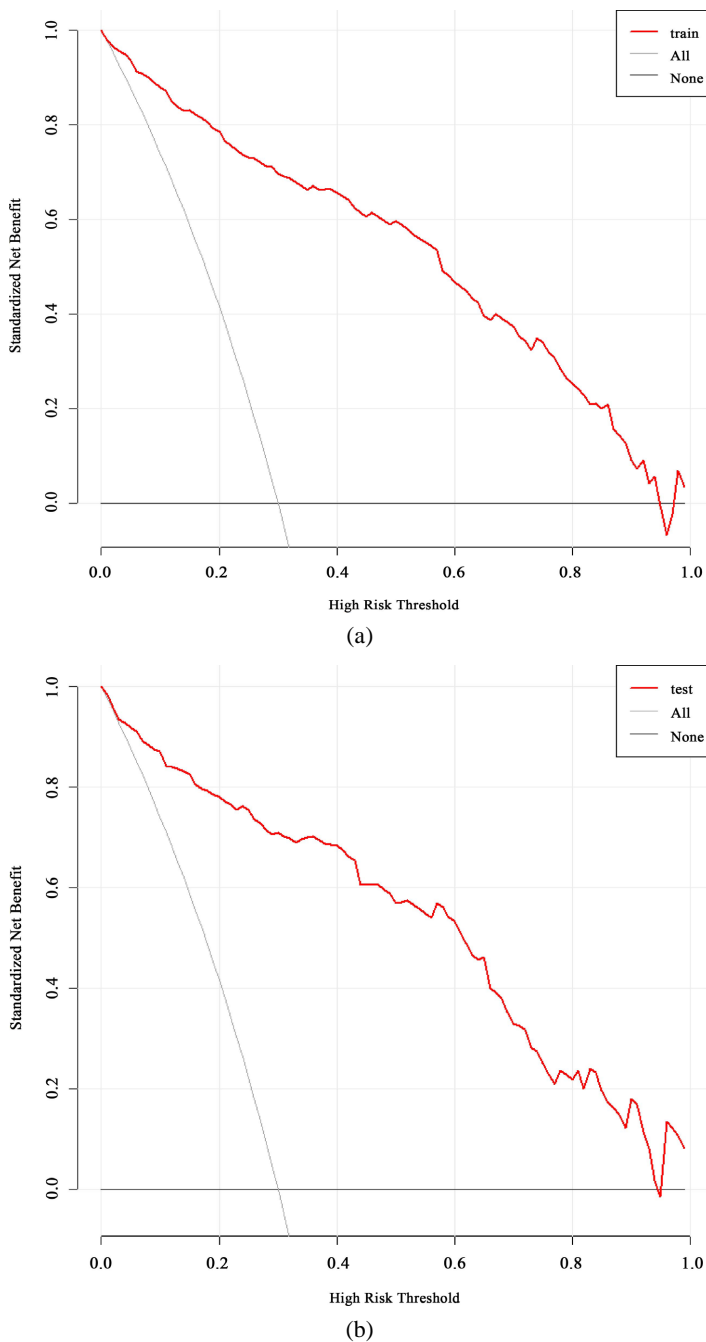


Figure 9. Clinical decision curves constructed by training group and validation group models

图 9. 训练组和验证组模型构建的临床决策曲线

4. 讨论

Alex Ward 研究发现[13], 全球范围内 60 岁以上人口有高达 42.0% 受 MCI 影响。近年来, 国内外专家学者均开展了 MCI 患病率的流行病学调查研究, 研究显示老年人群中 MCI 的发生率为 13%~50.15% [14], 国内外研究均发现 MCI 患病率有逐年提高的趋势[1]; 本研究基于住院老年患者调查研究, MCI 检出率为 21.6% (220/1019), 男性 22.5% (106/472)、女性 20.8% (114/547), 与夏艳秋等[15]基于综合医院老年住院患者 MCI 检出率 24.05% 结果一致, 高于贾龙飞等[14]研究报告中国 60 岁及以上老年人 MCI 的总体患病率 15.5%, 低于叶青芳等[16]中青年高血压住院患者 MCI 的检出率 37.56%。

国内外专家共识已证实高龄是老年人 MCI 的危险因素, MCI 患病率随之年龄的增加而增加[17]。Longfei Jia [14]一项大型老年人群研究证实, MCI 的患病率在 60~69 岁人群中为 11.9%、70~79 岁为 19.3%、80~89 岁为 24.4%, 90 岁或以上为 33.1%。本研究年龄每增加一岁, MCI 的患病风险增加 1.34 倍[OR = 1.34, 95% CI (1.28, 1.41)]与前者结果一致。可能原因为: 1) 随着年龄增加, 与认知功能有关的区域脑内侧前额叶皮层的灰质萎缩, 2) 老年人身体各器官和组织也都发生了退行性变, 这也间接导致 MCI 的发生[18]。

日常生活能力是反映老年人生活质量的重要指标, 其水平的高低对 MCI 具有重要的影响, 日常生活能力与 MCI 两者互为因果关系。贾丛康[19]等研究发现认知功能障碍患者日常活动能力必然要低, 反之日常活动能力低会增加认知功能障碍的发生率。认知功能障碍的患者的日常生活能力受多种因素影响, 不仅与认知功能下降有关系, 与患者的年龄、自身基础疾病等也有密切关系[20]。潘惠英等[21]研究发现, 日常生活能力量表得分低者的生活质量越差, 本研究发现日常生活能力评分中度依赖相对于无依赖发生 MCI 风险为 4.84 倍(OR = 4.84, 95% CI: 1.91~12.56); 重度依赖相对于无依赖发生 MCI 风险为 5.79 倍(OR = 5.79, 95% CI: 2.06~16.94), 依赖程度越高, 量表得分越低, MCI 发生风险更高, 结果同前者研究结果一致。可以因为以下原因: 1) MCI 患者复杂社会功能存在一定程度的损害, 日常生活能力下降, 躯体功能继而下降, 患者完成相同日常活动所消耗的时间明显长于正常老年人。2) 相关研究表明, 日常生活能力下降损伤可能会导致认知功能域的细微缺陷[22] [23]。Bronnick K 等[24]研究发现, 注意力的下降和日常生活活动能力损伤密切相关。3) 日常生活活动能力损伤和认知功能特定功能域障碍之间的密切相关。Paula 等[25]关于认知功能域和日常生活活动能力的损伤研究中发现, 执行能力和记忆能力是日常生活活动能力整体结构中损害有关的认知域。除此以外还发现语言能力、即时记忆和视觉空间能力可能在一定程度上和日常生活能力相关。

既往研究结果一致显示, 受教育年限对认知功能有保护作用[5], MCI 的发生率在受教育低的人群中显著高于受教育高的人群[26], Deng Y [27]等 meta 分析研究发现, 文盲是 MCI 的高危因素, 受教育高是 MCI 的保护因素, 本研究发现文化程度小学相对于文盲发生 MCI 风险为 0.38 倍(OR = 0.38, 95% CI: 0.18~0.80); 文化程度中学及以上相对于文盲发生 MCI 风险为 0.20 倍(OR = 0.20, 95% CI: 0.10~0.41), 文化程度越高, MCI 发生的风险更低, 本研究与前者研究结果一致, 这可能因为受教育程度年限较长的人有着更高的认知储备和更发达的神经网络, 知识储备和大脑的认知网络越发达, 健康信息素养也越高, 从而保护认知功能有关[28]。

MCI 被认为是痴呆早期干预的最佳窗口期, 临床情况下, 根据最新专家共识和诊断指南[1], 评估 MCI 的最准确方法认知功能评估和日常和社会能力评估、非认知性神经精神症状评估、体液检查、影像学检查、结合相关生物学标志物的检测。因此构建适用于临床老年患者 MCI 风险评价的预测模型具有重要意义。

目前国内外针对老年人群认知功能方面建立的风险预测模型有很多类型, 主要包括 Cox 比例风险回归模型[29]、神经网络模型[30]、logistic 回归模型[31]、Rothman-Keller 模型[32]、竞争风险模型[33]和贝

叶斯网络模型[34]等。Choi H 等[35]在卷积神经网络模型基础上建立自动图像系统,系统通过识别脑部脱氧葡萄糖和 β -淀粉样斑块 PET 示踪剂,预测 MCI 患者未来认知能力的下降情况,该模型预测准确性较高但需要海量的成像数据加以验证,且对老年人身体造成的损伤具有不可逆性,因此很难对其进行推广普及。

临床预测模型列线图是一个对用户友好的图形表示工具,可以用来计算每个个体的特定事件的发生概率。通过此次模型的构建与验证,本研究通过 LASSO 回归控制变量,分析得到年龄增加、日常生活能力下降、受教育年限较少(3 项指标)的老年人发生 MCI 风险较高,经 ROC 曲线分析和 Calibration 校准曲线,临床预测模型的预测效能较好。该 3 项指标属于无创操作,且在入院体格检查快速采集,不但发现重要的风险预测作用,在临床应用上也具有重要价值。模型构建时结合了住院老年患者体格检查相关指标进行筛选,临床医生可以实施更多对有益的措施,在临床应用上和老年人居家自检也具有重要价值。

本研究也有一些局限性:1) 本研究基于单中心住院老年人进行 MCI 的调查研究,无法避免存在选择偏移;2) 研究样本量较少,代表性比较局限。故未来期待有多中心、大样本研究对此临床预测模型进行进一步外部验证和完善,集中于影响分析和使用机器学习技术来评估住院老年 MCI 患者的风险和预测为临床医务人员提供可靠、便捷的早期识别住院老年患者 MCI 风险的评估工具。

综上所述,基于住院老年患者 MCI 的影响因素构建临床预测模型,年龄增加、日常生活能力下降、受教育年限较少(3 项指标)明显增加老年人发生 MCI 的风险。经过系列验证提示该模型的训练组和验证组均具有净收益范围,一致性和预测效能较好,为临床医务人员早期筛查 MCI 提供依据。

基金项目

新疆维吾尔自治区卫生健康青年医学科技人才专项科研项目(WJWY-202148);新疆护理学会年度科研项目(2022XH16, 2023XH040);新疆医科大学第一附属医院“青年科研起航”专项(2022YFY-QNRC-07)。

参考文献

- [1] 中华医学会神经病学分会痴呆与认知障碍学组. 阿尔茨海默病源性轻度认知障碍诊疗中国专家共识 2021 [J]. 中华神经科杂志, 2022, 55(5): 421-440.
- [2] 王黎, 周莲, 杨燕妮. 住院高血压患者伴发轻度认知功能障碍的风险预测模型构建[J]. 陆军军医大学学报, 2022, 44(8): 835-841.
- [3] 王志勇, 刘剑刚, 李浩. 阿尔茨海默病和轻度认知功能障碍相关生物学标志物[J]. 中华老年医学杂志, 2016, 35(12): 1345-1349.
- [4] 中华医学会老年医学分会老年神经病学组, 老年人认知障碍诊治专家共识撰写组. 中国老年人认知障碍诊治流程专家建议[J]. 中华老年医学杂志, 2014, 33(8): 817-825.
- [5] 马万瑞, 李漫娜, 王立群, 等. ≥ 55 岁高血压患者发生轻度认知功能障碍的影响因素研究[J]. 中国全科医学, 2023, 26(9): 1075-1079+1085.
- [6] 贺楚梅, 阳前军, 戢秋明, 等. 精神病患者基于日常生活能力评定量表分级护理的效果观察[J]. 护理学报, 2017, 24(8): 40-43.
- [7] 刘墩秀, 丁福, 朱跃平, 等. 汉化版 Morse 跌倒评估量表和 HendrichII 跌倒因素模型量表在老年住院患者跌倒风险评估中的适用性比较[J]. 中国护理管理, 2020, 20(8): 1168-1172.
- [8] 韩君, 王君俏, 谢博钦, 等. Fried 衰弱表型和 FRAIL 量表及埃德蒙顿衰弱评估量表在社区高龄老年人衰弱筛查中一致性和适用性的比较研究[J]. 中国全科医学, 2021, 24(21): 2669-2675.
- [9] 李娜, 徐晨雪, 韩然然, 等. 基于随机森林模型的乳腺癌病人心理资本现状及影响因素[J]. 护理研究, 2023, 37(8): 1325-1331.
- [10] Angel, Y. and McCabe, M.F. (2022) Machine Learning Strategies for the Retrieval of Leaf-Chlorophyll Dynamics: Model Choice, Sequential Versus Retraining Learning, and Hyperspectral Predictors. *Frontiers in Plant Science*, 13, Article 722442. <https://doi.org/10.3389/fpls.2022.722442>

- [11] 周锦辉. 中国老年人认知功能受损的6年发生风险预测模型建立和验证[D]: [硕士学位论文]. 北京: 中国疾病预防控制中心, 2020.
- [12] 王雅琪, 苏艺伟, 刘移民. 百草枯中毒预后的 Lasso-Logistic 回归分析预测模型的建立[J]. 职业卫生与应急救援, 2022, 40(3): 259-264.
- [13] Ward, A., Arrighi, H.M., Michels, S., *et al.* (2012) Mild Cognitive Impairment: Disparity of Incidence and Prevalence Estimates. *Alzheimer's & Dementia*, **8**, 14-21. <https://doi.org/10.1016/j.jalz.2011.01.002>
- [14] Jia, L., Du, Y., Chu, L., *et al.* (2020) Prevalence, Risk Factors, and Management of Dementia and Mild Cognitive Impairment in Adults Aged 60 Years or Older in China: A Cross-Sectional Study. *The Lancet Public Health*, **5**, E661-E671. [https://doi.org/10.1016/S2468-2667\(20\)30185-7](https://doi.org/10.1016/S2468-2667(20)30185-7)
- [15] 夏艳秋, 崔丽君, 魏丽萍, 等. 综合医院老年住院患者轻度认知功能障碍与焦虑抑郁关系的问卷调查[J]. 中国病案, 2020, 21(3): 83-86.
- [16] 叶青芳, 王旖旎, 李玲, 等. 中青年高血压住院患者轻度认知功能障碍发生现状及影响因素研究[J]. 中国全科医学, 2023, 26(2): 154-159+167.
- [17] Unverzagt, F.W., Ogunniyi, A., Taler, V., *et al.* (2011) Incidence and Risk Factors for Cognitive Impairment No Dementia and Mild Cognitive Impairment in African Americans. *Alzheimer Disease & Associated Disorders*, **25**, 4-10. <https://doi.org/10.1097/WAD.0b013e3181f1c8b1>
- [18] 史路平, 姚水洪, 王薇. 中国老年人群轻度认知障碍患病率及发展趋势的 Meta 分析[J]. 中国全科医学, 2022, 25(1): 109-114.
- [19] 贾从康, 张绍刚, 张美艳, 等. 驻津部队离休干部轻度认知功能障碍的现况调查及影响因素分析[J]. 解放军预防医学杂志, 2020, 38(4): 25-27+30.
- [20] 李强, 林翠茹, 张玉莲. 轻度老年性痴呆患者日常生活能力与中医证候相关性研究[J]. 光明中医, 2023, 38(9): 1617-1620.
- [21] 潘惠英. 金华市社区老年人轻度认知功能障碍的现况调查和干预性研究[D]: [硕士学位论文]. 复旦: 复旦大学, 2012.
- [22] Scholtissen-In De Braek, D.M.J.M., Hurks, P.P.M., Van Boxtel, M.P.J., *et al.* (2011) The Identification of Attention Complaints in the General Population and Their Effect on Quality of Life. *Journal of Attention Disorders*, **15**, 46-55. <https://doi.org/10.1177/1087054709347260>
- [23] Johnson, J.K., Lui, L.-Y. and Yaffe, K. (2007) Executive Function, More than Global Cognition, Predicts Functional Decline and Mortality in Elderly Women. *The Journals of Gerontology: Series A*, **62**, 1134-1141. <https://doi.org/10.1093/gerona/62.10.1134>
- [24] Bronnick, K., Ehrt, U., Emre, M., *et al.* (2006) Attentional Deficits Affect Activities of Daily Living in Dementia-Associated with Parkinson's Disease. *Journal of Neurology, Neurosurgery & Psychiatry*, **77**, 1136-1142. <https://doi.org/10.1136/jnnp.2006.093146>
- [25] De Paula, J.J., Diniz, B.S., Bicalho, M.A., *et al.* (2015) Specific Cognitive Functions and Depressive Symptoms as Predictors of Activities of Daily Living in Older Adults with Heterogeneous Cognitive Backgrounds. *Frontiers in Aging Neuroscience*, **7**, Article 139. <https://doi.org/10.3389/fnagi.2015.00139>
- [26] 杨凤姣, 李艳, 曹仲珍. 社区老年人轻度认知障碍的调查分析[J]. 护理实践与研究, 2010, 7(17): 123-125.
- [27] Deng, Y., Zhao, S., Cheng, G., *et al.* (2021) The Prevalence of Mild Cognitive Impairment among Chinese People: A Meta-Analysis. *Neuroepidemiology*, **55**, 79-91. <https://doi.org/10.1159/000512597>
- [28] 周良佳, 马菲. 老年轻度认知功能障碍影响因素的研究进展[J]. 现代预防医学, 2013, 40(2): 306-311.
- [29] Wang, M., Greenberg, M., Forkert, N.D., *et al.* (2022) Dementia Risk Prediction in Individuals with Mild Cognitive Impairment: A Comparison of Cox Regression and Machine Learning Models. *BMC Medical Research Methodology*, **22**, Article No. 284. <https://doi.org/10.1186/s12874-022-01754-y>
- [30] 骆文, 刘育青, 劳钰钲, 等. 基于 BP 神经网络的阿尔茨海默病预测模型研究[J]. 中华医学图书情报杂志, 2022, 31(1): 32-37.
- [31] 魏月清, 李红, 李芸, 等. ICU 后认知障碍风险预测模型的构建及验证[J]. 中华护理杂志, 2021, 56(1): 14-20.
- [32] 滕菲, 井洪, 王素珍, 等. 中国老年人轻度认知功能障碍风险评估的 Rothman-Keller 模型构建及应用[J]. 现代预防医学, 2023, 50(11): 1932-1938.
- [33] Van Oudenhoven, F.M., Swinkels, S.H.N., Hartmann, T., *et al.* (2022) Modeling the Underlying Biological Processes in Alzheimer's Disease Using a Multivariate Competing Risk Joint Model. *Statistics in Medicine*, **41**, 3421-3433. <https://doi.org/10.1002/sim.9425>

- [34] Zhao, D., Han, X., Huan, J., *et al.* (2023) Forecasting and Analysis of the Effect of Lifestyle on Cognitive Dysfunction Induced by Occupational Aluminum Exposure Based on Bayesian Networks. *Environmental Toxicology and Pharmacology*, **97**, Article 104035. <https://doi.org/10.1016/j.etap.2022.104035>
- [35] Choi, H. and Jin, K.H. (2018) Predicting Cognitive Decline with Deep Learning of Brain Metabolism and Amyloid Imaging. *Behavioural Brain Research*, **344**, 103-109. <https://doi.org/10.1016/j.bbr.2018.02.017>