

基于随机森林算法的心电图引导手臂输液港 静脉导管头端精准定位的影响因素研究

徐湘妍^{1,2}, 吉昕宇¹, 宋昀杰¹, 徐晓霜¹, 徐海萍³, 张汝阳¹

(1. 南京医科大学公共卫生学院 生物统计学系, 江苏 南京, 211166;

2. 复旦大学公共卫生学院 流行病学与卫生统计学系, 上海, 200032;

3. 南京医科大学第一附属医院 乳腺科, 江苏 南京, 210036)

摘要: 目的 基于随机森林算法探讨心电图技术引导手臂输液港静脉导管头端精准定位的影响因素。方法 选取338例植入输液港的女性乳腺癌化疗患者作为研究对象,应用心电图引导静脉导管头端定位,并采集患者的相关信息。采用随机森林方法探讨心电图引导手臂输液港静脉导管头端精准定位的影响因素,应用分类树拟合导管头端定位精准性模型。结果 身高、体质量和年龄与导管头端定位精准性相关。所构建模型的准确率(0.84)、灵敏度(0.98)、阳性预测值(0.85)尚可,特异度(0.19)、阴性预测值(0.54)、曲线下面积(AUC)(0.67)较低。结论 临床人员应重点关注低身高(<155 cm)、高体质量(>52 kg)、年龄49~63岁的乳腺癌患者,因为此类患者易发生静脉导管头端定位不精准的情况。

关键词: 随机森林; 心电图; 手臂输液港; 静脉导管定位; 影响因素

中图分类号: R 473.73; R 195.1 文献标志码: A 文章编号: 1672-2353(2021)19-007-05 DOI: 10.7619/jcmp.20213117

A study on influence factors of accurate positioning of venous catheter tip in arm infusion port guided by electrocardiogram based on random forest algorithm

XU Xiangyanyu^{1,2}, JI Xinyu¹, SONG Yunjie¹,
XU Xiaoshuang¹, XU Haiping³, ZHANG Ruyang¹

(1. Department of Biostatistics, School of Public Health of Nanjing Medical University, Nanjing, Jiangsu, 211166; 2. Department of Epidemiology and Biostatistics, School of Public Health of Fudan University, Shanghai, 200032; 3. Department of Breast Surgery, the First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu, 210036)

Abstract: Objective To explore influence factors of accurate positioning of venous catheter tip in arm infusion port guided by electrocardiogram based on random forest algorithm. **Methods** A total of 338 female breast cancer patients with chemotherapy implanted in transfusion port were selected as research objects. Electrocardiogram was used to guide the positioning of the venous catheter tip, and relevant information of the patients was collected. Random forest method was used to explore the influencing factors of electrocardiogram guided precise positioning of intravenous catheter tip in arm infusion port, and classification tree was used to fit the precise positioning model of catheter tip. **Results** Height, weight and age were significantly associated with the accuracy of catheter head positioning. The accuracy (0.84), sensitivity (0.98), positive predictive value (0.85) of the constructed model were acceptable, and the specificity (0.19), negative predictive value (0.54) and area under the Curve (AUC) (0.67) were low. **Conclusion** Clinical staff should pay attention to breast cancer patients with low height (<155 cm), high body mass (>52 kg) and age of 49 to 63 years, because such patients are prone to occur inaccurate positioning of the head of venous catheter.

Key words: random forest; electrocardiogram; arm infusion port; positioning of venous catheter; influencing factors

收稿日期: 2021-08-03

基金项目: 江苏省妇幼健康科研项目(F201821); 江苏省自然科学基金面上项目(BK20191354); 江苏省人民医院“临床能力提升工程”(JSPH-NB-2020-5); 南京医科大学护理研究中心肿瘤护理研究分中心开放课题(NYZLKF-2-202104)

通信作者: 张汝阳, E-mail: zhangruiyang@njmu.edu.cn; 徐海萍, E-mail: xhphappy9@126.com

手臂输液港是恶性肿瘤患者的常用治疗措施,临床常采用 X 射线定位静脉导管头端。近年来,采用心电图进行静脉导管头端定位者逐渐增多。完全植入式静脉输液港(TIVAP)是完全植入人体内的闭合输液装置,包括头端位于上腔静脉的导管部分和埋植于皮下的港座部分,其中静脉导管部分与经外周静脉置入中心静脉导管(PICC)相似。临床应用输液港可减少许多并发症的发生,如静脉炎、血栓、纤维蛋白鞘等^[1]。与 PICC 相比,输液港增加了 1 个港座,保留的时间更长,更适合长期化疗^[2]。输液港应用过程中的关键问题之一是保证导管头端定位的精准性,若静脉导管头端位置定位不精准,可能会引发相关并发症^[3]。心电图引导技术具有安全性高、精准率高、无辐射等优点^[4-7],近年来已被广泛应用于临床。LIU G 等^[8]的一项关于心电图引导 PICC 研究的 Meta 分析结果显示,成年患者置入 PICC 过程中采用心电图引导相较于 X 射线法具有更高的定位精度,精准度可达 89.7%,且没有重大的安全性问题。与传统的解剖标志定位相比,心电图引导方法也具有明显的优势^[9-10]。本研究基于随机森林算法探讨心电图技术引导手臂输液港静脉导管头端精准定位的影响因素,现报告如下。

1 资料与方法

1.1 一般资料

选取 2018 年 9 月—2019 年 9 月南京医科大学第一附属医院乳腺病科收治的 338 例植入输液港的女性乳腺癌化疗患者作为研究对象。应用心电图引导静脉导管头端定位,并收集患者的相关信息。患者年龄、婚姻状况、文化程度、家庭月收入、医疗费用支付方式的资料来自问卷调查。新辅助治疗、手术方式来自患者在医院接受治疗时的资料。身高、体质量、输液港植入侧臂围数据于置港前由护士统一测量所得。导管置入位置、导管置入静脉、导管置入长度、穿刺点至胸锁关节长度数据在植入手臂港过程中获取。

1.2 导管长度及头端定位精准性的判断方法

导管进入上腔静脉时,出现高尖 P 波。导管到达上腔静脉与右心房的上壁交界连接点(CAJ)时,P 波达到最高峰。若导管进入右心房,则会出现负向 P 波。本研究中,当 P 波振幅达到 QRS 波高度 50% 时,将此时的导管长度视为导管置入静脉的长度。手臂输液港囊袋中的港座与导管连接帽的长度为 1cm,最终记录的导管长度为导管

置入静脉的长度再加上 1 cm。

2016 年美国静脉输液护理学会发布的新版《输液治疗实践标准》^[11]指出,导管的头端位置位于 CAJ 为安全性最佳。本研究中,导管头端位置在 CAJ 或位于 CAJ 上端 1 个椎体,均视为导管头端位置精准。

1.3 随机森林法筛选重要变量

随机森林是一种机器学习算法,由 BREIMAN L^[12]于 2001 年提出,其基本思想是从 1 个样本量为 N 、特征数为 M 的数据集中,利用 Bootstrap 法有放回地抽样产生 1 个样本量为 N 、特征数为 m [常设为 \sqrt{M}] 的 Bootstrap 数据用于模型训练。理论上,原始数据中约有 1/3 的样本不会被抽取,称为袋外(OOB)数据。重复以上过程 k 次,可得含有 k 棵决策树的随机森林。本研究设置 5 000 棵决策树。

决策树首先选定所有样本为根节点,继而根据某变量生成 2 个子节点。根据节点中各结局的频率,可计算出根节点与 2 个子节点之间基尼不纯度的差值。建树时,选取使基尼不纯度变化最大的变量划分根节点。每个子节点都重复上述过程进行划分,直至子节点中基尼不纯度不再减小。在整个森林中,可以计算出某变量使整个森林平均减少的基尼不纯度,称为平均不纯度减小(可作为变量的重要性评分)。

本研究采用滑动窗口序贯向前选择法(SWSFS)筛选变量^[13-14]。按照重要性评分从大到小依次纳入变量,每加入 1 个变量重新进行随机森林分析。选择使得 OOB 错误率最小时对应变量集进行后续分析。

1.4 基于重要变量构建决策树

首先,将原数据集按照 2:1 比例随机划分为训练集与测试集。基于训练集,采用所筛选的变量构建“导管头端定位精准”的决策树。在测试集中,评价决策树的灵敏度、特异度、阳性预测值、阴性预测值、准确率和受试者工作特征(ROC)曲线的曲线下面积(AUC)。重复上述步骤 1 000 次,得到各统计量的抽样分布。

1.5 统计学分析

根据导管头端定位精准性将患者分为精准组、不精准组,对 2 组患者各个指标进行描述与比较。定量资料服从正态分布的变量采用 $(\bar{x} \pm s)$ 描述,组间方差齐性时,组间比较采用 t 检验,否则采用 t' 检验;定性资料采用 $[n(\%)]$ 描述,组间比较采用 χ^2 检验或 Fisher 确切概率法;有序分类变

量采用[$n(\%)$]描述,组间比较采用 Wilcoxon 秩和检验。所有假设检验 P 值取双侧,差异性检验水准设为 0.05,方差齐性检验水准设为 0.10,置信区间(CI)可信度设为 95%。采用 R 软件(版本号 4.0.2)对数据进行统计分析。描述性统计采用 tableone 包,随机森林采用 randomForest 包,决策树采用 rpart 包和 rattle 包,灵敏度与特异度等采用 caret 包和 pROC 包实现。

2 结果

2.1 2 组患者基线资料描述性统计分析

6 个连续变量(年龄、身高、体质量、穿刺点至胸锁关节长度、导管置入长度、输液港植入侧臂围)的 t 检验结果提示,2 组身高、导管置入长度比较,差异有统计学意义($P < 0.05$)。7 个无序分类变量(导管置入位置、导管置入静脉、婚姻状

况、文化程度、医疗费用支付方式、新辅助治疗、手术方式)的卡方检验结果提示,2 组上述指标比较,差异无统计学意义($P > 0.05$)。等级变量(家庭月收入)的秩和检验结果显示,2 组家庭月收入比较,差异无统计学意义($P > 0.05$)。见表 1。

2.2 随机森林筛选重要变量

采用随机森林法评价 13 个变量的重要性评分,由高至低依次为:身高 17.7 分,年龄 17.3 分,体质量 13.5 分,输液港植入侧臂围 12.1 分,穿刺点至胸锁关节长度 11.9 分,家庭月收入 5.4 分,教育程度 3.8 分,医疗费用支付方式 3.6 分,置入位置 2.8 分,新辅助治疗 2.2 分,置入静脉 2.1 分,手术方式 1.9 分,婚姻状况 1.2 分。当进入模型的变量数量为 3 时,模型错判率最低(见图 1),即前 3 个变量使得模型 OOB 数据错误率最低,因此重要变量分别为身高、年龄和体质量。

表 1 精准组与不精准组患者的基线资料比较($\bar{x} \pm s$) [$n(\%)$]

指标	不精准组($n=59$)	精准组($n=279$)	$t/\chi^2/W$	P
导管置入长度/cm	36.08 ± 2.58	37.46 ± 2.41	-3.94	<0.01
年龄/岁	51.78 ± 9.74	49.80 ± 10.42	1.34	0.18
身高/cm	155.42 ± 4.90	157.94 ± 5.24	-3.40	<0.01
体质量/kg	58.97 ± 6.11	58.59 ± 7.48	0.35	0.72
穿刺点至胸锁关节长度/cm	30.39 ± 2.35	30.80 ± 2.36	-1.20	0.23
输液港植入侧臂围/cm	26.75 ± 2.33	26.19 ± 2.41	1.63	0.10
置入位置	资料缺失 0	1(0.36)	2.99	0.78
	左上臂 27(45.76)	138(49.46)		
	右上臂 32(54.24)	140(50.18)		
置入静脉	资料缺失 0	1(0.36)	3.26	0.80
	贵要静脉 50(84.75)	242(86.74)		
	肱静脉 9(15.25)	36(12.90)		
婚姻状况	资料缺失 0	3(1.08)	6.66	0.85
	未婚 1(1.69)	4(1.43)		
	已婚 57(96.61)	269(96.42)		
	离异 0	1(0.36)		
	丧偶 1(1.69)	2(0.72)		
医疗费用支付方式	资料缺失 0	4(1.43)	3.13	0.32
	农保 1(1.69)	7(2.51)		
	居民保险/职工医保 30(50.85)	136(48.75)		
	商业保险 1(1.69)	0		
	公费 1(1.69)	4(1.43)		
	自费 26(44.07)	128(45.88)		
新辅助治疗	资料缺失 0	2(0.72)	3.08	0.79
	否 43(72.88)	206(73.84)		
	是 16(27.12)	71(25.45)		
教育程度	资料缺失 0	2(0.72)	8 118.00	0.93
	小学及以下 13(22.03)	62(22.22)		
	初中及高中 32(54.24)	146(52.33)		
	大学及以上 14(23.73)	69(24.73)		
家庭月收入	资料缺失 1(1.69)	4(1.43)	7 413.50	0.37
	≤3 000 元人民币 3(5.08)	14(5.02)		
	>3 000 ~ 5 000 元人民币 21(35.59)	83(29.75)		
	>5 000 ~ 10 000 元人民币 20(33.90)	99(35.48)		
	>10 000 元人民币 14(23.73)	79(28.32)		
手术方式	资料缺失 0	4(1.43)	0.79	0.33
	保乳治疗 7(11.86)	50(17.92)		
	改良根治术 52(88.14)	225(80.65)		

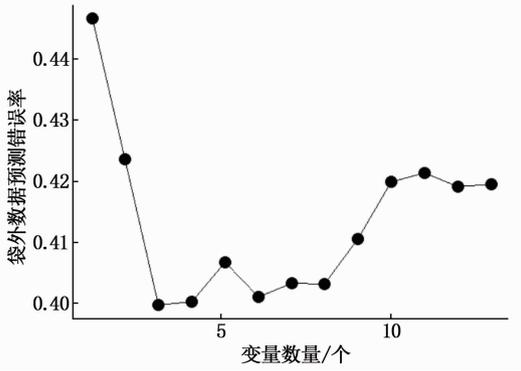


图 1 各变量的袋外数据预测错误率

2.3 基于重要变量构建决策树

在训练集中,基于重要变量构建决策树。在测试集和全人群数据中,评价该决策树的效果。结果提示,所构建模型的准确率(0.84)、灵敏度(0.98)、阳性预测值(0.85)尚可,特异度(0.19)、阴性预测值(0.54)、曲线下面积(AUC)(0.67)较低。见图 2、表 2。

表 2 决策树模型在测试集及全人群中的各统计量结果

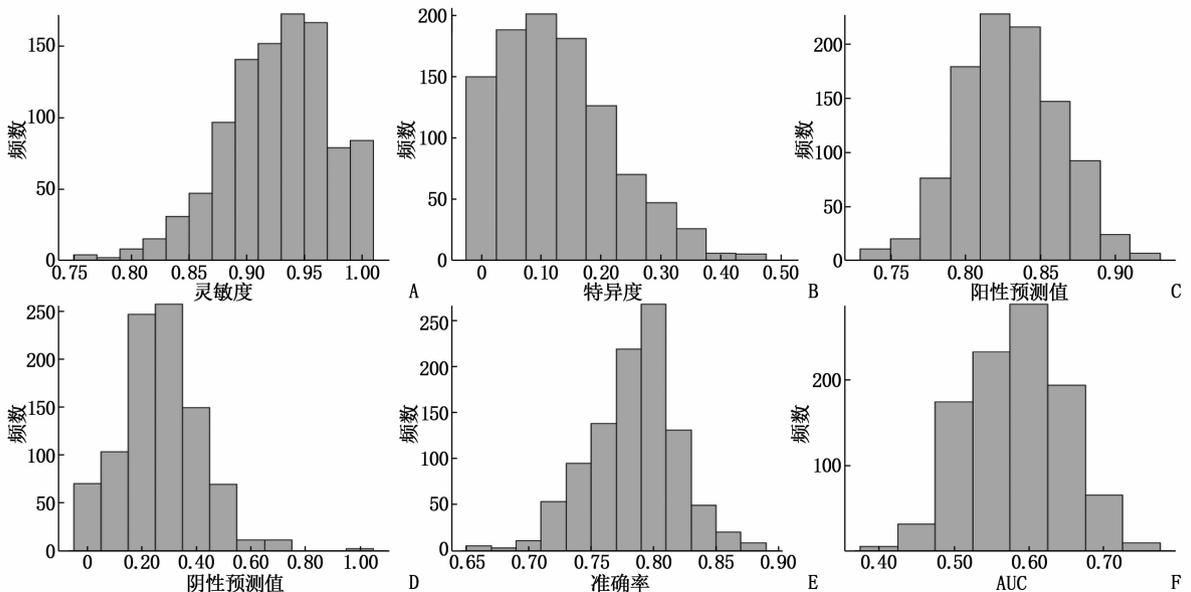
统计量	测试集			全人群
	2.5%分位数	中位数	97.5%分位数	
灵敏度	0.83	0.93	1.00	0.98
特异度	0.00	0.12	0.33	0.19
阳性预测值	0.77	0.83	0.89	0.85
阴性预测值	0.00	0.29	0.63	0.54
准确率	0.72	0.79	0.85	0.84
AUC	0.46	0.59	0.70	0.67

AUC: 曲线下面积。

基于全人群数据集的决策树提示,身高 < 155 cm、体质量 > 52 kg、年龄 49 ~ 63 岁的患者更易发生手臂输液港静脉导管头端定位不精准的情况。见图 3。

3 讨论

本研究随机森林和决策树模型的结果提示,身高、体质量和年龄可能是手臂输液港静脉导管



A: 灵敏度的抽样分布; B: 特异度的抽样分布; C: 阳性预测值的抽样分布; D: 阴性预测值的抽样分布; E: 准确率的抽样分布; F: AUC 的抽样分布。

图 2 决策树模型在测试集中各统计量的抽样分布

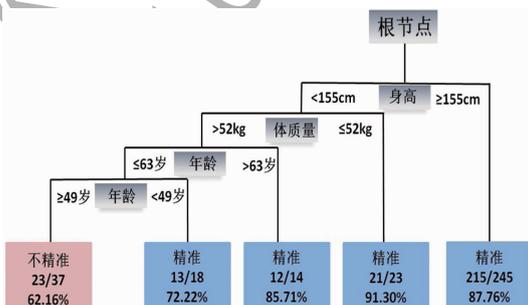


图 3 决策树模型在全人群中的可视化图

头端定位不精准的影响因素。身高较低的患者,上腔静脉一般较短,因此操作过程中导管的移动空间较小,易导致头端定位不精准,与徐海萍等^[15] Logistic 回归模型研究结论一致。除了患者本身的特征外,心电图图像的稳定性也是影响导管头端定位精准性的重要因素。郭素萍等^[16] 研究表明,约 5% 的患者在心电图引导静脉导管置入术中不能获得稳定心电图。此外有学者^[17] 指出,除了心电图信号本身外,其他无线电信号和周围

医疗设备的干扰也可能对心电图的图像产生影响。

多项研究^[15, 18-20]结果显示,不同体质量、年龄对心电图 P 波改变或静脉导管头端定位精准性比较,差异无统计学意义($P > 0.05$)。这可能是由于大多数研究只考虑了体质量、年龄的主效应,而忽视了其与其他因素间的联合效应或交互作用。本研究采用的决策树模型能够很好地识别变量之间的复杂效应。然而,体质量、年龄影响导管头端定位精准性的机制仍待进一步研究。本研究结果提示,临床医师应着重关注身高 < 155 cm、体质量 > 52 kg、年龄 49 ~ 63 岁的患者,因为此类患者更易发生静脉导管头端定位不精准的情况。

本研究采用随机森林算法评估了静脉导管尖端定位精确性的影响因素,随机森林可对变量进行重要性排序,排序结果显示了各变量对结局的贡献大小。作为一个初步探索性研究,本研究为后续确证性研究提供了思路。本研究局限性在于仅以女性乳腺癌化疗患者为研究对象,且只收集了常见的、易获取的研究指标,未来有待选取其他肿瘤患者为研究对象且扩大样本量、研究指标进行更深入的后续研究。

参考文献

- [1] 容桂荣,文琼仙. 外周化疗性静脉炎的发病机制及防治进展[J]. 华夏医学, 2020, 33(3): 202-205.
- [2] 夏薇,杨晓梅,刘佳红,等. 腔内心电图定位在卵巢癌患者手臂植入式输液港中的应用研究[J]. 西北国防医学杂志, 2020, 41(6): 387-390.
- [3] 周青,江智霞,代永娅,等. 特征性 P 波在腔内心电图引导 PICC 尖端定位中的应用研究进展[J]. 护理研究, 2020, 34(4): 641-646.
- [4] LI W, XU R, FAN D. Clinical application of electrocardiogram-guided tip positioning in peripheral inserted central catheters placement[J]. J Cancer Res Ther, 2018, 14(4): 887-891.
- [5] KRISHNAN A K, MENON P, GIREESH KUMAR K P, et al. Electrocardiogram-guided technique: an alternative method for confirming central venous catheter tip placement[J]. J Emerg Trauma Shock, 2018, 11(4): 276-281.
- [6] LI A, JIAO J, ZHANG Y, et al. A randomized controlled study of bedside electrocardiograph-guided tip location technique

and the traditional chest radiography tip location technique for peripherally inserted central venous catheter in cancer patients[J]. Indian J Med Res, 2018, 147(5): 477-483.

- [7] 杨晨,朱超,宋迪. 腔内心电图辅助 PICC 导管尖端实时定位技术在恶性肿瘤患者中的应用[J]. 当代护士: 学术版, 2021, 28(6): 113-115.
- [8] LIU G, HOU W, ZHOU C, et al. Meta-analysis of intracavitary electrocardiogram guidance for peripherally inserted central catheter placement[J]. J Vasc Access, 2019, 20(6): 577-582.
- [9] YUAN L, LI R, MENG A, et al. Superior success rate of intracavitary electrocardiogram guidance for peripherally inserted central catheter placement in patients with cancer: a randomized open-label controlled multicenter study[J]. PLoS One, 2017, 12(3): e0171630.
- [10] YIN Y X, GAO W, LI X Y, et al. Insertion of peripherally inserted central catheters with intracavitary electrocardiogram guidance: a randomized multicenter study in China[J]. J Vasc Access, 2019, 20(5): 524-529.
- [11] 彭娜. 2016 年 INS 输液治疗实践标准: 血管通路装置的选择和置入[J]. 现代医药卫生, 2017, 33(9): 1285-1287, 1291.
- [12] BREIMAN L. Random Forests[J]. Machine Learning, 2001, 45(1): 5-32.
- [13] ZHANG R, LAI L, HE J, et al. EGLN2 DNA methylation and expression interact with HIF1A to affect survival of early-stage NSCLC[J]. Epigenetics, 2019, 14(2): 118-129.
- [14] 张秋伊,赵杨,魏永越,等. 高维 DNA 甲基化数据的随机森林降维分析[J]. 中华疾病控制杂志, 2016, 20(6): 630-633.
- [15] 徐海萍,汪静雯,丁晓慧,等. 心电图引导手臂输液港静脉导管头端精准定位 111 例实证研究[J]. 中华肿瘤防治杂志, 2019, 26(23): 1800-1804.
- [16] 郭素萍,徐莹,王龙凤,等. 腔内心电图引导在 PICC 置管长度预测中的应用效果[J]. 2016, 22(20): 2940-2942, 2943.
- [17] 姚文坡,毛靖宁,刘铁兵. 心电图机检定中的工频干扰及抑制方法[J]. 医疗卫生装备, 2015, 36(5): 89-91.
- [18] 李真,邵珠美. PICC 置管腔内心电图 P 波改变及影响因素调查[J]. 当代护士: 上旬刊, 2019, 26(6): 22-24.
- [19] 郭玲,秦英,王国蓉,等. 成人中心静脉置管腔内心电图稳定性及影响因素研究[J]. 中华护理杂志, 2015, 50(6): 724-727.
- [20] 曹奕忠,张小惠. 成人中心静脉置管术中使用时心电图行导管头端定位时心电图稳定性及其影响因素[J]. 黑龙江医学, 2017, 41(4): 310-311.

(本文编辑:陆文娟)