

机器学习模型在预测肾结石输尿管软镜碎石术后早期结石清除率中的应用

朱学华^{1*}, 杨明钰^{2*}, 夏海缀¹, 何 为¹, 张智荧¹, 刘余庆^{1△}, 肖春雷^{1△}, 马路林¹, 卢 剑¹

(1. 北京大学第三医院泌尿外科, 北京 100191; 2. 北京大学信息科学与技术学院, 北京 100871)

[摘要] **目的:** 基于随机森林和 XGBoost 两种机器学习算法建立预测模型, 探讨其对肾结石患者行输尿管软镜碎石术 (flexible ureteroscopic lithotripsy, fURL) 后早期结石清除率 (stone-free rate, SFR) 的预测价值。 **方法:** 回顾性分析 201 例行 fURL 的肾结石患者的临床资料, 根据术后是否达到结石清除标准, 将患者分为结石清除组和结石残留组。比较两组患者年龄、体重指数 (body mass index, BMI)、结石数目、结石体积、结石密度和肾积水等因素的差异。对于肾下盏结石, 需测量肾脏解剖相关指标, 包括肾盂漏斗部夹角、肾下盏宽度、肾下盏长度及肾盂肾下盏高度。将上述潜在影响因素分别纳入随机森林和 XGBoost 算法建立预测模型, 绘制受试者工作曲线, 检验模型预测价值。前瞻性收集 71 例患者的临床资料对模型进行外部验证。 **结果:** 201 例 fURL 手术均顺利完成, 一期手术早期 SFR 为 61.2%。利用随机森林和 XGBoost 算法建立预测模型并得到不同变量预测重要性评分, 随机森林模型和 XGBoost 模型曲线下面积均为 0.77。应用 71 例样本对模型进行外部验证结果显示, 随机森林模型对检测样本的预测总准确率、特异度及灵敏度分别为 74.6%、82.6% 和 60.0%, XGBoost 模型对检测样本的预测总准确率、特异度及灵敏度分别为 80.3%、87.0% 和 68.0%。在两种模型中, 预测重要性评分排名前四位的变量均为结石体积、平均结石密度、最大结石密度和 BMI。 **结论:** 基于随机森林和 XGBoost 算法建立的机器学习模型可准确预测肾结石患者 fURL 术后早期结石清除状态, 有利于术前评估及临床决策。结石体积、平均结石密度、最大结石密度和 BMI 可能是影响肾结石 fURL 术后 SFR 的重要预测因素。

[关键词] 机器学习; 随机森林; XGBoost; 肾结石; 结石清除率; 预测模型

[中图分类号] R692.4 **[文献标志码]** A **[文章编号]** 1671-167X(2019)04-0653-07

doi: 10.19723/j.issn.1671-167X.2019.04.010

Application of machine learning models in predicting early stone-free rate after flexible ureteroscopic lithotripsy for renal stones

ZHU Xue-hua^{1*}, YANG Ming-yu^{2*}, XIA Hai-zhui¹, HE Wei¹, ZHANG Zhi-ying¹, LIU Yu-qing^{1△}, XIAO Chun-lei^{1△}, MA Lu-lin¹, LU Jian¹

(1. Department of Urology, Peking University Third Hospital, Beijing 100191, China; 2. School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China)

ABSTRACT Objective: To establish predictive models based on random forest and XGBoost machine learning algorithm and to investigate their value in predicting early stone-free rate (SFR) after flexible ureteroscopic lithotripsy (fURL) in patients with renal stones. **Methods:** The clinical data of 201 patients with renal stones who underwent fURL were retrospectively investigated. According to the stone-free standard, the patients were divided into stone-free group (SF group) and stone-residual group (SR group). We compared a number of factors including patient age, body mass index (BMI), stone number, stone volume, stone density and hydronephrosis between the two groups. For low calyceal calculi, renal anatomic parameters including infundibular angle (IPA), infundibular width (IW), infundibular length (IL) and pelvic calyceal height (PCH), would be measured. We brought above potential predictive factors into random forest and XGBoost machine learning algorithm respectively to develop two predictive models. The receiver operating characteristic curve (ROC curve) was established in order to test the predictive ability of the model. Clinical data of 71 patients were collected prospectively to validate the predictive models externally. **Results:** In this study, 201 fURL operations were successfully completed. The one-phase early SFR was 61.2%. We built two predictive models based on random forest and XGBoost machine learning algorithm. The predictive variables' importance scores were obtained. The area under

基金项目: 北京大学智慧医疗专项 (BMU2018ZHYL012) Supported by the Intelligent Medical Program of Peking University (BMU2018ZHYL012)

△ Corresponding author's e-mail, pku3uro@aliyun.com, xiaochunleixcl@sina.com

* These authors contributed equally to this work

网络出版时间: 2019-6-27 17:23:42 网络出版地址: <http://kns.cnki.net/kcms/detail/11.4691.R.20190624.1615.008.html>

the ROC curve (AUROC) of the two predictive models for early stone clearance status prediction was 0.77. In the study, 71 test samples were used for external validation. The results showed that the total predictive accuracy, predictive specificity and predictive sensitivity of the random forest and XGBoost models were 75.7%, 82.6%, 60.0%, and 81.4%, 87.0%, 68.0%, respectively. The first four predictive variables in importance were stone volume, mean stone density, maximal stone density and BMI in both random forest and XGBoost predictive models. **Conclusion:** The predictive models based on random forest and XGBoost machine learning algorithm can predict postoperative early stone status after fURL for renal stones accurately, which will facilitate preoperative evaluation and clinical decision-making. Stone volume, mean stone density, maximal stone density and BMI may be the important predictive factors affecting early SFR after fURL for renal stones.

KEY WORDS Machine learning; Random forest; XGBoost; Renal stones; Stone-free rate; Predictive model

泌尿系结石是常见的泌尿系疾病,中国人群的泌尿系结石发病率高达 5.8%^[1]。逆行肾内手术(retrograde intrarenal surgery, RIRS),尤其是输尿管软镜碎石术(flexible ureteroscopic lithotripsy, fURL),已成为处理肾结石安全、有效的治疗方式,在泌尿系结石治疗中的地位逐渐提高^[2]。但是,单次 fURL 术后的结石清除率(stone-free rate, SFR)存在较大差异(65%~79%),肾结石及肾集合系统的解剖因素都会影响 fURL 的手术效果^[3-4]。

目前已有的预测模型(如 RUSS 评分系统、S-ReSC 评分系统、Ito 列线图 and R. I. R. S. 评分系统等)的建立主要基于 Logistic 回归分析^[5-8]。Logistic 回归分析属于线性概率模型中的一种,其要求研究数据缺失值少,且需符合线性、正态分布、方差齐等条件,但临床数据易存在缺失值,影响 fURL 效果的结石及肾脏解剖因素较多,因素之间的作用方式复杂,因此,基于 Logistic 回归分析建立的预测模型在临床应用中存在较大局限性。

机器学习(machine learning)是人工智能(artificial intelligence, AI)领域探索分析大数据的强大工具。决策树算法是机器学习中一种非常经典的分类算法,经过数据集的训练,其能够高效判断出一个数据项所属的类别。随机森林和 XGBoost[极端梯度上升(extreme gradient boosting)的简称]均是基于决策树的机器学习算法,随机森林利用多棵决策树对样本数据进行训练、分类及预测^[9],XGBoost 的基本思想是为数众多的、准确率较低的分类器叠加组合为准确率较高的分类器^[10]。在医学领域,随机森林和 XGBoost 已被应用于肺部毛玻璃样结节良恶性预测和急诊科泌尿系感染风险预测等方面,具有很高的预测准确性^[11-12]。目前尚未见将机器学习应用于预测肾结石 fURL 术后 SFR 的研究,本研究基于随机森林和 XGBoost 算法建立预测模型,以期辅助预后评估及术前临床决策。

1 资料与方法

1.1 纳入及排除标准

纳入标准:术前完善泌尿系 CT 平扫检查确诊为单纯肾结石,行 fURL 治疗且术后行泌尿系 X 线平片(kidney-ureter-bladder, KUB)或 CT 评估结石清除情况的患者。

排除标准:肾结石合并输尿管结石或泌尿系畸形的患者,如合并肾盂输尿管连接部狭窄、马蹄肾、盆位肾、移植肾、髓质海绵肾等;脊柱畸形的患者,如严重的脊柱侧弯;肾盏憩室内结石的患者;存在无法纠正的尿路感染、心肺功能障碍、凝血功能障碍而无法耐受手术的患者;临床资料不完整的患者。

1.2 一般资料

回顾性分析 2016 年 1 月至 2017 年 12 月于北京大学第三医院泌尿外科行 fURL 的共 201 例肾结石患者的临床资料,其中男 119 例(59.2%),女 82 例(40.8%)。左肾结石 93 例(46.2%),右肾结石 64 例(31.8%),双肾结石 44 例(21.9%),其中双肾结石同期手术者 8 例,存在肾下盏结石者 126 例(62.7%)。前瞻性收集 2018 年 1 月至 2019 年 1 月共 71 例肾结石患者的临床资料,所有患者术前接受病史询问、常规术前检查、实验室检查及泌尿系 CT 平扫检查以评估肾结石情况及泌尿系解剖情况。

1.3 手术方法

麻醉方式为全身麻醉,患者取截石位,首先用输尿管硬镜探查输尿管并上行至肾盂输尿管交界处,直视下将 0.035 英寸(1 英寸=25.4 mm)黑泥鳅导丝置入肾盂,撤镜,保留导丝,沿导丝逆行置入 F11~13 或 F12~14 导引鞘,使其前端接近肾盂输尿管交界处(男性患者使用 45 cm 导引鞘,女性患者使用 35 cm 导引鞘)。撤离导丝,从鞘内直视下置入纤维或电子纤维输尿管软镜,探查肾盂及各组肾盏。找到结石后,由输尿管软镜工作通道置入 200 μm 钬激光光纤,根据具体情况调配输出功率(功率设

置 10~40 W) 将结石粉碎。由工作通道置入套石篮,套取较大结石碎片留作标本。直视下再次将导丝置入肾盂,撤出软镜及导引鞘,沿导丝逆行留置 F4.7 双 J 管,留置尿管术后毕^[13]。

1.4 术后评估

术后第 1~3 天行 KUB 或泌尿系 CT 平扫检查以评估术后早期结石清除情况^[14],其中 3 例患者因术前检查发现为 X 线阴性结石而术后行 CT 平扫检查。患者的基本信息包括性别、年龄、体重指数 (body mass index, BMI)。根据既往文献研究结果,结合临床经验,纳入的 fURL 术后 SFR 的潜在影响因素包括 BMI、肾积水情况、结石体积、结石数目、结石密度(最大密度、平均密度)和结石位置。

肾下盏结石需测量肾盂漏斗部夹角 (infundibulopelvic angle, IPA)、肾下盏宽度 (infundibular width, IW)、肾下盏长度 (infundibular length, IL) 及肾盂肾下盏高度 (pelvic caliceal height, PCH) 四项指标,具体计算及测量方法为^[15-16] (图 1): IPA ($^{\circ}$):肾下盏漏斗轴(连接肾盂中心点与结石累及肾盏底部中点)与输尿管肾盂轴(连接肾盂中心点与肾下极所对输尿管中点)的夹角;IW (mm):结石累及肾下盏漏斗部最窄处的宽度;IL (mm):结石累及肾下盏最远端至肾盂下极中心点的距离;PCH (mm):结石累及肾下盏底部水平至肾盂下极中心点的垂直距离。

双肾结石同期手术者,两侧结石情况分别评估。定义结石直径 <4 mm 为结石清除, ≥ 4 mm 为结石残留^[14]。结石体积 (mm^3) = 长度 (mm) \times 宽度 (mm) \times 高度 (mm) $\times \pi \times 1/6$ 。

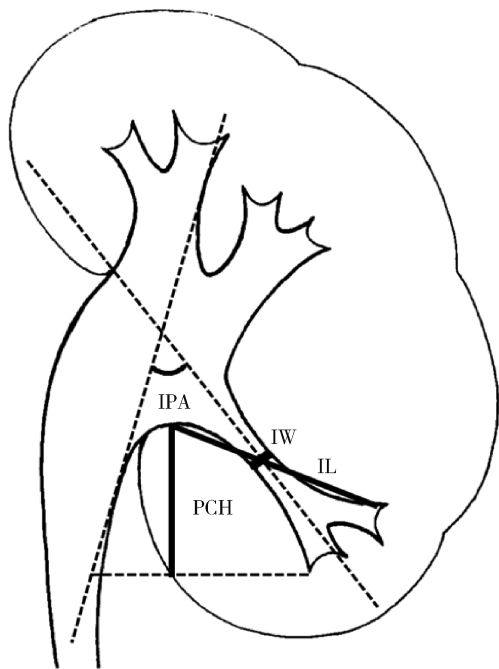
1.5 统计学分析

应用 Excel 软件建立数据库,根据术后早期是否达到结石清除标准将患者分为结石清除组和结石残留组,录入患者基本信息及术后 SFR 潜在影响因素。应用 SPSS 20.0 软件对数据进行统计学分析,比较两组之间可能影响 SFR 的变量的差异,包括患者年龄、BMI、肾积水情况、结石体积、结石数目、结石密度(最大密度、平均密度)、PCH、IPA、IL 和 IW。连续变量若符合正态分布,则用 $\bar{x} \pm s$ 表示,采用独立样本 t 检验分析数据;若连续性变量不符合正态分布,则采用中位数及范围区间表示,应用 Mann-Whitney U 检验进行数据分析;两组之间率的比较采用卡方检验。 $P < 0.05$ 为差异有统计学意义。

1.6 随机森林和 XGBoost 模型的建立

1.6.1 预测模型数据分割方法 原始数据集含有

201 例样本,采用放回随机抽样法,保证每个样本在下次抽样时仍可能被抽取。每次从原始数据中抽取约 2/3 的数据生成一个自主样本集作为训练集数据,每次未被抽到的约 1/3 的样本组成袋外数据作为测试集数据,用于估计分类误差和计算变量重要性评分。重复操作上述步骤 1 000 次,共得到 1 000 个包含 201 个样本(训练集样本量为 134,测试集样本量为 67)的数据集,构建决策树。



IPA, infundibulopelvic angle; IW, infundibular width; IL, infundibular length; PCH, pelvic caliceal height.

图 1 肾脏解剖结构相关指标测量方法示意图

Figure 1 Measurement methods for renal anatomic parameters

1.6.2 随机森林模型的构建 在训练每棵决策树时,只选取训练集中的一部分对其进行训练,以保证每棵决策树之间较好的独立性。利用每棵决策树对数据进行决策分类,随机森林对所有单棵树的分类结果进行综合判定。

1.6.3 XGBoost 模型的构建 XGBoost 同样是将多棵决策树进行集成的算法,但在训练新的决策树时,会针对性地增加之前训练过程中效果欠佳的数据点的权重,并将每棵决策树在训练集上的分类表现作为这个决策树最后的参考权重。

1.6.4 预测变量重要性评分 将变量重要性评分用于评价变量对于结局发生的影响,变量的重要性评分越高,则表明该变量越有能力对结局变量进行分类,具体实现过程参考文献[17]。

1.6.5 模型调参及外部验证 在应用随机森林和 XGBoost 算法时,有许多需要人为设定的变量,如构建分类树的个数 (ntree) 和决策树最大层数 (max-

length)等,计算机科学中将这 些 变 量 称 为 超 参 数。对于每个模型,我们以每组数据实验的平均受试者工作特征曲线下面积(the area under the receiver operating characteristic curve, AUROC)值作为评价指标。在对 201 例数据的训练过程中得到一组较优的超参数后,将这些超参数保存下来,将全部 201 例数据作为训练数据构建模型,然后用训练数据之外的 71 例数据对训练好的模型进行外部验证。

2 结果

201 例 fURL 手术均顺利完成,患者平均年龄(51.2 ± 13.7)岁,平均结石数目为(1.25 ± 0.80)个,平均结石体积为(1 325.5 ± 880.8) mm³,平均结石密度为(620.7 ± 297.2) Hu。结石清除组 123 例,结石残留组 78 例,一期手术早期 SFR 为 61.2%,两组之间相关临床数据的比较见表 1、2。

表 1 结石清除组与结石残留组患者基本信息及结石因素比较

Table 1 Comparison of demographic data and stone characteristics between stone-free group and stone-residual group

Items	Stone-free group (n = 123)	Stone-residual group (n = 78)	P value
Male, n (%)	75.0 (61.0)	44.0 (56.4)	0.163 ^a
Age/year, $\bar{x} \pm s$	51.4 ± 14.4	50.5 ± 13.4	0.655 ^b
BMI/(kg/m ²), median (min - max)	25.0 (18.0 - 33.1)	24.5 (17.7 - 35.0)	0.475 ^c
History of urologic operation, n (%)	32.0 (26.0)	27.0 (34.6)	0.192 ^a
Hydronephrosis, n (%)	67.0 (54.5)	53.0 (68.0)	0.058 ^a
Low calyceal stones, n (%)	75 (61.0)	51.0 (65.4)	0.887 ^a
Stone number, median (min - max)	1.0 (1.0 - 5.0)	1.5 (1.0 - 5.0)	0.000 ^c
Stone volume/mm ³ , median (min - max)	492.7 (8.37 - 15 055.4)	1 406.7 (133.2 - 11 992.0)	0.000 ^c
Mean stone density/Hu, $\bar{x} \pm s$	548.7 ± 273.7	734.2 ± 299.1	0.000 ^b
Maximal stone density/Hu, median (min - max)	1 049.0 (240.0 - 2 134.0)	1 266.00 (304.0 - 1 819.0)	0.000 ^c
Operation time/min, median (min - max)	89.0 (20.0 - 326.0)	111.0 (22.0 - 282.0)	0.000 ^c

Statistical significance was set at $P < 0.05$; a, Chi-square test; b, Student's *t* test; c, Mann-Whitney *U* test. BMI, body mass index.

表 2 结石清除组与结石残留组肾下盏解剖因素比较

Table 2 Comparison of low calyceal anatomic parameters between stone-free group and stone-residual group

Items	Stone-free group (n = 75)	Stone-residual group (n = 51)	P value
PCH/mm, $\bar{x} \pm s$	17.0 ± 4.3	21.8 ± 5.3	0.000 ^b
IPA/(°), median (min - max)	44.3 (29.2 - 71.5)	34.9 (22.3 - 44.6)	0.000 ^c
IL/mm, $\bar{x} \pm s$	22.9 ± 3.7	27.8 ± 5.0	0.000 ^b
IW/mm, median (min - max)	7.3 (1.0 - 22.0)	6.8 (1.0 - 9.5)	0.076 ^c

Statistical significance was set at $P < 0.05$; b, Student's *t* test; c, Mann-Whitney *U* test. PCH, pelvic calyceal height; IPA, infundibulopelvic angle; IL, infundibular length; IW, infundibular width.

单因素分析显示,两组之间在结石数目、结石体积、结石密度、PCH、IPA、IL 和手术时间等方面差异有统计学意义,而在性别、BMI、结石手术史、IW、肾积水及肾下盏结石等方面差异无统计学意义。将上述指标分别纳入随机森林和 XGBoost 算法建立预测模型,并计算预测变量重要性评分(图 2)。

模型在 201 例数据上进行 1 000 次数据实验,将每次数据实验训练得到的模型在测试集上的表现进行平均,绘制受试者工作特征曲线(receiver operating characteristic curve, ROC)。系统计算随机森林模型的 AUROC 为 0.77(95% CI:0.73 ~ 0.82),

XGBoost 模型的 AUROC 为 0.77(95% CI:0.72 ~ 0.82),表明模型具有良好的预测准确性(图 3)。

应用建立的预测模型对前瞻性收集的 71 例样本的术后结局进行预测,随机森林模型对检测样本的预测总准确率、特异度及灵敏度分别为 74.6%、82.6% 和 60.0%,XGBoost 模型对检测样本的预测总准确率、特异度及灵敏度分别为 80.3%、87.0% 和 68.0%(表 3)。

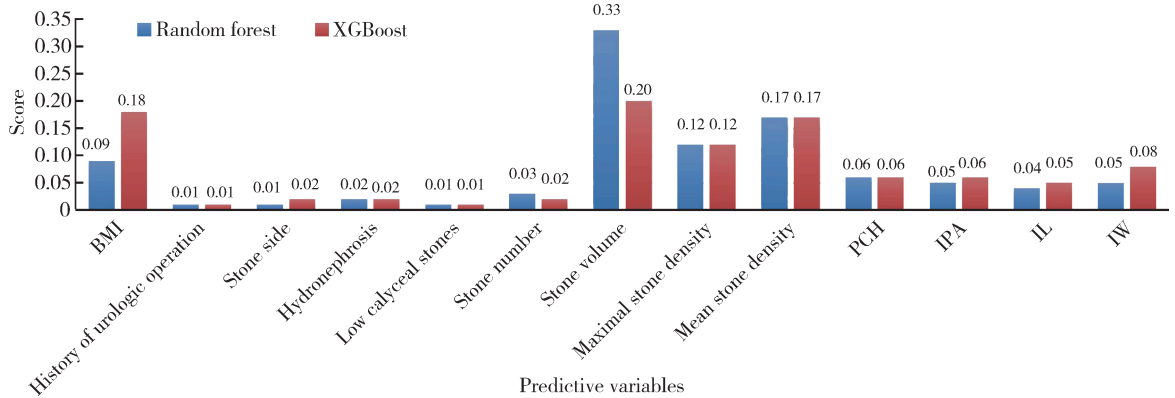
3 讨论

fURL 在临床应用中受到越来越多泌尿外科医

生的青睐。技术的提高以及碎石、取石设备的进步使其成为一种具有成功率高、并发症少、术后恢复快和住院时间短等优势的手术方式。目前,已有多家中心采用 fURL 治疗直径 ≥ 2 cm 的肾结石并取得了不错的效果^[18],但 fURL 的手术效果受到包括结石和肾脏解剖等多种复杂因素的影响。本研究通过检索既往文献以及结合临床经验,筛选出的潜在影响

因素包括 BMI、结石数目、结石体积、肾积水情况、PCH、IPA、IW 和 IL^[19-20]。

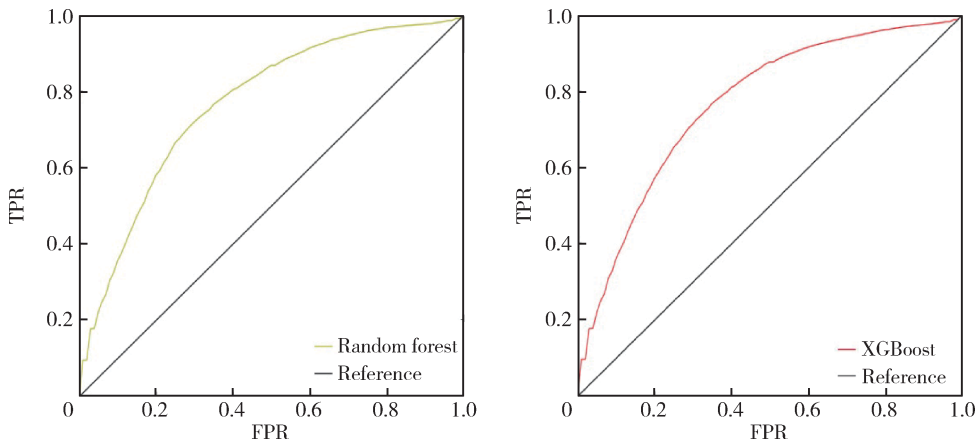
基于上述因素,我们利用随机森林和 XGBoost 算法建立预测模型,希望术前精确预测患者术后的结石清除状态,以便术前告知患者手术效果,充分尊重患者知情权及选择权,帮助患者选择最佳的治疗方式。



BMI, body mass index; IPA, infundibulopelvic angle; IW, infundibular width; IL, infundibular length; PCH, pelvic caliceal height.

图2 随机森林和 XGBoost 模型中的变量重要性评分

Figure 2 Importance score of predictive variables in random forest and XGBoost models



TPR (false positive rate) = Sensitivity; FPR (true positive rate) = 1 - Specificity; ROC, receiver operating characteristic curve.

图3 随机森林和 XGBoost 模型预测结石清除率的 ROC 曲线

Figure 3 ROC for stone clearance prediction of random forest and XGBoost models

表3 随机森林和 XGBoost 模型的预测价值

Table 3 The predictive value of random forest and XGBoost model

Items	Stone-residual	Stone-free	Total
Random forest model, n			
Positive	15	8	23
Negative	10	38	48
Total	25	46	71
XGBoost model, n			
Positive	17	6	23
Negative	8	40	48
Total	25	46	71

既往已有多项研究通过建立不同的评分系统或列线图预测肾结石 RIRS 术后 SFR。Resorlu 等^[5]的研究通过 Logistic 回归模型对多项指标进行分析,发现结石大小、结石数目、肾脏解剖异常及 IPA 是影响 SFR 的独立危险因素,该研究基于上述四个危险因素建立了 RUSS 评分系统,模型 AUROC 为 0.707。此评分系统简单易用,但各项影响因素评分均为 1 分,未考虑不同影响因素对结局影响的权重问题,0~4 分的得分范围不利于其在复杂的临床环境中对 SFR 做出精准评估。

相比于传统的 Logistic 回归分析,随机森林及

XGBoost 算法具有如下优势:对数据分布无线性要求,自动发现并利用相关因素之间的交互效应及非线性关系;可以充分利用有缺失的数据,无须事前对数据进行填补,更加真实地反映数据的原貌;通过大量随机选择样本的方法平衡了样本误差的影响,由此产生的大量数据训练出的模型,较仅以单个测试样本进行拟合的 Logistic 回归模型的结果更为可靠;能够处理大样本数据,并在分析处理数据的过程中不断学习和记忆,不断提高其预测能力。理论和实践均证明,随机森林和 XGBoost 算法是对大多数分类及回归问题都有效的通用模型,具有对异常值、缺失值和噪声容忍度高,不易出现过拟合等优点,在预测分类问题上具有很高的应用价值^[9-10]。本研究尝试将随机森林和 XGBoost 算法应用于肾结石 fURL 术后结石清除结局的预测,取得了很好的预测效果。

随机森林和 XGBoost 是目前医学研究领域最常用的机器学习算法,在疾病风险预测、预后评估及临床决策中发挥着重要作用。理论上,XGBoost 比随机森林具有以下优势:针对性地增加上一轮训练中算法效果欠佳的数据点的损失函数权重,以弥补模型的短板;最后对所有决策树进行集成时以每棵决策树的分类准确性作为权重,突出质量高的决策树对最终结果的影响;增加了复杂度函数和二阶连续可导的目标函数,提升了算法的泛化能力^[21]。本研究运用两种算法分别建立模型,以期得到预测性能更优越的模型,结果显示两种模型的 AUROC 相同。在外部验证中,XGBoost 模型对检测样本的预测灵敏度、特异度及总准确率均略优于随机森林模型,但差异无统计学意义($P > 0.05$)。然而,随着纳入样本量的增加、算法的不断学习及内部参数的不断调整,两种模型的预测性能是否会出现显著差异,还有待进一步研究。

随机森林和 XGBoost 算法在对数据进行分类的同时,还可以给出各个变量的重要性评分,评估各个变量在分类决策中所起的作用。在两种模型中,预测重要性评分排名前四位的变量均为结石体积、平均结石密度、最大结石密度和 BMI,呈现出良好的一致性。结石体积增加会使结石负荷增加,而结石密度增加会导致术中碎石难度增加,这均会导致术后结石残留风险升高。另外,在两种模型中,肾下盏解剖结构相关的四项指标在分类决策中也具有重要权重(重要性评分排名 5~8 位)。已有多项研究证明,若患者存在肾下盏结石,肾下盏解剖结构会对碎石取石效果产生影响。例如,Sari 等^[20]的研究发现

IPA 及 PCH 是影响肾下盏结石清除效果的独立预测因素,狭窄的 IPA 及长的 PCH 会导致结石残留概率增加。结合临床实践我们认为,狭窄的 IPA 及较长的 PCH 会造成术中软镜寻找结石、碎石及取石的难度增加,术中需要不断翻转镜身、调整镜头才能进入肾下盏,狭窄的 IPA 也会造成视野受限及术后结石排出困难。Jessen 等^[22]的多因素研究分析发现,IL 是唯一影响结石排出的肾下盏解剖因素,较长 IL 导致 SFR 升高。另外,我们发现 BMI 在两种模型中的重要性评分较高,既往研究暂未发现 BMI 会影响肾结石患者 fURL 术后 SFR。BMI 作为随机森林和 XGBoost 模型筛选出的高权重变量,目前研究尚无合理解释,这为后续研究提供了方向,但需要注意的是,BMI 的高权重可能是由于数据本身的限制及自变量之间的相关性所导致的^[23]。随着样本量的增加,包括 BMI 在内的各因素在分类决策中的权重是否会发生变化,以及是否与术后结局独立相关,有待进一步研究。

本研究的局限性主要有以下几个方面,首先,模型的建立是基于单中心的回顾性资料;其次,本研究排除了肾结石合并输尿管结石、泌尿系畸形的患者,但临床上这部分患者并不少见,故预测模型的推广应用范围受到一定限制;第三,关于术后何时评估结石清除率,不同研究选择的评估节点不同^[14],本研究选择术后第 1~3 天评估早期结石清除情况,其结果更有利于直接反映结石及肾脏解剖对手术效果的即刻影响,减少了术后其他因素的影响,但随着排石进程,该模型能否准确预测术后 1 个月或更长时间后的 SFR 尚需后续研究;第四,基于研究数据所限及研究目的考虑,我们未通过多因素分析的方法研究各因素与结石清除结局的相关性,这也是我们未来研究的重点;最后,随机森林及 XGBoost 模型的稳定性及可靠性还需要多中心的外部资料进行验证,并在推广过程中检验和提高其预测性能。

综上所述,随机森林和 XGBoost 模型可于术前精确预测肾结石患者 fURL 术后早期 SFR,有利于临床决策及预后评估,但两种模型的预测稳定性及可靠性还需要多中心数据的外部验证,结石体积、结石密度、BMI 及肾下盏解剖因素在两种模型分类决策中具有重要权重,但需继续研究其与结石清除结局是否存在独立相关性。

参考文献

- [1] Zeng G, Mai Z, Xia S, et al. Prevalence of kidney stones in China: an ultrasonography based cross-sectional study [J]. BJU Int,

- 2017, 120(1): 109 - 116.
- [2] Sanguedolce F, Bozzini G, Chew B, et al. The evolving role of retrograde intrarenal surgery in the treatment of urolithiasis [J]. *Eur Urol Focus*, 2017, 3(1): 46 - 55.
- [3] Berardinelli F, Proietti S, Cindolo L, et al. A prospective multi-center European study on flexible ureterorenoscopy for the management of renal stone [J]. *Int Braz J Urol*, 2016, 42(3): 479 - 486.
- [4] 杨波, 胡卫国, 胡浩, 等. 逆行肾内手术治疗肾结石失败的原因分析及其对策 [J]. *北京大学学报(医学版)*, 2014, 46(5): 794 - 797.
- [5] Resorlu B, Unsal A, Gulec H, et al. A new scoring system for predicting stone-free rate after retrograde intrarenal surgery: the "resorlu-unsal stone score" [J]. *Urology*, 2012, 80(3): 512 - 518.
- [6] Jung JW, Lee BK, Park YH, et al. Modified Seoul National University Renal Stone Complexity score for retrograde intrarenal surgery [J]. *Urolithiasis*, 2014, 42(4): 335 - 340.
- [7] Ito H, Sakamaki K, Kawahara T, et al. Development and internal validation of a nomogram for predicting stone-free status after flexible ureteroscopy for renal stones [J]. *BJU Int*, 2015, 115(3): 446 - 451.
- [8] Xiao YL, Li D, Chen L, et al. The R. I. R. S. scoring system: An innovative scoring system for predicting stone-free rate following retrograde intrarenal surgery [J]. *BMC Urol*, 2017, 17(1): 105.
- [9] Breiman L. Random forests [J]. *Mach Learn*, 2001, 45(1): 5 - 32.
- [10] Chen T, Guestrin C. XGboost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining [C]*. San Francisco, California, USA, 2016: 785 - 794.
- [11] Mei X, Wang R, Yang W, et al. Predicting malignancy of pulmonary ground-glass nodules and their invasiveness by random forest [J]. *J Thorac Dis*, 2018, 10(1): 458 - 463.
- [12] Taylor RA, Moore CL, Cheung KH, et al. Predicting urinary tract infection in the emergency department with machine learning [J]. *PLoS One*, 2018, 13(3): e0194085.
- [13] 刘可, 肖春雷, 刘余庆, 等. 标准化技术输尿管软镜钬激光碎石术治疗输尿管近端及肾结石:单中心 140 例报告 [J]. *中国微创外科杂志*, 2015, 15(12): 1065 - 1068, 1087.
- [14] Ghani KR, Wolf JS Jr. What is the stone-free rate following flexible ureteroscopy for kidney stones? [J]. *Nat Rev Urol*, 2015, 12(5): 281 - 288.
- [15] Ito H, Kawahara T, Terao H, et al. The most reliable preoperative assessment of renal stone burden as a predictor of stone-free status after flexible ureteroscopy with holmium laser lithotripsy: a single-center experience [J]. *Urology*, 2012, 80(3): 524 - 528.
- [16] Resorlu B, Oguz U, Resorlu EB, et al. The impact of pelvicaliceal anatomy on the success of retrograde intrarenal surgery in patients with lower pole renal stone [J]. *Urology*, 2012, 79(1): 61 - 66.
- [17] 曹文哲, 应俊, 陈广飞, 等. 基于 Logistic 回归和随机森林算法的 2 型糖尿病并发视网膜病变风险预测及对比研究 [J]. *中国医疗设备*, 2016, 31(3): 33 - 38, 69.
- [18] Kang SK, Cho KS, Kang DH, et al. Systematic review and meta-analysis to compare success rate of retrograde intrarenal surgery versus percutaneous nephrolithotomy for renal stones > 2 cm: An update [J]. *Medicine (Baltimore)*, 2017, 96(49): e9119.
- [19] Inoue T, Murota T, Okada S, et al. Influence of pelvicaliceal anatomy on stone clearance after flexible ureteroscopy and holmium laser lithotripsy for large renal stones [J]. *J Endourol*, 2015, 29(9): 998 - 1005.
- [20] Sari S, Ozok Hu, Topaloglu H, et al. The association of a number of anatomical factors with the success of retrograde intrarenal surgery in lower calyceal stones [J]. *Urol J*, 2017, 14(4): 4008 - 4014.
- [21] 王训师. XGBoost 机器学习模型在缺血性卒中后早期认知损害诊断中的应用研究 [D]. 杭州:浙江大学, 2018.
- [22] Jessen JP, Honeck P, Knoll T, et al. Flexible ureterorenoscopy for lower pole stones: influence of the collecting system's anatomy [J]. *J Endourol*, 2014, 28(2): 146 - 151.
- [23] Nicodemus KK, Malley JD, Strobl C, et al. The behavior of random forest permutation-based variable importance measures under predictor correlation [J]. *BMC Bioinformatics*, 2010, 11: 110.

(2019-03-14 收稿)

(本文编辑:任英慧)