

多元线性回归融合模型评估学龄前儿童近视发生风险的核心因素与应用

王庆金,张沧霞,葛含笑,马越,卢杰

引用:王庆金,张沧霞,葛含笑,等. 多元线性回归融合模型评估学龄前儿童近视发生风险的核心因素与应用. 国际眼科杂志, 2026,26(11):2018–2025.

基金项目:河北省中医药管理局项目(No.2025169);河北省重点研发计划项目卫生健康创新专项(No.21377730D)

作者单位:(061001)中国河北省沧州市,河北省沧州中西医结合医院眼科

作者简介:王庆金,女,硕士研究生,副主任医师,研究方向:青少年儿童近视防控、屈光手术、顽固性干眼。

通讯作者:张沧霞,女,本科,主任医师,研究方向:青少年儿童近视防控、屈光手术、疑难眼病眼外伤诊治与微创玻璃体视网膜手术. czxiaxia@126.com

收稿日期:2026–05–27 修回日期:2026–09–20

摘要

目的:构建学龄前儿童近视发生风险的评估模型,并评估其在临床应用中的价值。

方法:回顾性分析 2020 年 1 月至 2024 年 12 月在我院建立屈光档案的学龄前儿童资料(选取屈光度较低眼,若双眼屈光度相同,选取右眼),按照 7:3 比例分为训练集与验证集。详细收集患儿各项资料,以儿童等效球镜度数(SE)为因变量,采用多元线性回归模型,筛选对学龄前儿童近视相关生物学指标有影响的指标。随后根据是否发生近视(根据屈光度评估,SE≤-0.50 D 可诊断为近视)将训练集儿童分为近视组与未近视组。随后将多元线性回归模型筛选的影响因素纳入作为评估变量,采用 Logistic 回归构建风险评估方程。通过受试者工作特征(ROC)曲线、Hosmer–Lemeshow 检验及决策曲线,验证评估模型的区分度、校准度及临床应用价值。

结果:本研究纳入学龄前儿童 550 例 550 眼,训练集 385 例 385 眼(男 190 例,女 195 例),验证集 165 例 165 眼(男 82 例,女 83 例)。训练集儿童中近视组 34 例 34 眼(男 16 例,女 18 例),未近视组 351 例 351 眼(男 174 例,女 177 例)。多元线性回归分析显示,眼轴长度(AL)/角膜曲率半径(CR)比值、脉络膜厚度(SFCT)、前房深度(ACD)、年龄、父母近视史、户外活动时间为学龄前儿童近视发生的影响因素(均 $P<0.05$)。根据上述影响因素,构建 Logistic 回归风险评估模型,获得评估模型 $\text{Logit}(P) = 0.894 \times \text{AL}/\text{CR} - 1.312 \times \text{SFCT} + 1.182 \times \text{ACD} + 2.063 \times \text{年龄} + 0.763 \times \text{父母近视史} - 5.868$ 。评估基于训练集所构建的学龄前儿童近视风险评估模型的效能,结果显示,ROC 曲线下面积(AUC)为 0.939(95%CI:0.912–0.965),H–L 检验为 $\chi^2 = 1.487$, $P = 0.993$,决策曲线提示在阈值 0.03–0.81、0.89–0.97 范围内,净收益率>0,提示模型评估学龄前儿童近视风险与实际风险具有较好区分度、校准度及临床应用价值。在独立验证集中进行内部验证,结果显示,ROC 曲线的 AUC 为 0.967(95%CI:0.919–1.000),H–L 检验为 $\chi^2 = 3.818$, $P =$

0.873,决策曲线提示在阈值 0.01–0.99 范围内,净收益率>0,进一步提示模型评估与实际风险具有较好一致性。

结论:年龄、父母近视史、AL/CR、SFCT、ACD 为学龄前儿童近视发生的影响因素,据此构建的 Logistic 回归风险评估模型评估价值较高,可有效评估个体近视发生概率,为个体化预防学龄前儿童近视发生风险提供参考。

关键词:近视;学龄前儿童;多元线性回归;风险评估

DOI:10.3980/j.issn.1672–5123.2026.11.25

The core factors and application of multiple linear regression fusion model in evaluating the risk of myopia in preschool children

Wang Qingjin, Zhang Cangxia, Ge Hanxiao, Ma Yue, Lu Jie

Foundation items: Project of Hebei Provincial Administration of Traditional Chinese Medicine (No.2025169); Health Innovation Special Project of Hebei Provincial Key Research and Development Program (No.21377730D)

Department of Ophthalmology, Cangzhou Hospital of Integrated Traditional Chinese and Western Medicine, Cangzhou 061001, Hebei Province, China

Correspondence to: Zhang Cangxia. Department of Ophthalmology, Cangzhou Hospital of Integrated Traditional Chinese and Western Medicine, Cangzhou 061001, Hebei Province, China. czxiaxia@126.com

Received:2026–05–27 Accepted:2026–09–20

Abstract

• **AIM:** To construct an assessment model for the risk of myopia in preschool children and evaluate its clinical value.

• **METHODS:** A retrospective analysis was conducted on the data of preschool children who established refractive profiles in our hospital from January 2020 to December 2024. The eye with lower refractive error was selected. The right eye was chosen in cases where the refractive error of both eyes was identical. The data were divided into a training set ($n=385$) and a validation set ($n=165$) in a 7:3 ratio. The data of the children were collected in detail, and the children’s spherical equivalent (SE) was used as the dependent variable. The multiple linear regression model was used to screen the indicators that affected the biological indicators related to myopia in preschool children. Subsequently, the children in the training set were divided into a myopia group and a normal vision

group based on whether myopia occurred (SE ≤ -0.50 D can be diagnosed as myopia according to refractive power assessment). Subsequently, the influencing factors selected by the multiple linear regression model were included as predictive variables, and a risk assessment equation was constructed using Logistic regression. The discrimination, calibration and clinical application value of the assessment model were verified through the receiver operating characteristic (ROC) curve, Hosmer - Lemeshow test and decision curve.

• **RESULTS:** This study included 550 preschool children (550 eyes). The training set comprised 385 children (385 eyes; 190 boys and 195 girls), and the validation set comprised 165 children (165 eyes; 82 boys and 83 girls). Among the children in the training set, the myopia group included 34 children (34 eyes; 16 boys and 18 girls), and the non-myopic group included 351 children (351 eyes; 174 boys and 177 girls). Multiple linear regression analysis showed that the ratio of axial length (AL) to corneal curvature radius (CR), subfoveal choroidal thickness (SFCT), anterior chamber depth (ACD), age, parental myopia history, and outdoor activity time were risk factors for myopia in preschool children (all $P < 0.05$). Based on the above risk factors, a Logistic regression risk assessment model was constructed, and the evaluation model Logit (P) was obtained as $0.894 \times AL/CR - 1.312 \times SFCT + 1.182 \times ACD + 2.063 \times age + 0.763 \times parental\ myopia\ history - 5.868$. The effectiveness of the myopia risk assessment model for preschool children based on the training set was evaluated. The results showed that the area under the ROC curve (AUC) was 0.939 (95% CI: 0.912-0.965), the H-L test was $\chi^2 = 1.487$, $P = 0.993$, and the decision curve indicated that within the threshold range of 0.03-0.81 and 0.89-0.97, the net return rate was greater than 0. This suggested that the model had good discrimination, calibration, and clinical application value in evaluating the myopia risk and actual risk of preschool children. Internal validation was conducted in an independent validation set, and the results showed that the AUC of the ROC curve was 0.967 (95% CI: 0.919-1.000), the H-L test was $\chi^2 = 3.818$, $P = 0.873$, and the decision curve indicated that within the threshold range of 0.01-0.99, the net return rate was greater than 0, further suggesting that the model assessment had good consistency with the actual risk.

• **CONCLUSION:** Age, parental history of myopia, AL/CR, SFCT, and ACD are risk factors for myopia in preschool children. The Logistic regression risk assessment model constructed based on these factors has high evaluation value and can effectively evaluate the probability of myopia occurrence in individuals, providing reference for individualized prevention of myopia risk in preschool children.

• **KEYWORDS:** myopia; preschool children; multiple linear regression; risk assessment

Citation: Wang QJ, Zhang CX, Ge HX, et al. The core factors and application of multiple linear regression fusion model in evaluating the risk of myopia in preschool children. Guoji Yanke Zazhi(Int Eye Sci), 2026,26(11):2018-2025.

0 引言

近年来,儿童青少年近视患病率持续攀升,呈现低龄化、高发化的严峻趋势^[1]。学龄前阶段是眼球屈光状态发育的关键期,该阶段出现的远视储备不足或过早向近视漂移,往往预示着后续高度近视及相关并发症风险的显著增加^[2-3]。因此,早期识别学龄前儿童近视发生的高危因素,构建精准、个体化的风险评估模型,对于延缓近视发生、降低最终近视程度具有重要的公共卫生与临床意义。目前,国内外关于近视评估模型的研究多聚焦于学龄期儿童及青少年,且采用的建模方法多以 Logistic 回归为主^[4-5]。然而,学龄前儿童屈光状态尚处于动态发育之中,其近视发生机制与学龄儿童存在差异,单纯依赖单一指标或传统回归模型往往难以充分捕捉各因素间的复杂交互作用^[6]。此外,现有模型较少同时融合多元线性回归筛选关键生物学参数,并进一步构建风险评估方程,导致模型的解释力受限。基于此,本研究通过回顾性分析 550 例学龄前儿童的临床资料,采用多元线性回归筛选对近视相关生物学指标有影响的核心因素,进而将筛选出的影响因素纳入 Logistic 回归,构建风险评估融合模型,并通过受试者工作特征(ROC)曲线、校准曲线及决策曲线对模型的区分度、校准度和临床净收益进行内部验证。旨在为学龄前儿童近视风险的早期识别提供一种高效、可量化的评估工具,助力制定个体化预防策略。

1 对象和方法

1.1 对象 回顾性分析 2020 年 1 月至 2024 年 12 月在我院建立屈光档案的学龄前儿童资料(选取屈光度较低眼,若双眼屈光度相同,选取右眼),按照 7:3 比例,采用随机数字表法将患儿资料分为训练集与验证集。纳入标准:(1)年龄为 5-6 岁;(2)认知功能正常,已配合完成眼部生物学测量;(3)就诊前未接受过任何近视干预治疗,包括但不限于低浓度阿托品滴眼液、角膜塑形镜、周边离焦框架镜或软性接触镜,且无系统性使用影响屈光发育的药物史;(4)各项检查资料及基础资料均记录完整。排除标准:(1)患有先天性眼部疾病,如先天性白内障、先天性青光眼、视网膜病变等;(2)存在眼部外伤史或手术史;(3)患有影响屈光状态的全身性疾病,如糖尿病、甲状腺功能异常等;(4)近期使用过可能影响眼部发育的药物或光学干预手段;(5)检查数据存在明显缺失。经医院医学伦理委员会审核豁免伦理批准及知情同意。

1.2 方法

1.2.1 样本量计算 根据样本量计算公式 $n = Z_{\alpha}^2 p(1-p) / \delta^2$ 计算样本量, n 代表纳入学龄前儿童的总样本量, P 代表学龄前儿童近视的发生率,查阅相关文献^[7],学龄前儿童近视的发生率约 4.40%; α 取 0.05,则双侧检验 Z_{α} 取值 1.96, δ 代表总体比例估计值的容许误差,本研究中取 0.025,得到纳入的总样本量至少为 259 例。

1.2.2 临床资料收集 查阅医院电子病历系统,详细收集纳入研究儿童的基本信息、眼部生物学参数。基本信息包括性别、年龄、户外活动时间(每日 <2 h,每日 ≥2 h)、父母近视史(双近视父母、单近视父母、父母均无近视)。眼部生物学参数通过标准化流程测量,包括眼轴长度(axial length,AL)/角膜曲率半径(CR)、AL、脉络膜厚度(subfoveal choroidal thickness,SFCT)、前房深度(ACD)、等效球镜度(spherical equivalent,SE)。其中 AL/CR、AL、ACD 均为 IOL Master 500 光学生物测量仪(Carl Zeiss

Meditec AG) 测量, SFCT 为光学相干断层扫描(OCT) 仪(Cirrus HD – OCT 5000) 测量, SE 由自动电脑验光仪(Topcon KR-8800) 在睫状肌麻痹下(复方托吡卡胺滴眼液, 每 5 min 1 次, 连续点 4 次, 20 min 验光) 验光获取(SE=球镜度数+1/2 柱镜度数)。所有测量均由 2 名经验丰富的眼科医师完成, 对同一指标测量 3 次, 取平均值作为最终结果, 若两次测量结果差异超过 0.05 mm(AL、CR、ACD) 或 5 μm(SFCT), 则进行第 4 次测量, 取 4 次测量结果的中位数。

1.2.3 诊断标准与分组 将散瞳后检查 SE≤-0.50 D 作为学龄前儿童近视的诊断界值^[8]。根据此标准, 将训练集学龄前儿童资料分为两组: 近视组(SE≤-0.50 D) 和未近视组(SE>-0.50 D)。分组过程由 2 名经过培训的研究人员独立进行, 对纳入儿童的 SE 值进行判定并划分组别, 若出现分组不一致的情况, 则通过讨论或由第三位资深眼科医师复核确定最终分组, 以确保分组的准确性和一致性。

统计学分析: 使用 SPSS23.0 软件进行数据处理。计量资料均经 K-S 检验, 符合正态分布的计量资料以 $\bar{x} \pm s$ 表示, 两组间比较采用独立样本 *t* 检验, 多组间比较采用单因素方差检验; 计数资料以 *n*(%) 表示, 组间比较用 χ^2 检验。采用 Pearson 相关性分析各连续变量与 SE 的相关性, 采用点二列相关性分析年龄、户外活动时间与 SE 的相关性, 采用卡方检验 Cramer's V 表示父母近视史与 SE 的关联强度, 采用多元线性回归模型筛选对学龄前儿童近视相关生物学指标有影响的指标。将多元线性回归模型筛选的影响因素作为评估变量, 采用 Logistic 回归构建风险评估方程。通过 ROC 曲线、Hosmer-Lemeshow(H-L) 检验及决策曲线, 验证评估模型的区分度、校准度及临床应用价值。 *P*<0.05 为差异有统计学意义。

2 结果

2.1 训练集与验证集学龄前儿童临床资料比较 本研究

纳入学龄前儿童 550 例 550 眼, 训练集 385 例 385 眼, 验证集 165 例 165 眼。训练集与验证集组间年龄、性别、户外活动时间等临床资料比较, 差异无统计学意义(均 *P*>0.05), 见表 1。

2.2 训练集中近视与未近视学龄前儿童临床资料比较 训练集学龄前儿童中近视组 34 例 34 眼, 未近视组 351 例 351 眼。近视与未近视学龄前儿童年龄、户外活动时间、父母近视史、AL、AL/CR、SFCT、ACD、SE 比较, 差异均有统计学意义(*P*<0.001), 两组性别对比差异无统计学意义(*P*>0.05), 见表 2。

2.3 学龄前儿童 SE 与相关因素的相关性分析 对 SE 与 AL、AL/CR、SFCT、ACD 进行 Pearson 相关性分析, 得出 AL、AL/CR、SFCT、ACD 与 SE 有相关性(均 *P*<0.05); 对 SE 与年龄、户外活动时间进行点二列相关性分析, 采用卡方检验 Cramer's V 分析父母近视史与 SE 相关性, 显示年龄、户外活动时间、父母近视史与 SE 有相关性(均 *P*<0.05), 见表 3, 各连续变量与 SE 的散点图见图 1。

2.4 LASSO 回归分析 将表 2 中差异有统计学意义的指标及表 3 中与 SE 相关的连续变量纳入 LASSO 回归分析, 结果显示, 随着惩罚系数 λ 的增大, 纳入变量的回归系数逐渐被压缩, 部分系数被压缩至 0。采用十折交叉验证法确定最优 λ 值, 当 λ 为 0.023, 最终筛选出年龄、户外活动时间、父母近视史、AL/CR、SFCT、ACD 共 6 个非零系数的特征变量, 提示上述 6 个指标是学龄前儿童近视发生的影响因素, 见图 2。

2.5 影响学龄前儿童近视相关生物学指标的多元线性回归分析 以学龄前儿童近视相关生物学指标(SE) 为因变量, 将表 2 中差异有统计学意义的指标及表 3 中与 SE 相关的连续变量纳入行多元线性回归分析, 结果显示, 年龄、父母近视史、户外活动时间、AL/CR、SFCT、ACD 为影响学龄前儿童近视相关生物学指标的因素(均 *P*<0.05), 见表 4。

表 1 训练集与验证集学龄前儿童临床资料比较

项目	训练集(<i>n</i> = 385)	验证集(<i>n</i> = 165)	χ^2/t	<i>P</i>
年龄(例, %)			0.018	0.894
5 岁	154(40.0)	67(40.6)		
6 岁	231(60.0)	98(59.4)		
性别(例, %)			0.006	0.941
男	190(49.4)	82(49.7)		
女	195(50.6)	83(50.3)		
户外活动时间(例, %)			0.000	0.984
每日<2 h	250(64.9)	107(64.8)		
每日≥2 h	135(35.1)	58(35.2)		
父母近视史(例, %)			0.004	0.998
双近视父母	30(7.8)	13(7.9)		
单近视父母	120(31.2)	51(30.9)		
父母均无近视	235(61.0)	101(61.2)		
AL($\bar{x} \pm s$, mm)	20.55±0.97	20.61±0.97	0.696	0.486
AL/CR($\bar{x} \pm s$)	2.94±0.12	2.96±0.15	1.896	0.058
SFCT($\bar{x} \pm s$, μm)	320.71±40.57	324.29±39.61	0.956	0.340
ACD($\bar{x} \pm s$, mm)	2.79±0.29	2.84±0.28	1.798	0.073
SE($\bar{x} \pm s$, D)	0.24±0.75	0.25±0.74	0.275	0.784

表 2 训练集中近视与未近视学龄前儿童临床资料比较

项目	近视组 (n=34)	未近视组 (n=351)	χ^2/t	P
年龄(例,%)			18.088	<0.001
5岁	2(5.9)	152(43.3)		
6岁	32(94.1)	199(56.7)		
性别(例,%)			0.078	0.780
男	16(47.1)	174(49.6)		
女	18(52.9)	177(50.4)		
户外活动时间(例,%)			13.949	<0.001
每日<2 h	32(94.1)	218(62.1)		
每日≥2 h	2(5.9)	133(37.9)		
父母近视史(例,%)			29.680	<0.001
双近视父母	10(29.4)	20(5.7)		
单近视父母	14(41.2)	106(30.2)		
父母均无近视	10(29.4)	225(64.1)		
AL($\bar{x}\pm s$,mm)	21.65±1.10	20.44±0.89	7.441	<0.001
AL/CR($\bar{x}\pm s$)	3.07±0.10	2.93±0.11	6.972	<0.001
SFCT($\bar{x}\pm s$,μm)	275.32±30.67	325.11±38.69	7.280	<0.001
ACD($\bar{x}\pm s$,mm)	3.15±0.15	2.75±0.28	13.276	<0.001
SE($\bar{x}\pm s$,D)	-1.04±0.30	0.36±0.66	22.694	<0.001

表 3 学龄前儿童 SE 与相关因素的相关性分析

统计值	AL	AL/CR	SFCT	ACD	年龄	户外活动时间	父母近视史
r/V	-0.719	-0.655	0.832	-0.854	-0.369	0.343	0.272
P	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.003

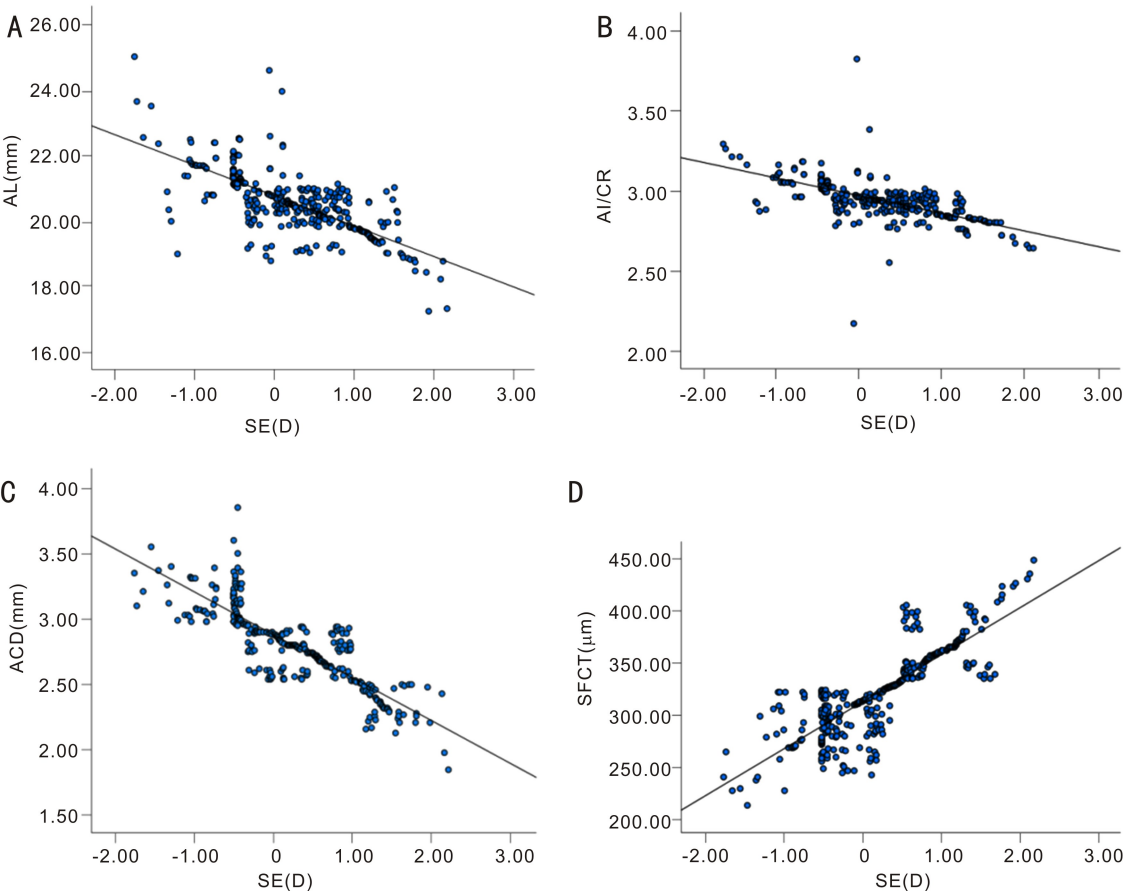


图 1 SE 与 AL、AL/CR、ACD、SFCT 的散点图 A:AL 与 SE 的相关性;B:AL/CR 与 SE 的相关性;C:ACD 与 SE 的相关性;D:SFCT 与 SE 的相关性。

2.6 Logistic 回归算法构建学龄前儿童近视风险的评估模型 根据多元线性回归筛选得到的影响因素纳入作为评估变量,融合构建学龄前儿童近视发生风险评估模型,采用 Logistic 回归算法拟合风险评估方程,获得评估模型 $\text{Logit}(P) = 0.894 \times \text{AL/CR} - 1.312 \times \text{SFCT} + 1.182 \times \text{ACD} + 2.063 \times \text{年龄} + 0.763 \times \text{父母近视史} - 5.868$,见表 5。

2.7 基于训练集构建的学龄前儿童近视风险评估模型的拟合效果评价 在训练集中,构建的风险评估模型 ROC 曲线的曲线下面积 (AUC) 为 0.939 (95% CI: 0.912 - 0.965), H-L 检验为 $\chi^2 = 1.487, P = 0.993$,决策曲线提示,在阈值 0.03-0.81、0.89-0.97 范围内,净收益率>0,最大净

收益为 0.080,见图 3。十折交叉验证重复 100 次后,模型平均预测准确率为 91.74%,平均 AUC 为 0.914,仅较原结果下降 0.025,模型不存在明显的过拟合,提示模型预测具有良好精准度。进一步采用 L2 正则化对模型参数进行约束,将参数权重的 L2 范数加入模型损失函数限制参数过度拟合样本噪声,调整后参数模型的 C-index 为 0.937, AUC 为 0.937,仍保持较高的预测效能,说明参数调整后未出现预测能力大幅下降,原模型结果稳定。采用 Bootstrap 重抽样法重复抽样 1 000 次,评估模型校准度,结果显示,模型的校准曲线斜率为 1.000,校准曲线与理想曲线相近,模型校准度良好。

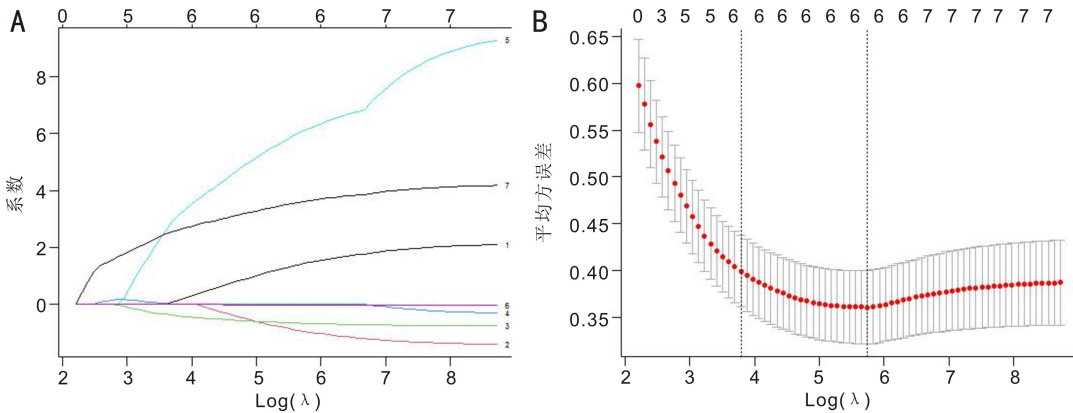


图2 LASSO 回归分析筛选变量图 A:系数剖面图;B:交叉验证曲线。

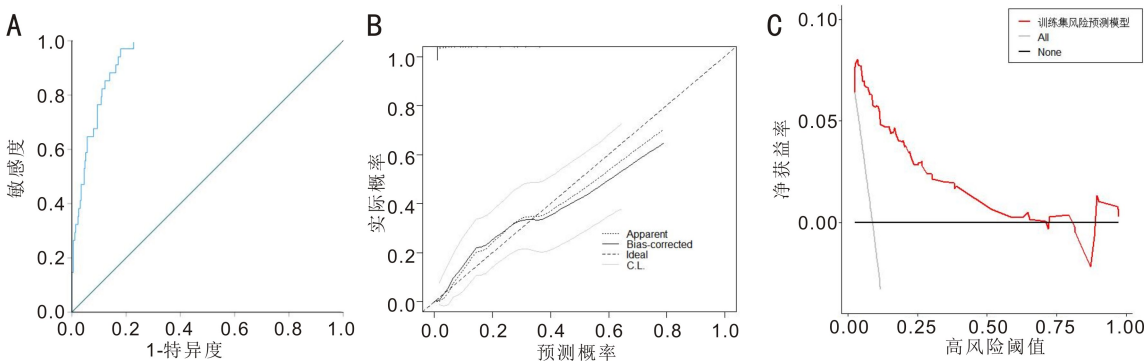


图3 训练集 ROC 曲线和校准曲线及决策曲线 A:ROC 曲线;B:校准曲线;C:决策曲线。

表 4 影响学龄前儿童近视相关生物学指标的多元线性回归分析

变量	非标准化回归系数(B)	B 的标准误差	标准回归系数(β)	t	P	VIF	R ²	校正 R ²
年龄	-0.079	0.031	-0.052	-2.541	0.011	0.816	0.873	0.871
户外活动时间	0.092	0.032	0.058	2.820	0.005	0.784		
父母近视史	0.106	0.022	0.091	4.771	<0.001	0.933		
AL/CR	-0.971	0.144	-0.156	-6.761	<0.001	0.632		
SFCT	0.007	0.000	0.388	14.415	<0.001	0.466		
ACD	-1.115	0.072	-0.433	-15.479	<0.001	0.430		

表 5 学龄前儿童近视风险的 Logistic 回归分析结果

变量	β	SE	Wald χ ²	P	OR	95% CI
年龄	2.063	1.030	4.008	0.045	7.868	1.044-59.291
户外活动时间	1.413	0.865	2.667	0.102	4.109	0.754-22.402
父母近视史	0.763	0.300	6.478	0.011	2.145	1.192-3.859
AL/CR(标准化值)	0.894	0.272	10.843	0.001	2.446	1.436-4.164
SFCT(标准化值)	-1.312	0.390	11.303	0.001	0.269	0.125-0.578
ACD(标准化值)	1.182	0.340	12.096	0.001	3.260	1.675-6.346

2.8 基于验证集对学龄前儿童近视风险评估模型的内部验证 将训练集构建的风险评估模型应用于验证集进行内部验证,结果显示,ROC 曲线的 AUC 为 0.967(95%CI: 0.919–1.000),H-L 检验为 $\chi^2=3.818,P=0.873$,决策曲线提示,在阈值 0.01–0.99 范围内,净收益率>0,最大净收益为 0.070,见图 4。十折交叉验证重复 100 次后,模型平均预测准确率为 96.74%,平均 AUC 为 0.928,仅较原结果下降 0.039,模型不存在明显的过拟合,提示模型评估学龄前儿童近视风险具有良好精准度。进一步采用 L2 正则化对模型参数进行了约束,将参数权重的 L2 范数加入模型损失函数限制参数过度拟合样本噪声,调整后参数模型的 C-index 为 0.968,曲线下面积为 0.968,仍保持较高的预测效能,说明参数调整后未出现预测能力大幅下降,原模型结果稳定。采用 Bootstrap 重抽样法重复抽样 1 000 次,评估模型校准度,结果显示,模型的校准曲线斜率为 1.000,校准曲线与理想曲线相近,模型校准度良好。

3 讨论

学龄前儿童处于视觉发育的关键期和敏感期,眼睛接受约 80%的外界信息,近视将直接影响儿童对周围环境的感知,且会导致注意力不集中、学习兴趣下降,直接降低学习效率^[9]。同时,低龄配戴眼镜可能影响儿童的性格发展,导致自卑心理产生,阻碍社交互动及自尊心的建立,可能引发焦虑、沮丧等心理问题^[10]。本研究中,训练集 385 例学龄前儿童,有 34 例儿童近视,发生率为 8.8%,与陈爱菊等^[11]研究中 9.8%(87/920)的屈光异常率相近。由此可见,近视已成为影响学龄前儿童视觉健康的公共卫生问题,早期识别近视高危个体并采取针对性干预措施,是延缓近视发生、控制近视进展的关键。

SE 是将散光折算成等效球镜后,用于综合评估眼球整体屈光状态的指标^[12]。本研究以 SE 为因变量,采用多元线性回归分析影响学龄前儿童近视相关生物学指标的因素,结果显示,年龄、父母近视史、AL/CR、SFCT、ACD、户外活动时间均为学龄前儿童近视相关生物学指标的影响因素,分析其原因:(1)年龄是影响学龄前儿童 SE 的重要因素,6 岁儿童的 SE 显著低于 5 岁儿童,这与学龄前儿童眼球发育的生理特点密切相关。随着年龄增长,儿童眼球逐渐发育,AL 不断增加,而角膜曲率、晶状体屈光力等代偿性变化则相对滞后,从而导致屈光状态更易逐渐向近视方向发展^[13]。(2)许晨等^[7]研究发现,父母近视为学龄前儿童近视的影响因素,本研究结果与之一致。父母近视史作为遗传因素,在学龄前儿童近视发生中扮演关键角色。双近视父母的儿童具有明显的家族聚集性,遗传因素可能

通过影响眼球的发育模式,如 AL 增长速度、角膜曲率形态等,增加儿童近视的易感性^[14]。(3)AL 是反映眼球前后径的重要参数,AL 越长,意味着眼球的前后径过度增长,AL/CR 则综合考虑了 AL 与 CR 的关系,能更全面地反映眼球的屈光状态^[15–16]。该比值增大,表明眼球的屈光力相对增强,从而更容易发生近视^[17]。(4)SFCT 即黄斑区视网膜下脉络膜厚度,脉络膜作为眼球壁的重要组成部分,其厚度变化可能通过影响巩膜的营养供应和代谢,进而影响 AL 的增长^[18]。SFCT 越小,表明脉络膜越薄,意味着其抑制 AL 过度生长的能力下降,从而允许 AL 进一步延长,导致近视发生或加深^[19]。(5)ACD 为前房深度,AL 异常延长时,睫状体被动后移,通过悬韧带牵拉晶状体–虹膜隔向后移位,导致 ACD 代偿性加深。因此,ACD 增大本质上是眼球后段轴向扩张的前段结构标记。在学龄前儿童中,ACD 若显著超过同龄正常范围,提示晶状体屈光力代偿已接近极限,后续 AL 进一步增长将直接表现为近视进展^[20]。值得注意的是,单因素分析及 LASSO 回归初步显示户外活动时间与学龄前儿童近视相关,但纳入 Logistic 回归模型后,未达到统计学显著水平,最终未能进入风险评估方程。这一结果提示,户外活动时间对学龄前儿童近视的影响,可能部分通过 AL、SFCT 等结构性生物学参数起中介传导作用,当模型中已包含 AL/CR、SFCT 等直接反映眼球结构改变的强相关变量后,户外活动时间作为行为学变量的效应被削弱。

然而,本研究中单一的多元线性回归模型对 SE 变异的解释能力有限,可能受到未纳入模型的其他混杂因素或复杂交互作用的影响,但该分析结果为后续构建融合模型提供了关键的评估变量。基于多元线性回归筛选出的核心影响因素,本研究构建了 Logistic 回归风险评估模型,得到预警模型 $\text{Logit}(P) = 0.894 \times \text{AL/CR} - 1.312 \times \text{SFCT} + 1.182 \times \text{ACD} + 2.063 \times \text{年龄} + 0.763 \times \text{父母近视史} - 5.868$ 。同时,通过训练集和验证集的内部验证,该模型展现出良好的区分度、校准度和临床净收益。ROC 曲线分析显示,模型的 AUC 值在训练集和验证集中均较高,分别为 0.939(95%CI:0.912–0.965)、0.967(95%CI:0.919–1.000),表明模型在本中心人群中具有较好的区分能力,但该结果需经外部数据进一步验证。H-L 检验结果提示模型的评估概率与实际发生概率之间具有较好的一致性,校准曲线拟合良好。决策曲线分析则证实,在一定的阈值概率范围内,应用该模型进行学龄前儿童近视风险评估可获得正的净收益,具有一定的临床实用价值。与顾泉成等^[21]基于 1 759 名中小学生(含初高中)的大样本机器学习预测模型

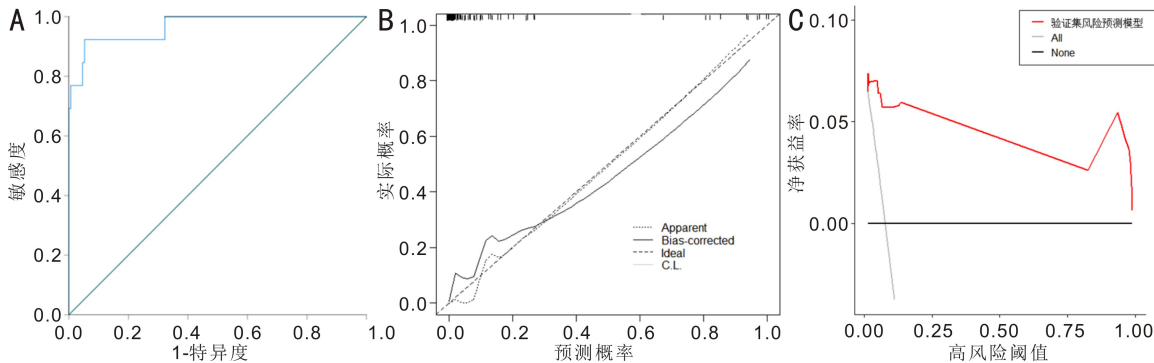


图 4 验证集 ROC 曲线和校准曲线及决策曲线 A:ROC 曲线;B:校准曲线;C:决策曲线。

不同,本研究聚焦于更早期的5-6岁学龄前儿童,且研究视角并非构建面向广泛人群的预测模型,而是探索一种基于常规非侵入性参数的近视风险初筛策略,在缺乏系统检查条件的基层场景中实现高危儿童的有效分流。故而,本模型可作为基层医疗机构视力筛查的辅助工具,通过非侵入性参数快速评估儿童近视风险,对于模型判定为高风险的儿童,建议转诊至眼科专科,由专业医师进一步实施散瞳验光等精标准检查以明确诊断,并根据最终确诊结果制定个体化随访或干预计划。对于低风险儿童,则可在一定程度上减少不必要的散瞳操作,优化医疗资源配置。但值得关注的是,本模型当前仅为初步探索性内部验证,其不同筛查场景中的适用性和成本效益尚需前瞻性研究进一步验证,且本模型不应作为独立诊断依据,临床决策始终应以专科检查结果为准。

值得关注的是,本研究在分析眼的选取上仅纳入单眼且未区分主导眼与非主导眼,可能对结果产生一定影响。现有文献关于主导眼在近视发生与进展中的作用尚存争议:部分研究提示在屈光参差性近视人群中,主导眼的近视程度更深、AL更长^[22]。本研究纳入的学龄前儿童双眼屈光参差较小,在双眼屈光度接近的情况下,主导眼对近视相关生物学指标的独立影响相对有限。但考虑到主导眼与屈光状态的关系在学龄前儿童中尚缺乏系统研究,本研究未进行主导眼分层分析仍属方法学局限,未来研究应进一步探讨主导眼在学龄前儿童近视发生中的作用。

本研究的创新性主要体现在:在研究视角上,聚焦5-6岁学龄前这一屈光高度可塑但研究相对薄弱的年龄段,填补低龄儿童近视风险评估工具的不足;在方法上,创新性先以多元线性回归筛选影响SE的生物学变量,再纳入Logistic回归构建分类评估方程,既保留连续变量间的线性数量关系,又实现近视发生风险的二分类概率输出,较单一回归模型具有更强的解释力和临床可操作性。在临床应用层面,本模型可作为基层医疗机构视力筛查的辅助决策工具,对于无条件开展散瞳验光的社区卫生中心或幼儿园保健室,可通过代入上述参数快速评估个体近视发生风险。

综上所述,年龄、父母近视史以及AL/CR、SFCT、ACD为学龄前儿童近视的影响因素,基于上述多元线性回归结果融合Logistic回归算法构建的评估模型 $\text{Logit}(P) = 0.894 \times \text{AL/CR} - 1.312 \times \text{SFCT} + 1.182 \times \text{ACD} + 2.063 \times \text{年龄} + 0.763 \times \text{父母近视史} - 5.868$,能够较为准确地评估学龄前儿童近视发生风险,可为临床医师早期识别近视高危儿童提供一种客观、量化的工具,有助于实现个体化的近视防控策略。然而,本研究仍存在一定局限性:(1)本研究为单中心回顾性研究,样本量相对有限,Logistic回归模型共纳入7个评估变量,而训练集内近视事件数为34例,低于传统推荐的 $\text{EPV} \geq 10$ 标准,样本量在模型参数估计方面可能存在一定不足,虽采用LASSO惩罚回归可缓解过拟合,但惩罚参数 λ 的选择依赖于交叉验证,其结果受样本随机划分影响较大,无法弥补因样本量不足所致的预测效能精确度欠缺。(2)模型中未纳入饮食、阅读时间、屏幕暴露时长、角膜散光等潜在影响因素,且从“连续测量”与“分

类决策”两个维度予以区分,但此种两步法在统计本质上仍不可避免地存在信息重叠,由此带来的过拟合风险尤为突出。(3)本研究仅采用单中心同源数据的内部拆分验证,未纳入独立外部队列,模型泛化能力尚不明确,加之横断面设计缺乏时序性,不能确立预测因子与近视发生之间的因果关联,亦无法追踪屈光动态演变轨迹,结论仅具探索性提示意义。(4)本研究仅纳入5、6岁学龄前儿童,未覆盖其他年龄段的屈光发育关键期,因此所得结论是否适用于低龄幼儿或学龄期青少年尚不清楚,年龄外推性有限。未来可开展多中心、前瞻性研究,进一步扩大样本量和拓展年龄范围,纳入更多相关变量,以优化模型结构,提高评估效能。同时,还可考虑采用岭回归替代普通线性回归,并结合1a随访数据进一步更新系数,动态优化评估模型。

利益冲突声明:本文不存在利益冲突。

作者贡献声明:王庆金、葛含笑论文选题与修改,初稿撰写;马越、卢杰文献检索,数据分析;张沧霞选题指导,论文修改及审阅。所有作者阅读并同意最终的文本。

参考文献

[1] Liang JH, Pu YQ, Chen JQ, et al. Global prevalence, trend and projection of myopia in children and adolescents from 1990 to 2050: a comprehensive systematic review and meta-analysis. *Br J Ophthalmol*, 2025,109(3):362-371.

[2] 焦永红, 卢海, 麻婧, 等. 重视儿童早发性高度近视眼致盲性并发病的管理. *中华眼科杂志*, 2025,61(1):7-11.

[3] Xu HY, Zong CX, Pu JN, et al. Reduction of hyperopia and accelerated myopia progression in preschool children: a four-year longitudinal study. *Optom Vis Sci*, 2026,103(2):e70028.

[4] Mu JF, Zhong HX, Jiang MJ, et al. Development of a nomogram for predicting myopia risk among school-age children: a case-control study. *Ann Med*, 2024,56:2331056.

[5] 李婷, 王梅, 张东红, 等. 沈阳市12~18岁青少年近视相关因素的探讨及Nomogram预测模型的建立. *中国卫生统计*, 2022,39(2):238-242.

[6] Xu SS, Li LL, Han WJ, et al. Association between myopia and pupil diameter in preschoolers: evidence from a machine learning approach based on a real-world large-scale dataset. *Ophthalmol Ther*, 2024,13(7):2009-2022.

[7] 许晨, 王勤美, 许建平, 等. 学龄前儿童近视的影响因素分析及低浓度阿托品应用效果. *中国妇幼保健*, 2022,37(21):3984-3987.

[8] 杨培增, 范先群, 孙兴怀, 等. *眼科学*(第9版). 北京:人民卫生出版社, 2018:74-75.

[9] 周亮, 师春立, 陈剑宇. 四川省儿童青少年近视现状及其影响因素研究. *华南预防医学*, 2023,49(4):407-411.

[10] Wang Y, Li L, Guo GL, et al. Investigation and analysis of the status of myopia and related behavior in children and adolescents in Northeast Sichuan. *Indian J Ophthalmol*, 2024,72(Suppl 3):S415-S422.

[11] 陈爱菊, 杨友谊, 张芳芳. 3~6岁儿童屈光异常的影响因素及临床干预对策研究. *中国妇幼保健*, 2020,35(6):1075-1078.

[12] Zhang XY, Fan Q, Zhang FY, et al. Gene-environment interaction in spherical equivalent and myopia: an evidence-based review. *Ophthalmic Epidemiol*, 2022,29(4):435-442.

[13] 项凯迪, 王菁菁, 潘臣炜, 等. 上海地区 4~14 岁儿童 1 年屈光进展及影响因素分析. 中国学校卫生, 2022, 43 (9): 1309-1313,1318.

[14] Qin X, Sun YY, Wang SN, et al. Risk factors for ocular biological parameters in Chinese preschool children: a cohort study from the Beijing whole childhood eye study. Front Med, 2025,12:1510124.

[15] Tañá-Rivero P, Tañá-Sanz S, Pastor-Pascual F, et al. Axial length measurement failure rates using optical biometry based on swept-source OCT in cataractous eyes. Expert Rev Med Devices, 2022,19(8): 633-640.

[16] Nakai Y, Hieda O, Nakamura Y, et al. Biometric risk factors for myopia onset in emmetropic school-age children. Jpn J Ophthalmol, 2025,69(5):687-693.

[17] 姚远, 许韶君, 苏娱, 等. 非体育类校外辅导与小学生筛查性近视及眼轴长度的关联. 中国学校卫生, 2025,46(11):1544-1548.

[18] Larsen MP, Iqbal H, Grauslund J, et al. Subfoveal choroidal thickness in pregnancy: a systematic review with meta-analyses. Acta Ophthalmol, 2025,103(5):493-504.

[19] 张海燕, 李洪霞, 单石蕾, 等. OCT 测量脉络膜厚度在学龄儿童近视进展中的研究. 中国斜视与小儿眼科杂志, 2024,32(4): 24-26,后插 1.

[20] 黄瑞琴, 杨丽霞, 司静彬, 等. 近视性屈光参差儿童周边离焦与眼生物参数的相关性. 中华眼视光学与视觉科学杂志(中英文), 2025,27(8):586-593.

[21] 顾泉成, 陈新丽, 陈曦, 等. 基于机器学习的中小学生近视影响因素分析及预测模型的开发与验证. 国际眼科杂志, 2025,25(2): 328-336.

[22] Robert R, Babu M, Unnikannan K. Sighting dominance, biometric parameters, and refractive status analyzing the role of ocular dominance. Beyoglu Eye J, 2023,8(2):104-109.