

# 构建中晚期早产儿严重脑室内出血的预测模型

张文清, 穆 锴, 平天越, 张淑敏, 宗鲁洁, 薛 江\*

山东大学齐鲁第二医院儿童医学中心, 山东 济南

收稿日期: 2026年1月12日; 录用日期: 2026年2月6日; 发布日期: 2026年2月24日

## 摘 要

目的: 严重脑室内出血(Severe Intraventricular Hemorrhage, sIVH)是早产儿死亡及神经发育障碍的重要原因。现有预测模型多针对胎龄32周以下的极早产儿, 对中晚期早产儿存在研究空白。本研究旨在构建机器学习模型, 以实现对该群体sIVH的精准早期预测。方法: 共纳入1831例中晚期早产儿, 收集其出生前及出生后48小时内的临床数据。采用多种算法进行特征选择, 比较了16种预测模型的性能, 评估并优化最佳模型。结果: 在纳入的患儿中, 共确诊32例sIVH (1.7%)。最佳模型为朴素贝叶斯模型。高C反应蛋白水平、入院后48小时之内应用儿茶酚胺类药物以及生后立即接受气管插管是sIVH的独立危险因素。结论: 本研究构建了一个机器学习预测模型, 帮助临床医生在中晚期早产儿生命早期识别sIVH高风险个体, 改善预后, 降低死亡率。

## 关键词

中晚期早产儿, 重度脑室内出血, 预测模型, 机器学习, SHAP

# Development of a Prediction Model for Severe Intraventricular Hemorrhage in Moderate-to-Late Preterm Infants

Wenqing Zhang, Kai Mu, Tianyue Ping, Shumin Zhang, Lujie Zong, Jiang Xue\*

Children's Medical Center, The Second Qilu Hospital of Shandong University, Jinan Shandong

Received: January 12, 2026; accepted: February 6, 2026; published: February 24, 2026

## Abstract

**Objective:** Severe Intraventricular Hemorrhage (sIVH) is a major cause of death and neurodevelopmental

\*通讯作者。

文章引用: 张文清, 穆锴, 平天越, 张淑敏, 宗鲁洁, 薛江. 构建中晚期早产儿严重脑室内出血的预测模型[J]. 临床医学进展, 2026, 16(2): 3145-3155. DOI: 10.12677/acm.2026.162727

**impairment in preterm infants. Existing prediction models primarily focus on very preterm infants under 32 weeks of gestation, leaving a research gap for moderate-to-late preterm infants. This study aims to utilize early multimodal data to construct a machine learning model for accurate and early prediction of sIVH in this specific population. Methods: A total of 1,831 moderate-to-late preterm infants were included in this study. Clinical data collected from the prenatal period and the first 48 hours after birth were analyzed. Multiple algorithms were employed for feature selection, and the predictive performance of 16 different models was compared to identify and optimize the best-performing model. Results: Among the included pediatric patients, a total of 32 cases of sIVH (1.7%) were diagnosed. The optimal model was the Naive Bayes model. High C-reactive protein levels, administration of catecholamine drugs within 48 hours of admission, and receiving tracheal intubation immediately after birth were independent risk factors for sIVH. Conclusions: This study successfully developed an effective machine learning prediction model. It can assist clinicians in early identification of moderate-to-late preterm infants at high risk for sIVH, thereby improving prognosis and reducing mortality.**

## Keywords

**Moderate-to-Late Preterm Infants, Severe Intraventricular Hemorrhage, Prediction Model, Machine Learning, SHAP**

Copyright © 2026 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

严重脑室内出血(Severe Intraventricular Hemorrhage, sIVH), 即 Papile 分级 III 级及以上的颅内出血[1], 是早产儿死亡及神经发育障碍的重要原因[2]。80%至 85%的早产儿出生在妊娠 32 至 36 + 6 周之间[3], 但临床关注点和医疗资源长期向极早产儿(胎龄 < 32 周)倾斜。尽管国内外研究已在极早产儿群体中建立了若干脑室内出血预测模型[4] [5], 或针对该群体探索了新型预测标志物[6] [7], 但对于中晚期早产儿的研究仍有待进一步探讨。机器学习(Machine Learning, ML)因其强大的特征识别和复杂模式挖掘能力, 在医学预测模型构建中展现出巨大潜力[8]。但其“黑箱”特性降低了模型的可解释性[9], 在一定程度上阻碍了临床医生的理解与信任。

本研究旨在通过单中心回顾性研究, 构建中晚期早产儿发生 sIVH 的预测模型。采用多种算法进行特征筛选, 构建并比较 16 种预测模型, 选取综合效能最佳者作为最终模型, 经 2 种概率校准方法校准后通过内部验证, 最终开发出一个 sIVH 风险预测工具, 为后续的研究奠定基础。

## 2. 资料与方法

### 2.1. 一般资料

本研究用于模型开发及验证的数据均来源于山东大学齐鲁第二医院新生儿重症监护室(NICU)的电子健康记录系统。研究对象为该院 NICU 收治的所有中晚期早产儿, 结局指标为住院期间诊断为 III、IV 级 IVH。研究时间为 2017 年 1 月至 2025 年 7 月, 共纳入 1831 例中晚期早产儿。基于头颅超声、CT 及磁共振结果进行诊断, 并结合该院丰富的临床经验制定治疗策略, 具体包括应用止血药物、降低颅内压、手术治疗等。

纳入标准: 1) 胎龄介于 32 至 36 + 6 周的中晚期早产儿; 2) 出生后 24 小时内入院。排除标准: 1)

关键临床病历资料缺失或不完整；2) 患有严重先天性畸形或遗传代谢性疾病。数据收集过程中，将变量进行统一命名和单位标准化，对分类变量进行编码，对连续变量进行归一化处理，排查异常值，处理缺失值。删除缺失率 >30%的变量，填补缺失率 <30%的缺失值(连续变量使用均值填补，类别变量使用众数填补)。

## 2.2. 结局指标

模型开发的主要结局指标是住院期间 sIVH 的发生情况。诊断标准：住院期间任意一次颅脑超声、CT 或磁共振检查提示 sIVH 者即诊断为 sIVH，以最严重的一次检查结果判定其严重程度。患儿颅脑超声、CT 或磁共振检查结果均由本院影像科医生出具，分级标准按照 Papile 分级方法：I 级：单侧或双侧室管膜下出血；II 级：室管膜下出血穿破室管膜进入脑室腔，合并脑室内出血；III 级：脑室内出血，伴脑室扩张；IV 级：脑室内出血，伴脑室周围出血性梗死。

## 2.3. 预测因子

基于系统性综述、Meta 分析[10]、既往类似研究[11]和专家临床意见[12]，本研究初步筛选出 28 个可能影响中晚期早产儿 sIVH 发生的因素，这些因子涵盖以下五个方面：① 母婴基本情况：性别、胎龄、出生体重、身长、头围、胸围、母亲肥胖情况、母亲宫颈机能不全情况、分娩方式、是否为小于胎龄儿；② 围产期状况与干预：1 分钟与 5 分钟 Apgar 评分、是否接受心肺复苏、是否接受气管插管、是否接受持续气道正压通气呼吸支持；③ 实验室与生理指标：降钙素原、C 反应蛋白、血小板数目、pH 值、二氧化碳分压、高碳酸血症；④ 药物应用：是否使用肾上腺素、儿茶酚胺类药物、布洛芬、咖啡因、酚磺乙胺、苯巴比妥；⑤ 合并症：有无呼吸窘迫综合征。所有数据均收集于患儿入院前及入院后 48 小时之内。详见表 1。

通过 RF [13]、Lasso 回归[14]、Boruta 算法[15]、XGBoost [16]共 4 种方法进行预测因子筛选，根据这些变量对结局的影响程度和具有非零系数的变量进行排序，并通过取四种方法选出的变量交集来识别共同变量。

## 2.4. 样本量

适当的样本量通常取决于指定的统计假设和少数研究设计参数[17]。本研究的 prevalence 设定为 10.0%，cstatistic 为 0.8，预测因子为 4 个，则模型开发所需的最小样本量为 305 例，最小阳性样本量为 31 例。对于机器学习而言，样本量越大，构建的预测模型越稳健。本研究的样本量符合分析要求。所有统计分析均使用 R 统计软件版本 4.3.3 完成。

## 2.5. 模型开发和比较

将数据集随机划分为训练集(70%)和测试集(30%)，在训练中使用合成少数类过采样计数(SMOTE)方法平衡正负样本比例。本研究采用 16 种机器学习模型预测中晚期早产儿严重脑室内出血的发生，包括轻量梯度提升机、CatBoost 分类器、随机森林分类器、梯度提升分类器、极限梯度提升(XGBoost)、AdaBoost 分类器、额外树分类器、逻辑回归、决策树分类器、线性判别分析、岭分类器、K 近邻分类器、朴素贝叶斯(Naive Bayes, NB)、支持向量机(线性核)、虚拟分类器、二次判别分析。通过召回率、受试者工作特征曲线下面积(AUC)及决策曲线分析(DCA)对这 16 个模型进行评估与比较。针对最优模型的内部验证，则采用敏感度、特异性、马修斯相关系数及 F1 分数等指标进行综合评估。此外，为优化模型预测概率的准确性，采用概率校准方法对最优模型进行进一步校准。

**Table 1.** Patient baseline characteristics analysis

**表 1.** 患者基线资料分析

变量名	Total (n = 1831)	Severe IVH (n = 32)	Non-Severe IVH (n = 1799)	p 值
性别[男/女, 例(%)]	988 (54%)/843 (46%)	15 (46.9%)/17 (53.1%)	973 (54.1%)/826 (45.9%)	0.527
接受心肺复苏[例(%)]	18 (1%)	6 (18.8%)	12 (0.7%)	<0.001
母亲肥胖情况[例(%)]	7 (0.4%)	0 (0%)	7 (0.4%)	1.000
呼吸窘迫综合症[例(%)]	388 (21.2%)	20 (62.5%)	368 (20.5%)	<0.001
使用苯巴比妥[例(%)]	20 (1.1%)	2 (6.2%)	18 (1%)	0.046
使用酚磺乙胺[例(%)]	10 (0.5%)	1 (3.1%)	9 (0.5%)	0.162
母亲宫颈机能不全情况[例(%)]	9 (0.5%)	1 (3.1%)	8 (0.4%)	0.147
接受气管插管[例(%)]	111 (6.1%)	11 (34.4%)	100 (5.6%)	<0.001
分娩方式[顺产/剖宫产, 例(%)]	408 (22.3%)	15 (46.9%)	393 (21.8%)	0.002
接受持续气道正压通气呼吸支持[例(%)]	22 (1.2%)	9 (28.1%)	13 (0.7%)	<0.001
小于胎龄儿[例(%)]	180 (9.8%)	5 (15.6%)	175 (9.7%)	0.236
使用肾上腺素[例(%)]	25 (1.4%)	4 (12.5%)	21 (1.2%)	<0.001
使用布洛芬[例(%)]	14 (0.8%)	2 (6.2%)	12 (0.7%)	0.024
使用儿茶酚胺类药物[例(%)]	61 (3.3%)	7 (21.9%)	54 (3%)	<0.001
高碳酸血症[例(%)]	955 (52.2%)	23 (71.9%)	932 (51.8%)	0.038
使用咖啡因[例(%)]	209 (11.4%)	9 (28.1%)	200 (11.1%)	0.007
胎龄(天)	34.8 ± 1.3	34.2 ± 1.4	34.8 ± 1.3	0.025
1 分钟 Apgar 评分	9.4 ± 1.4	7.7 ± 2.8	9.4 ± 1.3	0.001
5 分钟 Apgar 评分	9.8 ± 0.9	8.4 ± 2.5	9.8 ± 0.8	0.003
出生体重(kg)	2.3 ± 0.5	2.3 ± 0.7	2.3 ± 0.5	0.660
胸围(cm)	30.1 ± 2.5	29.7 ± 3.7	30.1 ± 2.5	0.589
头围(cm)	31.8 ± 2.0	31.3 ± 2.5	31.8 ± 2.0	0.284
身长(cm)	45.8 ± 3.1	45.3 ± 4.6	45.9 ± 3.1	0.493
降钙素原水平(ng/ml)	5.4 ± 11.1	24.6 ± 27.7	5.1 ± 10.2	<0.001
血小板数目(10 <sup>9</sup> /L)	227.6 ± 75.6	179.9 ± 87.1	228.4 ± 75.1	0.004
pH	7.3 ± 0.1	7.3 ± 0.1	7.3 ± 0.1	0.071
C 反应蛋白水平(mg/L)	4.5 ± 12.2	29.5 ± 42.4	4.1 ± 10.5	0.002
二氧化碳分压(mmHg)	45.8 ± 11.7	53.5 ± 15.1	45.7 ± 11.6	0.006

注：分类变量以 n (%)表示，连续变量以均值 ± 标准差表示。IVH：脑室内出血。

### 3. 结果

#### 3.1. 人口特征

本研究回顾性分析了 2017 年 1 月至 2025 年 7 月在我院新生儿科住院的所有中晚期早产儿，共纳入 1831 名早产儿进行分析，其中阳性组 32 例，对照组 1799 例(图 1)。这一发生率与 2021 年发表的一项基于全球数据的权威荟萃分析结果高度吻合，该研究报道的晚期早产儿 sIVH 患病率为 1.8% (95% CI

1.2%~2.8%) [7]。

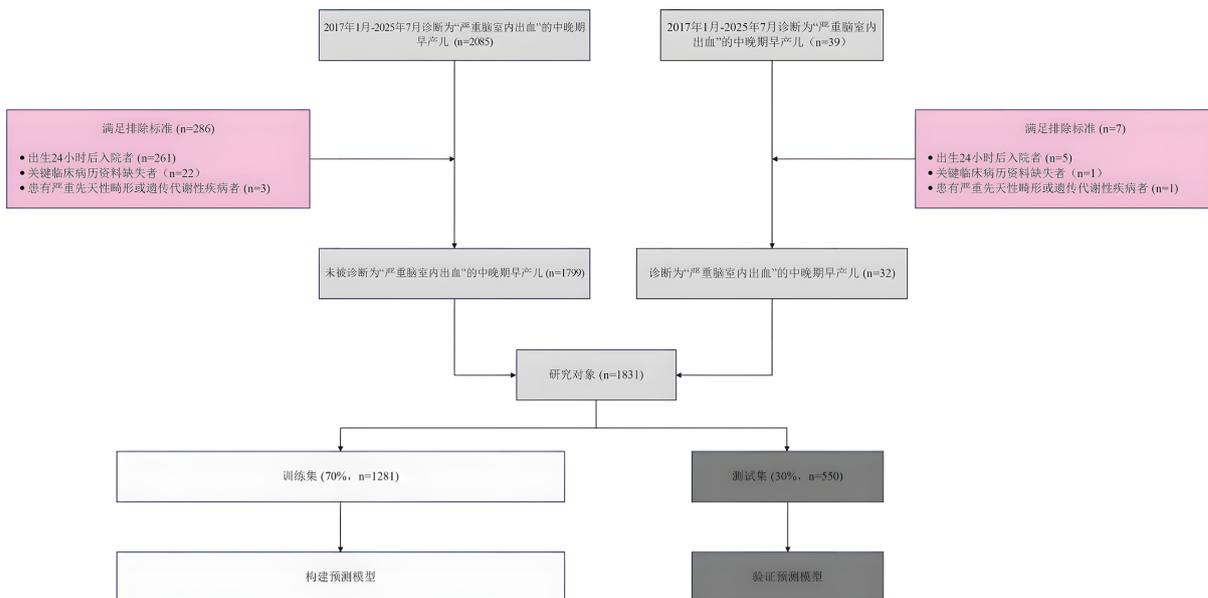


Figure 1. Flowchart of the study  
图 1. 本研究流程图

### 3.2. 变量筛选

本研究通过运用 RF 算法、Lasso 回归、Boruta 算法、XGBoost 分别筛选出 4 组重要预测变量，通过取四类方法结果的交集，共筛选出 5 个预测因子，分别为：降钙素原水平、C 反应蛋白水平、生后立即接受心肺复苏、生后立即接受持续气道正压通气呼吸支持、入院 48 小时内应用儿茶酚胺类药物。韦恩图详见图 2。尽管 C 反应蛋白与降钙素原同属炎症标志物且存在显著共线性，但降钙素原在非感染状态下可能出现生理性升高，特异性相对较低[18]。因此，为避免潜在的偏倚并基于其在新生儿炎症评估中的稳定表现，本研究最终选择反应蛋白水平作为预测因子纳入分析。同时，结合临床经验，以 Apgar 5 分钟评分替代出生后需进行心肺复苏这一变量，因二者均可反映患儿存在围生期不良事件或需要医疗支持的风险。但 Apgar 评分作为一种连续变量，能提供更精细、更标准化的生理状态信息，从而更有效地量化此风险谱系。虽然在初步筛选中发现持续气道正压通气(CPAP)呼吸支持与不良结局相关，但考虑到 CPAP 在临床实践中多用于早期呼吸支持，常具有保护性作用，而与气管插管相比，其侵入性和风险程度明显较低。因此，经综合评估后，本研究最终将生后立即接受气管插管确定为独立危险因素纳入分析。最终模型纳入的 4 个预测因子：C 反应蛋白水平、生后立即接受气管插管呼吸支持、入院后 48 小时内应用儿茶酚胺类药物及 Apgar 5 分钟评分——分别从感染炎症、呼吸功能、循环稳定性及围产期过渡状态四个核心病理生理维度，共同构建了一个全面且互不重叠的风险预测框架。

### 3.3. 模型开发和性能比较

在 AUC 指标上，Naive Bayes (NB)模型以 0.8446 的表现位居榜首，当阈值为 0.1 时，其召回率为 0.8，为模型中最优水平，故将 Naive Bayes 模型选为最佳模型。各个模型的受试者工作曲线汇总见图 3。为提高模型输出概率的准确性与可信度，采用 Platt 缩放法(Platt Scaling) [19]与等张回归法(Isotonic Regression) [20]对其进行系统的概率校准。基于在处理大样本数据时的理论适应性，以及在本研究数据集上所呈现出的更优越校准性能，选择了等张回归进行模型校准。决策曲线分析(Decision Curve Analysis, DCA) [21]结果显示

(校准前与 2 种方法校准后的 DCA 比较见图 4), 经校准后, 本预测模型在 0 至 0.180 的阈值概率区间内, 其净获益(Net Benefit)持续高于“对所有患者进行干预”(Treat All)和“不对任何患者进行干预”(Treat None)两种策略。这表明, 在该区间内, 临床医生使用本模型指导决策能够获得更优的临床净收益, 即在有效识别目标患者的同时, 显著减少不必要的干预。此外, 在验证集中, 阈值为 0.10 时, NB 模型的 AUC 为 0.9106, 召回率为 0.7000。这些数据表明, 在验证集中, 该模型对重度脑室内出血的识别能力表现出色。

### 3.4. 模型的解释

本研究采用 SHAP 方法对模型输出进行解释, 通过量化各变量对预测结果的贡献度, 揭示模型的决策机制。图 5 为显示每个特征的 SHAP 值分布的蜂群图, 提供了每个特征如何在不同值范围内影响模型预测结果的视觉表示。SHAP 值用颜色编码(蓝色表示较低的特征值, 红色表示较高的特征值), 以揭示特

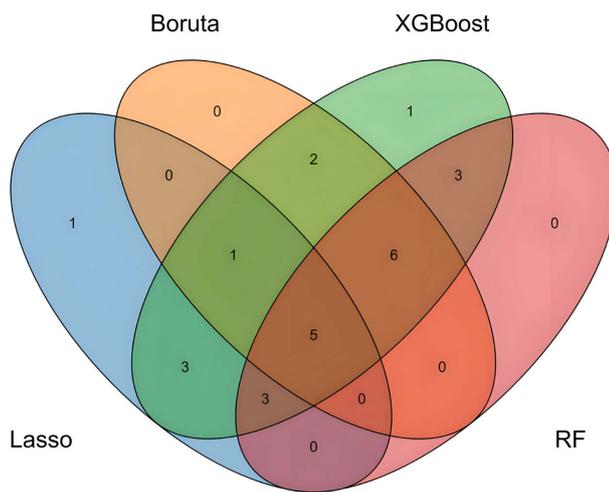


Figure 2. Venn diagram of predictors selected by 4 methods  
图 2. 4 种方法筛选预测因子的韦恩图

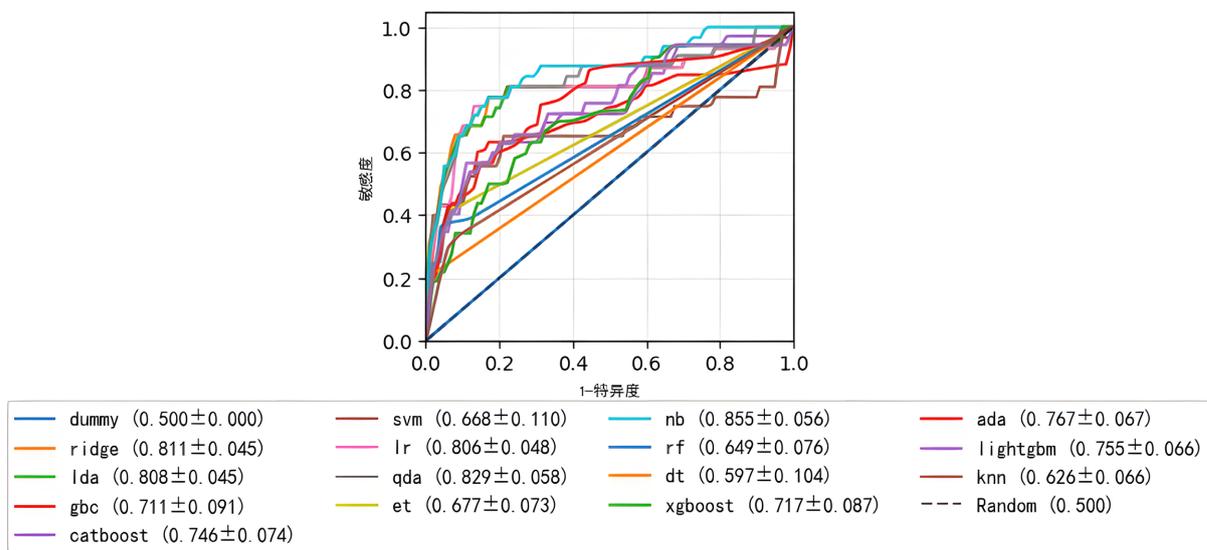


Figure 3. ROC curves of the 16 models  
图 3. 16 种模型的受试者工作曲线图

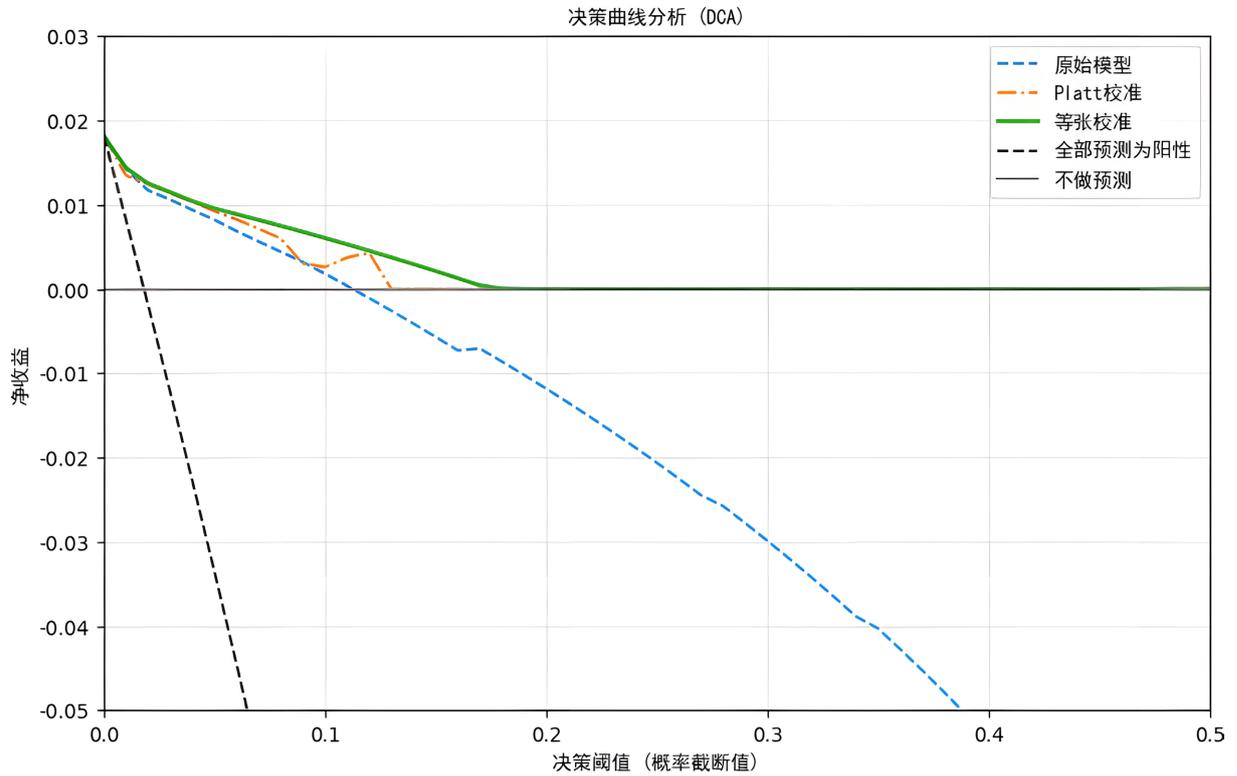


Figure 4. Decision curve analysis of the original model before and after calibration

图 4. 原始模型与校准前后的决策曲线分析图

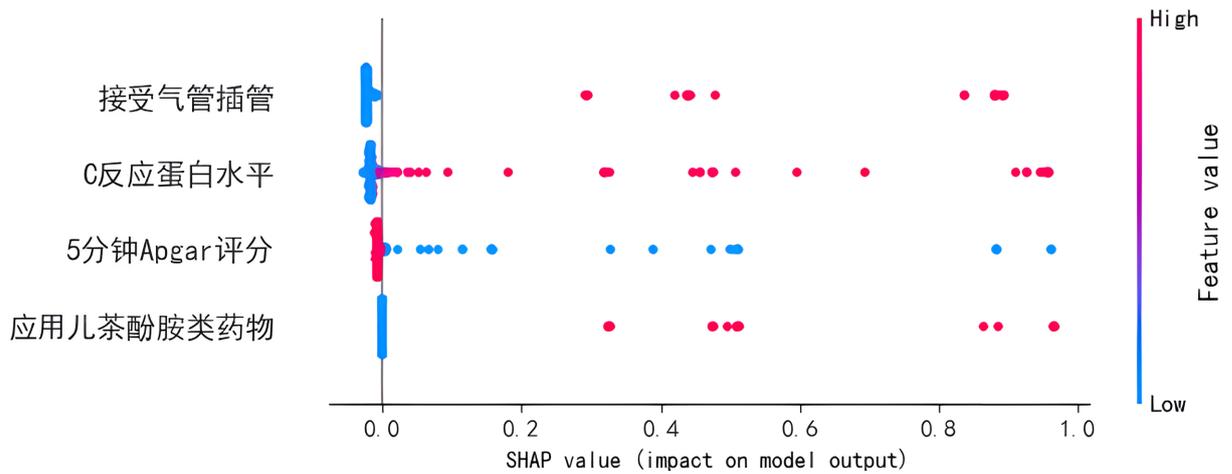


Figure 5. SHAP beeswarm plot

图 5. SHAP 蜂群图

征值和预测结果之间的关系。在所有特征中，气管插管是影响最显著的特征，当患者接受气管插管时，SHAP 值多为正值，提示其对疾病发生风险具有正向驱动作用；高 C 反应蛋白水平与应用儿茶酚胺类药物也表现出明显的正向贡献，且特征值越高，对模型输出的正向影响越强；而 5 分钟 Apgar 评分则呈现负向作用，即较高的 5 分钟 Apgar 评分会降低疾病发生的预测风险。这些结果表明，产后立即接受气管插管、C 反应蛋白升高及入院后 48 小时内应用儿茶酚胺类药物这类临床指标是疾病发生的高危因素，会增加疾病发生的概率；高水平的 5 分钟 Apgar 评分则是疾病发生的保护因素。

## 4. 讨论

本研究基于山东大学齐鲁第二医院 NICU 的电子病历数据开发了机器学习模型, 该模型在预测中晚期早产儿发生严重脑室内出血(sIVH)方面展现出强大的判别能力和临床应用价值。基于电子病历数据, 通过运用随机森林、Lasso 回归、Boruta 和 XGBoost 算法, 结合临床经验, 从数据库中筛选出 4 个关键变量, 开发并对外验证了针对中晚期早产儿发生严重脑室内出血的预测模型。本研究旨在开发适用于临床的实用预测模型, 因此预测变量均为患儿入院后 48 小时内采集的数据。根据文献综述和荟萃分析结果, 这 4 个变量在同类研究中具有普遍适用性。本研究数据覆盖 8 年患者信息, 采用回顾性分析方式, 且所选变量是电子病历中最全面、最易获取的记录。为确定最佳拟合效果, 分别对 4 个预测变量应用了 16 种机器学习算法进行单独拟合。最终 NB 模型表现最为优异, 其召回率高达 0.7、临床净效益达 1.82%、曲线下面积(AUC)为 0.8446。该模型在内部验证集上同样展现出出色性能。

通过运用随机森林、套索回归、Boruta 和 XGBoost 共 4 种算法, 结合临床医生的建议, 筛选出 4 个预测因子, 确定高 C 反应蛋白水平、入院后 48 小时内应用儿茶酚胺类药物、生后立即接受气管插管是发生 sIVH 的危险因素。感染与炎症在 sIVH 发生中的作用不容忽视。C 反应蛋白作为炎症标志物, 能够反映感染的严重程度和全身炎症反应的状态。感染激发的免疫反应和炎症瀑布反应导致炎症因子增多, 脑耗氧量增加, 引起脑血管调节功能紊乱, 最终增加血管破裂风险[22]。本研究证实, 高 5 分钟 Apgar 评分是 sIVH 的保护因素, 这与既往研究结果一致[23]。5 分钟 Apgar 评分反映了婴儿出生后的适应能力和复苏效果。较高的 5 分钟 Apgar 评分(通常 $\geq 7$ 分)表明婴儿成功完成了从宫内到宫外生活的过渡, 没有经历严重的围生期窒息。血流动力学稳定性在 sIVH 预防中至关重要。儿茶酚胺类药物可能导致全身血压的快速变化, 进而影响脑血流[24]。由于早产儿大脑血管结构脆弱, 脑血管的自主调节功能不成熟, 这种脑血流的剧烈波动很容易导致生发基质内脆弱的毛细血管破裂, 引发脑室内出血[25]。接受气管插管通常标志着患儿存在严重的呼吸系统疾病或围产期窒息等基础高危状态[26]。插管操作及后续的有创机械通气本身也可能直接增加出血风险: 早产儿尤其伴有缺氧酸中毒时, 常存在压力被动性脑血流, 使其生发基质内脆弱的毛细血管无法耐受喉镜刺激、气道吸引等操作引发的急剧血流动力学波动[27]。

王露冉等人开展了一项针对极早产儿不良预后的临床研究[4], 重点分析了出生胎龄  $< 32$  周的早产儿在生后 1 周内发生死亡或严重脑室内出血的围生期危险因素。该研究回顾性收集了 256 例极早产儿的临床资料, 通过多因素 logistic 回归分析确定了 5 个独立危险因素(包括产前激素使用、出生体重、脐动脉血 pH 值等)。研究结果显示, 基于这些因素构建的预测模型具有较好的区分度, AUC 值达到 0.85。然而, 该研究仅采用了传统的回归分析方法, 未尝试更复杂的机器学习算法, 可能限制了模型预测性能的进一步提升。2023 年刘希等学者进行的多中心研究[5], 旨在构建极/超早产儿中重度脑室周围 - 脑室内出血的预测模型。该研究纳入了 538 例患儿数据, 筛选出 8 个预测因子(包括胎龄、出生体重、产前激素使用等), 并建立了风险预测列线图。验证结果显示该模型的 AUC 值为 0.83, 校准曲线具有良好的拟合度。值得注意的是, 该研究虽然验证了模型的性能, 但未将使用儿茶酚胺类药物等重要的血流动力学指标纳入分析, 可能影响了模型在临床实践中的预警价值。在研究对象选择方面, 现有关于早产儿 sIVH 的预测模型研究, 绝大多数集中于极早产儿或超低出生体重儿。然而, 数量更为庞大的中晚期早产儿(胎龄  $\geq 32$  周)的生理特点与医疗问题独具特征, 却长期被视为“低风险”群体而缺乏专门的风险评估工具。本研究精准聚焦于这一被相对忽视的群体, 致力于解决其特有的临床需求, 填补了该领域预测模型研究的空白。在筛选预测因子方面, 本研究使用 4 种算法取交集筛选特征, 确保了特征的强鲁棒性, 减少了过拟合风险。与已发表的现有模型相比, 通过严格的变量筛选、模型构建及验证, 这一创新模型实现了更优的预测准确性。在模型开发过程中, 通过交叉验证、内部验证等方法, 预先设置随机种子以最大限度提高模

型的准确性与可信度。构建模型后,采用校准策略,校准后的模型在决策曲线分析(DCA)中展现出更高的临床净收益,这意味着其在辅助临床决策(如针对高危患儿加强监护)时具有更明确的实用价值。

机器学习在临床医学领域显示出巨大前景。将机器学习融入临床环境带来了一系列优势,比如提高诊断精度、简化管理任务和个性化治疗计划等。机器学习技术的集成促进了明智的临床决策过程,有望实现更快速精准的诊断、个体化的治疗方案以及医疗成本的降低[28]。机器学习在分析医疗数据、处理异质性多维数据方面表现突出,本研究中纳入的4个预测变量(如C反应蛋白水平、5分钟Apgar评分等指标)就充分体现了这种特性,这些数据本身具有高度异质性。此外,本研究采用SHAP方法来阐释预测模型并开发实用的网络应用。这一步骤对于消除黑盒模型的担忧、重建医学领域对机器学习的信心至关重要[29]。

本研究主要存在以下局限。第一,本研究为单中心回顾性设计,数据来源于单一机构,其患者特征与临床实践具有特定背景,可能导致选择偏倚,并影响模型的外推性。尽管内部验证效果良好,但模型在更广泛人群与其他医疗机构中的效能仍需通过多中心前瞻性研究加以验证。第二,模型中部分连续变量(如C反应蛋白水平、5分钟Apgar评分)所使用的最佳截断值是基于当前数据优化得出,其临床适用性仍需在前瞻性队列中进一步确认。为应对上述局限,我们正在整合来自多家医院的近期数据,以期在后续研究中优化模型并增强其泛化能力。

总之,本研究确定了4个关键预测因子(C反应蛋白水平、5分钟Apgar评分、入院后48小时内使用儿茶酚胺类药物、生后立即接受气管插管),并开发了一个AUC高达0.8446的预测模型。该模型为中晚期早产儿sIVH的早期预测提供了实用的预测工具,从而使医生能够主动做出有针对性的临床决策。

## 声明

本研究严格遵循《赫尔辛基宣言》的宗旨,并获得山东大学齐鲁第二医院伦理委员会批准,伦理批准号: KYLL202478。研究者发起的临床研究-观察性研究的备案编号: MR-37-24-044775。

## 利益冲突与作者贡献声明

本研究的资助者未参与研究设计、数据收集、分析和解释,也未参与论文的撰写。所有作者都有权查阅原始数据,通讯作者负责决定是否发表。

## 参考文献

- [1] Papile, L., Burstein, J., Burstein, R. and Koffler, H. (1978) Incidence and Evolution of Subependymal and Intraventricular Hemorrhage: A Study of Infants with Birth Weights Less than 1,500 gm. *The Journal of Pediatrics*, **92**, 529-534. [https://doi.org/10.1016/s0022-3476\(78\)80282-0](https://doi.org/10.1016/s0022-3476(78)80282-0)
- [2] Rees, P., Callan, C., Chadda, K.R., Vaal, M., Diviney, J., Sabti, S., et al. (2022) Preterm Brain Injury and Neurodevelopmental Outcomes: A Meta-Analysis. *Pediatrics*, **150**, e2022057442. <https://doi.org/10.1542/peds.2022-057442>
- [3] Martínez-Nadal, S., Ginovart Galiana, G., Morales Luengo, F., Rodríguez Revuelta, M.J., García Reymundo, M., Ansó Oliván, S., et al. (2025) Recommendations for the Perinatal Management and Follow-Up of Moderate and Late Preterm Infants. *Anales de Pediatría (English Edition)*, **102**, Article ID: 503714. <https://doi.org/10.1016/j.anpede.2025.503714>
- [4] 王露冉, 刘学丽, 杨翠红, 等. 出生胎龄 < 32 周早产儿生后 1 周内死亡或严重脑室内出血的围生期因素分析[J]. 中华儿科杂志, 2025, 63(4): 387-393.
- [5] 刘希, 乔丹, 贝斐. 极/超早产儿中重度脑室周围-脑室内出血预测模型的研究[J]. 中华新生儿科杂志(中英文), 2023, 38(12): 715-720.
- [6] Zhang, C., Zhu, Z., Wang, K., Wang, L., Lu, J., Lu, L., et al. (2024) Predicting Neurodevelopmental Outcomes in Extremely Preterm Neonates with Low-Grade Germinal Matrix-Intraventricular Hemorrhage Using Synthetic MRI. *Frontiers in Neuroscience*, **18**, Article 1386340. <https://doi.org/10.3389/fnins.2024.1386340>
- [7] Jiang, L., Yu, Q., Wang, F., Wu, M., Liu, F., Fu, M., et al. (2023) The Role of Blood Pressure Variability Indicators

- Combined with Cerebral Blood Flow Parameters in Predicting Intraventricular Hemorrhage in Very Low Birth Weight Preterm Infants. *Frontiers in Pediatrics*, **11**, Article 1241809. <https://doi.org/10.3389/fped.2023.1241809>
- [8] Kumar, P., Chauhan, S. and Awasthi, L.K. (2023) Artificial Intelligence in Healthcare: Review, Ethics, Trust Challenges & Future Research Directions. *Engineering Applications of Artificial Intelligence*, **120**, Article ID: 105894. <https://doi.org/10.1016/j.engappai.2023.105894>
- [9] Toumaj, S., Heidari, A. and Jafari Navimipour, N. (2025) Leveraging Explainable Artificial Intelligence for Transparent and Trustworthy Cancer Detection Systems. *Artificial Intelligence in Medicine*, **169**, Article ID: 103243. <https://doi.org/10.1016/j.artmed.2025.103243>
- [10] Ghaderi, M., Afraie, M., Pourahmad, B., Amirmanesh, N., Rahimi, A., Sheikahmadi, S., *et al.* (2025) Comprehensive Evaluation of Risk Factors for Intraventricular Hemorrhage in Preterm Neonates: A Systematic Review and Meta-Analysis. *European Journal of Medical Research*, **30**, Article No. 695. <https://doi.org/10.1186/s40001-025-02960-2>
- [11] Shen, F., Xu, J., Rong, H., Zhang, J., Yang, Y. and Li, X. (2025) Construction and Validation of a Risk Prediction Model for Early Severe Intraventricular Hemorrhage in Very Low Birth Weight Infants. *The Kaohsiung Journal of Medical Sciences*, **41**, e70037. <https://doi.org/10.1002/kjm2.70037>
- [12] 中国医师协会新生儿科医师分会, 北京医师协会新生儿科医师分会, 李秋平, 等. 早产儿脑室内出血预防专家共识(2025) [J]. 中华妇幼临床医学杂志(电子版), 2025, 21(1): 1.
- [13] Hastie, T., Tibshirani, R. and Friedman, J. (2009) *The Elements of Statistical Learning*. Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- [14] Sun R.T., Li, C.L., Jiang, Y.M., Hao, A.Y., Liu, K., Li, K., *et al.* (2025) A Radiomics-Clinical Predictive Model for Difficult Laparoscopic Cholecystectomy Based on Preoperative CT Imaging: A Retrospective Single Center Study. *World Journal of Emergency Surgery*, **20**, Article No. 62. <https://doi.org/10.1186/s13017-025-00635-1>
- [15] Lin, J., Chen, Y., Xu, M., Chen, J., Huang, Y., Chen, X., *et al.* (2024) Association and Predictive Ability between Significant Perioperative Cardiovascular Adverse Events and Stress Glucose Rise in Patients Undergoing Non-Cardiac Surgery. *Cardiovascular Diabetology*, **23**, Article No. 445. <https://doi.org/10.1186/s12933-024-02542-2>
- [16] Moore, A. and Bell, M. (2022) XGBoost, a Novel Explainable AI Technique, in the Prediction of Myocardial Infarction: A UK Biobank Cohort Study. *Clinical Medicine Insights: Cardiology*, **16**, 11795468221133611. <https://doi.org/10.1177/11795468221133611>
- [17] Wang, X. and Ji, X. (2020) Sample Size Estimation in Clinical Research: From Randomized Controlled Trials to Observational Studies. *Chest*, **158**, S12-S20. <https://doi.org/10.1016/j.chest.2020.03.010>
- [18] Ruan, L., Chen, G.Y., Liu, Z., Zhao, Y., Xu, G.Y., Li, S.-F., *et al.* (2018) The Combination of Procalcitonin and C-Reactive Protein or Presepsin Alone improves the Accuracy of Diagnosis of Neonatal Sepsis: A Meta-Analysis and Systematic Review. *Critical Care*, **22**, Article No. 316. <https://doi.org/10.1186/s13054-018-2236-1>
- [19] Odesola, P.A., Adegoke, A.A. and Babalola, I. (2025) Model Uncertainty Quantification: A Post Hoc Calibration Approach for Heart Disease Prediction. *medRxiv*, 2025.09.28.25336834. <https://doi.org/10.1101/2025.09.28.25336834>
- [20] 范双龙, 赵志强, 余红梅, 等. 基于概率校准的弥漫性大 B 细胞淋巴瘤患者死亡风险预测[J]. 中国卫生统计, 2021, 38(5): 670-674.
- [21] Vickers, A.J. and Elkin, E.B. (2006) Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. *Medical Decision Making*, **26**, 565-574. <https://doi.org/10.1177/0272989X06295361>
- [22] Andersen, M., Matthiesen, N.B., Murra, M., Nielsen, S.Y. and Henriksen, T.B. (2025) Early-Onset Neonatal Infection and Epilepsy in Children. *JAMA Network Open*, **8**, e2519090. <https://doi.org/10.1001/jamanetworkopen.2025.19090>
- [23] 林冰纯, 陈春, 黄智峰, 等. 超早产儿脑室周围-脑室内出血发生率及其危险因素分析-中华新生儿科杂志(中英文) [J]. 中华新生儿科杂志, 2022, 37(1): 12-16.
- [24] Abdul Aziz, A.N., Thomas, S., Murthy, P., Rabi, Y., Soraisham, A., Stritzke, A., *et al.* (2019) Early Inotropes Use Is Associated with Higher Risk of Death and/or Severe Brain Injury in Extremely Premature Infants. *The Journal of Maternal-Fetal & Neonatal Medicine*, **33**, 2751-2758. <https://doi.org/10.1080/14767058.2018.1560408>
- [25] Noori, S. and Seri, I. (2015) Hemodynamic Antecedents of Peri/Intraventricular Hemorrhage in Very Preterm Neonates. *Seminars in Fetal and Neonatal Medicine*, **20**, 232-237. <https://doi.org/10.1016/j.siny.2015.02.004>
- [26] Xiao, T., Hu, L., Chen, H., Gu, X., Zhou, J., Zhu, Y., *et al.* (2024) The Performance of the Practices Associated with the Occurrence of Severe Intraventricular Hemorrhage in the Very Premature Infants: Data Analysis from the Chinese Neonatal Network. *BMC Pediatrics*, **24**, Article No. 394. <https://doi.org/10.1186/s12887-024-04664-8>
- [27] 李俊霞, 徐煜皓, 陆超, 等. 超早产儿脑室内出血的风险因素分析[J]. 南京医科大学学报(自然科学版), 2023, 43(7): 990-994.

- 
- [28] Karalis, V.D. (2024) The Integration of Artificial Intelligence into Clinical Practice. *Applied Biosciences*, **3**, 14-44.  
<https://doi.org/10.3390/applbiosci3010002>
- [29] Narkhede, J. (2024) Comparative Evaluation of Post-Hoc Explainability Methods in AI: LIME, SHAP, and Grad-CAM. 2024 4th International Conference on Sustainable Expert Systems (ICSES), Kaski, 15-17 October 2024, 826-830.  
<https://doi.org/10.1109/icses63445.2024.10762963>