

基于机器学习的老年全身麻醉颈椎 术后谵妄发生风险模型的构建与验证

徐娟¹, 刘婷婷¹, 高俊², 蒋云芬¹

(1.中国人民解放军联勤保障部队第九〇四医院 骨科, 江苏 无锡 214000; 2.常州市中医医院 骨科, 江苏 常州 314000)

【摘要】目的 构建并筛选出评估效能最佳的老年全身麻醉颈椎术后谵妄发生风险模型, 为临床干预提供依据。**方法** 2021 年 5 月至 2023 年 10 月, 采用便利抽样法选取某院老年全身麻醉颈椎术后患者 323 例为训练集, 将同期另一所医院的同类患者 233 例为验证集。根据术后 3 d 是否发生谵妄, 将其分为谵妄组和非谵妄组。采用单因素及二元 Logistic 回归分析确定术后危险因素, 采用列线图、决策树 CART、随机森林构建风险模型, 通过受试者工作曲线(receiver operator characteristic, ROC) 分析不同模型的评估效能。**结果** 单因素分析及二元 Logistic 回归分析确定, 术中输血、术后脑脊液渗漏、糖尿病、手术时长等均为老年全身麻醉颈椎术后谵妄发生的危险因素(均 $P < 0.05$), 据此建立列线图模型。决策树 CART 模型分析显示年龄为根节点, 手术时长、糖尿病、术中输血等为子节点。随机森林模型筛选出年龄、手术时长、术中输血、糖尿病等 7 个重要变量。在训练集和验证集, 列线图模型、决策树 CART 模型、随机森林模型评估的曲线下面积分别为 0.888 和 0.895、0.874 和 0.884、0.911 和 0.911, 随机森林模型与决策树 CART 模型的差异均有统计学意义(均 $P < 0.05$)。**结论** 三种模型对老年全身麻醉颈椎术后谵妄均有评估价值, 随机森林模型效能最优。

【关键词】 全身麻醉颈椎术后; 谵妄; 列线图; 决策树 CART; 随机森林

doi: 10.3969/j.issn.2097-1826.2025.02.015

【中图分类号】 R473; R823 **【文献标识码】** A **【文章编号】** 2097-1826(2025)02-0058-05

Construction and Validation of a Risk Model for Postoperative Delirium in Elderly Patients Undergoing Cervical Spine Surgery Under General Anesthesia Based on Machine Learning

XU Juan¹, LIU Tingting¹, GAO Jun², JIANG Yunfen¹ (1. Department of Orthopedics, The 904th Hospital of Joint Logistic Support Force of PLA, Wuxi 214000, Jiangsu Province, China; 2. Department of Orthopedics, Changzhou Traditional Chinese Medicine Hospital, Changzhou, 213004, Jiangsu Province, China)

Corresponding author: JIANG Yunfen, Tel: 0510-85142341

【Abstract】Objective To construct and screen a risk model of postoperative delirium in elderly patients undergoing general anesthesia for cervical spine surgery with the best evaluation efficiency, so as to provide a basis for clinical intervention. **Methods** From May 2021 to October 2023, the convenience sampling method was used to select 323 elderly patients undergoing cervical general anesthesia in a hospital as the training set, and 233 similar patients in another hospital during the same period were selected as the validation set. According to whether delirium occurred 3 days after surgery, the patients were divided into delirium group and non-delirium group. Nomogram, decision tree CART and random forest were used to construct risk models. Receiver operator characteristic(ROC) curve was used to analyze the evaluation efficacy of different models. **Results** Univariate analysis and binary Logistic regression analysis showed that intraoperative blood transfusion, postoperative cerebrospinal fluid leakage, diabetes, and operation duration were risk factors for delirium after cervical general anesthesia in the elderly (all $P < 0.05$), and a nomogram model was established accordingly. Decision tree CART model analysis showed that age was the root node, and operation duration, diabetes, intraoperative blood transfusion were the child nodes. Seven important variables including age, operation duration, intraoperative blood transfusion and diabetes were selected by random forest model. In the training set and validation set, the area under the curve of the nomogram model, the decision tree CART model, and the random forest model were 0.888 and 0.895, 0.874 and 0.884, 0.911 and 0.911, respectively, and the differences between the random forest model and the decision tree CART model were statistically significant (all $P < 0.05$). **Conclusions** All the three models have assessment value for postoperative delirium after general anesthesia for cervical spondylosis in the elderly, and the random forest model has the best performance.

【Key words】 cervical general anesthesia; delirium; nomogram; decision tree CART; random forest

【收稿日期】 2024-06-11 **【修回日期】** 2024-12-14

[Mil Nurs, 2025, 42(02): 58-61, 66]

【基金项目】 江苏省重点研发计划专项课题(BE2019653); 第九〇四医院院管科研课题(MS202111); 江苏省高层次卫生人才“六个一工程”拔尖人才项目(LGY2019026)

【作者简介】 徐娟, 本科, 副主任护师, 电话: 0510-85142631

【通信作者】 蒋云芬, 电话: 0510-85142341

谵妄是急性发作的精神病症, 患者在意识清楚的情况下产生不真实或荒谬的观念、幻觉和妄想, 并且很难与现实进行区分, 可由感染、药物、代谢紊乱、

睡眠障碍等引起^[1]。老年全身麻醉手术患者谵妄发生率较高,是造成患者跌倒、功能障碍、住院费用增加甚至死亡的风险因素,但 30%~40%的谵妄可以预防的^[2],因此,精准识别谵妄患者,积极纠正可逆因素有重要意义。术后谵妄影响因素众多,包括患者自身易感因素、手术因素、疾病因素以及炎症和应激因素等,且部分研究根据以上影响因素构建了各种评估模型^[3-4]。各模型存在差异,进而影响模型评估效能,因此选择最佳模型对精准判断谵妄发生,并提示临床采取干预措施有积极意义。机器学习已被广泛应用于疾病诊断、并发症、预后等评估模型中,使得构建更为精准的谵妄发生评估模型成为可能。本研究分别基于列线图、决策树 CART、随机森林构建了老年全身麻醉颈椎术后谵妄发生风险模型,为临床筛选最佳模型,实施临床干预提供参考依据。

1 对象与方法

1.1 研究对象 2021 年 5 月至 2023 年 10 月,采用便利抽样法选取某院老年全身麻醉颈椎术后患者 323 例为训练集,其中谵妄 51 例、非谵妄 272 例;同期另一所医院的同类患者 233 例为验证集,其中谵妄 42 例,非谵妄 191 例。纳入标准:因颈椎创伤或颈椎疾病行全身麻醉手术的老年患者。排除标准:术前精神障碍、精神疾病、脑卒中、年龄<60 岁、颅脑损伤以及临床资料不全的患者。根据术后 3 d 是否发生谵妄将所有研究对象分为谵妄组和非谵妄组。本研究经医院伦理委员会批准(20240301005),所有研究对象签署知情同意书。样本量计算公式 $n = Z_{\alpha/2}^2 \times P(1 - P) / \delta^2$,按参考文献^[5], $P = 12.30\%$, $\alpha = 0.05$, $Z_{\alpha/2} = 1.96$, $\delta = 0.05$,因此最低样本量为 166,考虑到 10%左右的样本丢失,最低样本量为 183 例。

1.2 方法

1.2.1 谵妄的评估 经统一培训,由熟练掌握意识模糊评估法(confusion assessment method,CAM)^[6]的两名护理人员于术后 1~3 d 固定时间(8:00—12:00)对研究对象开展谵妄诊断。护理人员使用 CAM 判断谵妄的 Kappa 值为 0.810~0.875。

1.2.2 调查工具 搜索知网、万方、Pubmed、Web of Science 中关于颈椎术后谵妄的影响因素的文献,结合麻醉特点,在考虑实用性和可操作性后,经讨论纳入 18 项资料,包括年龄、性别、体质量指数(body mass index,BMI)、疾病类型、基础疾病等基本资料,血红蛋白(hemoglobin,Hb)、肾功能、低蛋白血症(白蛋白<25 g/L 或总蛋白<60 g/L)等术前资料,手术时长、术中输血、术中 PaO₂<83 mmHg(1 mmHg=0.133 kPa)等术中资料,术后镇静药物使用、脑脊液漏等术后资料。

1.2.3 资料收集 由 2 名护理人员通过电子病历系统中收集患者基本资料和术前资料,通过护理记录收集其术后资料;由手术医生和麻醉医生通过手术记录收集其术中资料;对于缺失变量则使用该变量的均值或相似样本均值来补充;资料填写完毕交叉核对。

1.3 统计学处理 采用 SPSS 22.0 统计软件,计数资料以例数和百分比表示,采用 χ^2 检验;正态分布的计量资料以 $\bar{x} \pm s$ 表示,采用独立样本 t 检验;非正态分布的计量资料采用 $M(P_{25}, P_{75})$ 表示,采用 Mann-Whitney U 检验。采用 Logistic 回归逐步筛选影响因素,先后通过列线图、决策树 CART、随机森林建立谵妄发生风险评估模型。采用受试者工作曲线(receiver operating characteristic curve,ROC)下面积(area under curve,AUC)在训练集和验证集对模型评估效能进行分析,以 $P < 0.05$ 或 $P < 0.01$ 为差异有统计学意义。

2 结果

2.1 老年颈椎全麻术后谵妄发生单因素分析 训练集中,两组患者在年龄、糖尿病、手术时长、术中输血、术中 PaO₂ 过低、术后脑脊液渗漏等项目上的差异均有统计学意义(均 $P < 0.05$),本文仅列出差异有统计学意义的项目,见表 1。

表 1 老年颈椎全麻术后谵妄发生单因素分析

变 量	谵妄组 (<i>n</i> =51)	非谵妄组 (<i>n</i> =272)	<i>U</i> 或 <i>t</i> 或 χ^2	<i>P</i>
年龄(岁, $\bar{x} \pm s$)	72.32±6.41	68.41±6.98	3.717	<0.001
糖尿病[<i>n</i> (%)]	28(54.90)	92(33.82)	8.172	0.004
手术时长[$\bar{x}$ /min, <i>MP</i> ₂₅ , <i>P</i> ₇₅]	140(80,180)	120(70,170)	3.741	<0.001
术中输血[<i>n</i> (%)]	14(27.45)	32(11.76)	8.653	0.003
术中 PaO ₂ 过低[<i>n</i> (%)]	10(19.61)	22(8.09)	6.385	0.012
术后脑脊液渗漏[<i>n</i> (%)]	9(17.65)	18(6.62)	5.456	0.019

2.2 老年颈椎全麻术后谵妄发生的多因素分析 以是否发生术后谵妄为因变量,单因素分析中有统计学差异的项目为自变量,纳入 Logistic 回归分析。结果显示,术中输血、术后脑脊液漏、糖尿病、手术时长、术中 PaO₂ 过低、年龄均为老年颈椎全麻术后谵妄发生的危险因素(均 $P < 0.05$),见表 2。

表 2 老年颈椎全麻术后谵妄发生的多因素分析

变 量	β	SE	Wald χ^2	<i>P</i>	OR	95%CI	
						下限	上限
常数	-17.098	2.939	33.840	—	—	—	—
术中输血	1.409	0.584	5.820	0.016	4.091	1.302	12.850
脑脊液漏	1.182	0.594	3.958	0.047	3.261	1.018	10.447
糖尿病	1.382	0.683	4.097	0.043	3.981	1.045	15.173
手术时长	0.030	0.008	13.685	0.000	1.030	1.014	1.047
术中 PaO ₂ 过低	1.341	0.604	4.925	0.026	3.825	1.170	12.506
年龄	0.180	0.036	24.768	0.000	1.198	1.116	1.286

2.3 老年颈椎全麻术后谵妄发生的列线图模型 列线图中线条长度代表对因变量评估能力大小,所

有分数累积后总得分对应的风险即为谵妄发生概率,见图 1。经校正曲线显示该模型校准度较好。

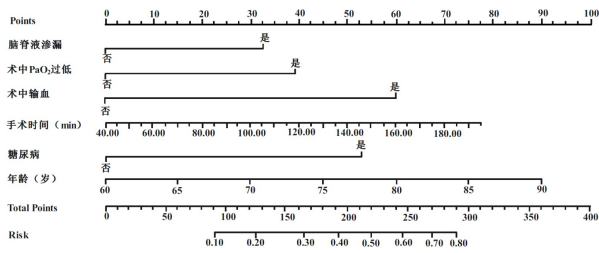


图 1 老年颈椎全麻术后谵妄发生的列线图模型

2.4 老年颈椎全麻术后谵妄决策树 CART 模型 以基尼系数进行树的生长,以最小代价复杂度减枝法进行减枝,结果显示决策树分为 3 层。与谵妄发生相关的影响因素依次为年龄、手术时长、糖尿病、术中输血、手术方式、术中 PaO₂ 过低,见图 2。

2.5 老年颈椎全麻术后谵妄随机森林模型 采用重复 5 次十折交叉验证对自变量进行取舍,当变量为 7 时口袋外数据最小,因此本研究选取前 7 项变量对模型进行优化,重要性评分居前 7 位的分别为年龄、手术时长、术中输血、糖尿病、术中 PaO₂ 过低、术后脑脊液渗漏、手术方式。

2.6 三种模型评估价值的比较 在训练集中,列线图、决策树 CART、随机森林模型评估价值的 AUC 分别为 0.888、0.874、0.911。利用 Medcalc 软件,Z 检验对不同模型 AUC 进行比较发现:随机森林模型 AUC 高于决策树 CART 模型,差异有统计学意义 ($Z=3.721, P=0.010$),其余模型间差异无统计学意义 (均 $P>0.05$)。在验证集中,上述 3 种模型评估 AUC 分别为 0.895、0.884、0.911,随机森林模型 AUC 高于决策树 CART 模型,差异有统计学意义 ($Z=3.266, P=0.031$),其余模型间差异无统计学意义 (均 $P>0.05$),具体见表 3。

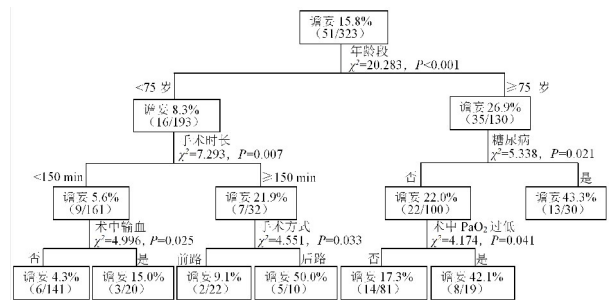


图 2 老年颈椎全麻术后谵妄决策树 CART 模型

表 3 训练集不同模型评估效能

模 型	训练集					验证集				
	AUC (95%CI)	敏感度 (%)	特异性 (%)	Youden 指数	正确率 (%)	AUC (95%CI)	敏感度 (%)	特异性 (%)	Youden 指数	正确率 (%)
列线图	0.888(0.826~0.950)	88.24	76.84	0.651	78.64	0.895(0.849~0.952)	92.86	70.68	0.635	74.68
决策树 CART	0.874(0.808~0.941)	88.24	76.10	0.643	78.02	0.884(0.831~0.937)	78.57	83.25	0.618	82.41
随机森林	0.911(0.852~0.969)	86.27	86.40	0.727	86.38	0.911(0.862~0.961)	88.10	84.82	0.729	85.41

3 讨论

3.1 老年全身麻醉颈椎术后谵妄发生现状 目前,关于术后谵妄研究报道较多,不同研究中谵妄发生率差异较大。Morino 等^[7]报道,髋关节置换术后谵妄发生率在 13.0%~41.0%。Kobayashi 等^[8]报道,90 岁以上高龄脊柱手术患者术后谵妄发生率为 31.43%。而孙小珊等^[9]报道,老年脊柱术后谵妄发生率为 10.47%,其中颈椎术后谵妄发生率高于腰椎。对于老年全身麻醉颈椎术后谵妄发生情况报道较少,本研究训练集谵妄发生率为 15.79%,验证集谵妄发生率为 18.03%。以上研究结果表明,不同疾病、年龄、手术部位谵妄的发生率不同,因此本研究针对老年全身麻醉颈椎术后患者进行研究,并据此构建三种风险评估模型。

3.2 3 种模型共同评估变量 本研究发现,年龄、手术时长、糖尿病、术中输血、术中 PaO₂ 均为 3 种模型共同评估变量。(1)年龄。Kolk 等^[10]报道,年龄是复杂头颈皮瓣术后谵妄的风险因素。随着年龄的增

加,神经细胞死亡以及神经退行性变,加之中枢神经递质及其受体减少,认知功能降低,促进了谵妄的发生。医护人员应该重视老龄患者术后谵妄的预防,尽早开展时间、地点、定向力认知功能训练,降低谵妄发生。(2)手术时间。张晓洁等^[11]报道,手术时间>120 min 是患者术后谵妄的风险因素。手术时间长可导致手术创伤引起的应激反应时间长,患者机体代偿能力降低,加之麻醉抑制了胆碱能神经递质,促进了谵妄的发生^[12]。医护人员应该充分做好术前准备、密切监测术中电生理和血压,尽量缩短手术时长。(3)糖尿病。姚璐璐等^[13]报道,糖尿病是主动脉瓣膜置换术后谵妄的风险因素,与研究^[14]的结果类似。糖尿病可导致神经递质传导减慢、皮质萎缩,影响了脑部神经传导和代谢物质的交换,从而促进谵妄的发生。针对糖尿病患者,临床应该密切监测其血糖控制情况,及时降低血糖水平,减少糖尿病所引起的谵妄风险。(4)术中 PaO₂ 过低。周鸿业等^[15]报道,术中 PaO₂ 过低是颈椎麻醉术后患者谵

妄的风险因素。当 PaO_2 降低时, Hb 氧合不足, 脑细胞摄氧不足, 从而导致脑细胞氧化代谢不足, 进而促进谵妄的发生。造成 PaO_2 降低的因素较多, 因此医护人员应对患者进行全面评估, 指导有吸烟史且戒烟超过 2 周患者正确咳嗽、咳痰, 有效运用呼吸训练装置, 强化训练重要性。指导并开展术前肺康复训练, 包括药物康复、心理康复、物理康复等, 降低术中 PaO_2 过低风险。(5) 术中输血。与术中 PaO_2 过低类似, 当术中出血量大或者基础血红蛋白偏低时, Hb 氧合不足, 需要输血, 因此术中输血同样反应了脑细胞摄氧不足, 从而引起谵妄的发生^[16]。术前护理人员应该详细询问患者的输血史、出血史和用药史, 并进行体格检查及实验室检测, 以准确评估患者的贫血状态和凝血功能。对于服用抗凝药、抗血小板药和其他可能影响止血功能药物的患者, 应根据手术类型、手术出血风险和药物特点采取相应的防治方案, 从而降低术中输血概率。

3.3 三种模型非共同评估变量的分析 (1) 术后脑脊液渗漏。该变量进入列线图、随机森林模型, 但未进入决策树 CART 模型。脑脊液渗漏对谵妄影响的研究报道较少。脑脊液的渗漏可以引起间脑和脑干的移位, 引起脑干肿胀和静脉充血, 进而导致意识和思维混乱^[17], 从而引起谵妄的发生。术后医护应嘱咐患者头高脚底半卧位, 且尽量避免可能导致颅内压急剧升高的活动, 如用力排便、咳嗽、打喷嚏等, 适当应用缓泻剂和粪便软化剂降低便秘风险。(2) 手术方式。该变量仅进入决策树 CART、随机森林模型中, 但未进入列线图模型。颈椎手术常包括颈椎前路手术和后路手术, 不同的手术方式适用指针有一定差异。目前暂无颈椎手术方式对术后谵妄影响的研究报道, 少量的研究主要集中于两种手术方式并发症的差异^[18]。申庆丰等^[19]报道, 颈椎后路手术的患者手术时长和出血量均高于颈椎前路手术的患者。因此可能由于后路手术导致的手术时间延长以及出血量增加, 从而引起谵妄风险升高。

3.4 模型的准确性与实用性分析 黄年华等^[20]基于列线图构建了 PCI 术后谵妄预测模型 ($\text{AUC} = 0.89$), 但该模型未经外部验证。王紫微等^[21]基于 Logistic 原理构建了脊髓损伤患者谵妄模型 ($\text{AUC} = 0.818$), 但模型计算公式复杂, 无法满足临床应用的要求。本研究发现在训练集, 列线图、决策树 CART、随机森林模型评估 AUC 分别为 0.888、0.874、0.911, 验证集模型评估 AUC 分别为 0.895、0.884、0.911, 随机森林模型评估效能最优。从模型算法和运用原理来看, 列线图模型对数据要求较高(如阳性样本数量、是否共线性), 容易因数据影响模

型评估效能^[22]。决策树 CART 模型呈现和应用过程较为复杂, 容易出现泛化和拟合能力偏弱的现象^[23]。随机森林模型可自动分析因子间的交互作用和非线性作用, 可通过有目的放飞对数据进行利用, 对数据要求低^[24]。临床医护人员可以采用随机森林模型, 及时开展谵妄评估, 针对高风险的患者, 可综合应用影响因素中对应的护理干预、药物治疗、心理干预等手段, 降低患者术后谵妄发生的风险。

4 小结

本研究基于临床资料利用列线图、决策树 CART、随机森林构建了老年全身麻醉颈椎术后谵妄风险评估模型。模型的构建规避了单一指标评估效能有限的弊端, 有助于临床及早识别术后谵妄的高风险人群, 及时采取干预措施, 降低谵妄的发生。本研究术前抑郁、焦虑、睡眠质量等变量未纳入模型分析中, 可能影响模型的评估效能, 未来研究应该纳入更全面的变量。此外, 模型仅在一个医院进行了验证, 存在泛化不足的局限, 因此未来需对模型评估效能进行多中心、大样本的进一步验证。

【参考文献】

- [1] LOBO-VALBUENA B, GORDO F, ABELLA A, et al. Risk factors associated with the development of delirium in general ICU patients. A prospective observational study[J/OL]. [2024-05-01]. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0255522>. DOI: 10.1371/journal.pone.0255522.
- [2] 李洁, 蒋运兰, 廖诗沁, 等. 老年人围术期睡眠障碍与术后谵妄发生风险关系的 Meta 分析[J]. 军事护理, 2023, 40(6): 96-100.
- [3] 王紫微, 张颖惠, 张丽玉, 等. ICU 颈脊髓损伤病人谵妄发生风险预测模型的构建与验证[J]. 护理研究, 2023, 37(18): 3307-3313.
- [4] 黄宛冰, 张玉芬, 吴前胜, 等. 基于 Cox 回归的 Stanford B 型主动脉夹层术后谵妄预测模型的构建[J]. 护理学杂志, 2023, 38(3): 27-31.
- [5] 李汉明, 李广州, 王清, 等. 老年患者颈椎手术后发生谵妄的危险因素分析[J]. 中国脊柱脊髓杂志, 2024, 34(5): 463-468.
- [6] ALDECOA C, BETTELLI G, BILOTTA F, et al. European Society of Anaesthesiology evidence-based and consensus-based guideline on postoperative delirium[J]. Eur J Anaesthesiol, 2017, 34(4): 192-214.
- [7] MORINO T, HINO M, YAMAOKA S, et al. Risk factors for delirium after spine surgery: an age-matched analysis[J]. Asian Spine J, 2018, 12(4): 703-709.
- [8] KOBAYASHI K, IMAGAMA S, SATO K, et al. Postoperative complications associated with spine surgery in patients older than 90 years: a multicenter retrospective study[J]. Global Spine J, 2018, 8(8): 887-891.
- [9] 孙小珊, 孔珉珉, 孟勇, 等. 老年脊柱手术患者术后谵妄发生情况调查及相关影响因素分析[J]. 现代生物医学进展, 2022, 22(13): 2559-2563.
- [10] KOLK A, SCHWARZER C, Wolff K D, et al. Factors associated with postoperative delirium in patients undergoing complex head and neck flap surgery[J]. J Oral Maxillofac Surg, 2022, 80(2): 372-379.
- [11] 张晓洁, 陈雪, 饶丽华, 等. 不同营养状况的全麻下行防旋型股骨近端髓内钉手术老年病人术后谵妄发生情况及影响因素分析[J]. 实用老年医学, 2023, 37(10): 1031-1034.

(下转第 66 页)