

Multi-modal ultrasound in diagnosis of renal diseases based on support vector machine

LI Guanghan¹, LIU Jian¹, WU Jingping¹, TIAN Yan¹, LIU Jiang²,
MA Liyong³, LIU Yuejun⁴, ZHANG Bo¹, ZHENG Min^{1*}

(1. Department of Ultrasound Medical, China Japan Friendship Hospital, Beijing 100029, China;

2. China Japan Friendship Clinical College of Peking University, Beijing 100029, China;

3. School of Information Science and Engineering, Harbin University of Science and Technology [Weihai], Weihai 264209, China; 4. School of Automation, Harbin University of Science and Technology, Harbin 150080, China)

[Abstract] **Objective** To compare the effectiveness of multi-modal ultrasound, including conventional ultrasound, color Doppler ultrasound and shear wave elastic imaging in diagnosis of renal diseases based on support vector machine support vector machine (SVM) and traditional *Logistic* regression. **Methods** Totally 94 patients with pathologically proved renal diseases (RD group) and 109 patients without renal diseases (control group) were collected and examined with conventional ultrasound, color Doppler ultrasound and shear wave elastic imaging, respectively. SVM and *Logistic* regression were used for modeling. Then all 203 patients were divided into 2 groups according to 3:1, then 153 cases were used as SVM's training samples for single factor variable judgment and model establishment, the other 50 cases were used as validation samples to evaluate the prediction effect of SVM model. **Results** The elastic hardness of left renal cortex and the width of right kidney entered the regression equation in *Logistic* regression. The accuracy of *Logistic* regression model for diagnosis of renal diseases was 83.74%, of SVM model was 85.10% ($\chi^2=0.17$, $P=0.68$). **Conclusion** Multimodal ultrasound has high effectiveness for diagnosis of renal diseases. SVM and *Logistic* models have similar diagnostic effectiveness.

[Keywords] kidney diseases; ultrasonography; elasticity imaging techniques; support vector machine; logistic models

DOI:10.13929/j.issn.1003-3289.2020.06.000

基于支持向量机多模态超声模型诊断肾疾病

李广涵¹, 刘建¹, 武敬平¹, 田艳¹, 刘将², 马立勇³,
刘跃军⁴, 张波¹, 郑敏^{1*}

[1. 中日友好医院超声医学科, 北京 100029; 2. 北京大学中日友好临床学院, 北京 100029;

3. 哈尔滨工业大学(威海)信息科学与工程学院, 山东威海 264209;

4. 哈尔滨理工大学自动化学院, 黑龙江哈尔滨 150080]

[摘要] **目的** 比较基于支持向量机(SVM)和传统 *Logistic* 回归法基于常规超声、彩色多普勒超声和弹性成像参数构建的多模态超声模型诊断肾脏疾病的效能。 **方法** 收集 94 例肾脏疾病患者(肾病组)及无肾脏疾病的对照组患者 109 名, 分别进行常规超声、彩色超声和剪切波弹性检查。采用 *Logistic* 回归法和 SVM 构建模型。利用随机数字法将全部 201 例患者按照 3:1 分为 2 组, 以其中 153 例为训练样本, 进行单因素变量判断和建立 SVM 模型; 以 50 例为验证样本, 评价 SVM 模型的预测效果。 **结果** *Logistic* 回归方程纳入左肾皮质弹性硬度和右肾宽度。 *Logistic* 回归模型预测肾脏疾

[基金项目] 国家政府间国际科技创新合作重点专项(2017YFE0110500)、山东省自然科学基金(ZR2018MF026)。

[第一作者] 李广涵(1988—), 男, 河北秦皇岛人, 在读博士, 医师。研究方向: 超声弹性及造影、介入超声、肌肉骨骼超声。E-mail: ak74ebr@163.com

[通信作者] 郑敏, 中日友好医院超声医学科, 100029。E-mail: zhengmin16@163.com

[收稿日期] 2020-02-15 **[修回日期]** 2020-03-18

病的准确率为 83.74%, SVM 模型为 85.10% ($\chi^2=0.21, P=0.65$)。结论 多模态超声对于肾脏疾病具有较高诊断效能; SVM 和 Logistic 模型的诊断效能相似。

[关键词] 肾疾病; 超声检查; 弹性成像技术; 支持向量机; Logistic 模型

[中图分类号] R692; R445.1 [文献标识码] A [文章编号] 1003-3289(2020)06-0000-00

肾脏疾病发病率较高, 易致终末期肾病^[1], 需早期诊断与治疗, 以延缓肾功能恶化。不同类型肾脏疾病的临床表现及治疗方法存在较大差异^[2]。常规超声、彩色多普勒超声等能观察肾脏的大小及血流等情况, 却难以判断肾脏组织的硬度。弹性成像通过无创方式判断组织硬度, 得到相应部位杨氏模量, 以客观反映其弹性, 其数值随组织硬度的增大而增大^[3], 近年来显示出对于肾脏疾病具有良好应用前景; 但由于肾脏位置较深, 结构复杂, 且测量硬度受血液和尿液的影响较大, 单独应用弹性成像评价肾脏疾病存在一定局限性。本研究观察基于常规超声、彩色多普勒超声和弹性成像的多模态超声诊断肾脏疾病的价值。

1 资料与方法

1.1 一般资料 选取 2019 年 5—12 月在中日友好医院接受超声检查的肾脏疾病患者, 排除无法屏气接受弹性成像、单侧肾脏缺如、肾盂积水或肾脏肿瘤者。以 94 例经肾脏粗针穿刺活检获得明确病理诊断者作为肾病组, 男 60 例, 女 34 例, 年龄 17~70 岁, 平均(42.1±13.1)岁; 包括原发肾小球肾炎者 53 例, 继发肾小球肾炎者 28 例, 小管-间质疾病 13 例。选取同期 109 例肌酐正常且排除可能影响肾脏的疾病的住院或门诊患者作为对照组, 男 59 名, 女 50 名, 年龄 16~79 岁, 平均(52.9±15.0)岁。患者均签署知情同意书。

1.2 仪器与方法 按照预先设计的调查问卷记录 2 组患者一般情况后, 采用迈瑞昆仑 R8 超声仪进行常规超声、彩色多普勒超声及弹性成像。嘱患者仰卧或侧卧, 先以常规超声分别测量双肾长度、宽度、皮质厚度、髓质厚度、实质厚度及距皮深度。随后行彩色多普勒检查, 于肾髓质间测量叶间动脉, 于肾动脉起始段或肾门处测量肾动脉主干, 分别测量上述血管的收缩及舒张期血流流速及血流阻力指数。最后进入剪切波弹性模式, 使探头尽可能与肾脏包膜垂直, 适当经探头施压

后行弹性检查, 待弹性图像稳定后嘱患者屏息, 尽量调小圆形 ROI 的面积, 确保其内彩色充满且均匀一致, 并在 5~10 s 内完成测量。每次测量后暂停弹性系统, 重新开启后再行下一次测量。依次测量右肾及左肾皮质、髓质和肾窦部剪切波速度, 剪切波速度最少测量 3 次, 取中位数为剪切波数值(图 1~3)。由 2 名具有 5 年以上超声腹部诊断经验和弹性成像操作经验的超声科主治医师完成上述操作。

1.2.1 数据处理与统计 使用 Microsoft Excel 2007 软件录入临床数据。按照 3:1 比例将全部 203 例患者随机分为 2 组, 以其中 153 例作为训练样本、50 例为验证样本。采用 SPSS 25.0 统计分析软件, 以独立样本 *t* 检验比较单因素连续变量, 筛选与肾病存在统计学关联的指标建模。以肾病组=1 和对照组=0 为因变量, 将筛选出的单因素变量纳入多因素 Logistic 回归, 变量纳入的标准设置为 0.05。采用 Logistic 混合逐步前向回归(LR 法)建立 Logistic 回归预测模型, 得到存在肾病与否的预测结果。P<0.05 为差异有统计学意义。

1.2.2 SVM 建模 采用 Logistic 回顾和 SVM 法, SVM 采用二类支持向量机, 其核函数包括线性核函数、高斯核函数和径向基函数, 通过优化方法对核函数及其参数获得最佳值, 从而建立诊断模型。于 Python 3.7 平台运行 SVM 模型, 以 scikit-learn 0.22 软件进

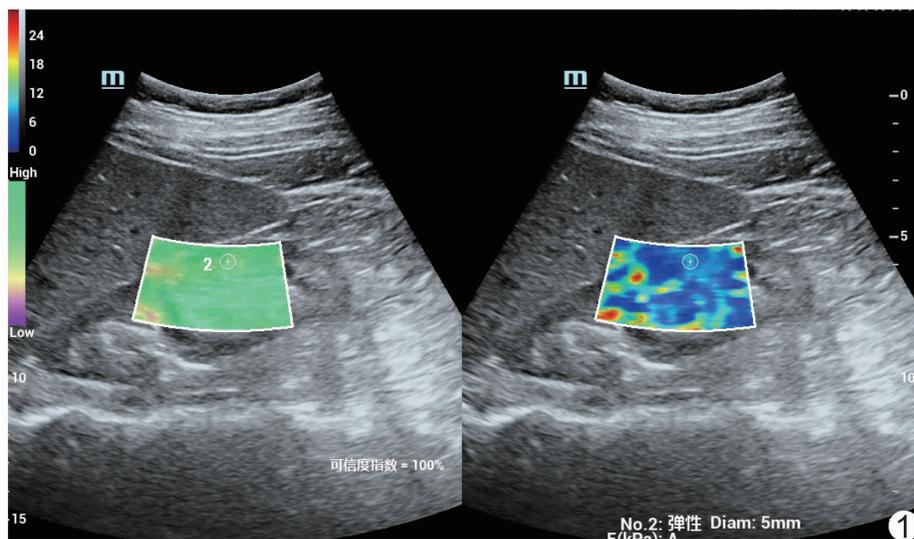


图 1 正常肾脏弹性成像图, 肾皮质弹性值 8.74 kPa

表 1 单因素变量检验结果

组别	右肾宽度 (cm)	右肾实质厚度 (cm)	右肾皮质厚度 (cm)	左肾实质厚度 (cm)	右肾皮质弹性 (kPa)	左肾皮质弹性 (kPa)
肾病组	4.87±0.60	1.87±0.30	0.91±0.17	1.91±0.30	6.37±1.41	6.32±1.61
对照组	4.41±0.75	1.73±0.29	0.86±0.18	1.79±0.27	9.07±2.36	9.14±1.93
t 值	-4.83	-3.38	-1.97	-3.06	10.03	11.18
P 值	<0.01	0.01	0.05	<0.01	<0.01	<0.01

表 2 多因素逐步 Logistic 回归结果

因素	B	标准差	Wald 值	P 值	Exp(B)
右肾实质厚度(cm)	2.090	0.79	6.99	0.01	8.08
右肾皮质弹性(kPa)	-0.611	0.14	18.76	<0.01	0.54
左肾皮质弹性(kPa)	-0.735	0.14	26.01	<0.01	0.48
常量	6.345	1.83	12.09	<0.01	569.38

行分析。采用网格优化方法搜索 SVM 算法的核函数和各参数,逐步得到准确率最高的模型架构,选择准确率最高的核函数和参数作为预测模型。以敏感度、特异度、准确率及 AUC 作为评价诊断效能的指标。

2 结果

2.1 单因素分析 6 个单因素变量与肾病存在统计学关联 ($P < 0.05$),分别为右肾宽度、右肾实质厚度、右肾皮质厚度、左肾实质厚度、左肾皮质弹性硬度和右肾皮质弹性硬度。其他观察指标与肾病无明显统计学关联。见表 1。

2.2 建立 Logistic 多因素回归模型 多因素回归预测模型中包括 3 个因素,即右肾实质厚度、右肾皮质弹性硬度和左肾皮质弹性硬度。Logit(p) = 2.090 × 右肾实质厚度 - 0.611 × 右肾皮质弹性硬度 - 0.735 × 左肾皮质弹性硬度 + 6.345。其中 P 为归于肾病组的概率,取 0.5 为判断阈值,即 $P > 0.5$ 为肾病组, $P < 0.5$ 归为对照组。见表 2。

2.3 模型诊断效能比较 对建立的 Logistic 回归模型进行测试,用敏感度、特异度、准确率和 AUC 评价诊断效能, SVM 模型准确率 85.10%,与 Logistic 回归模型 (83.74%) 接近 ($\chi^2 = 0.21, P = 0.65$)。见表 3。

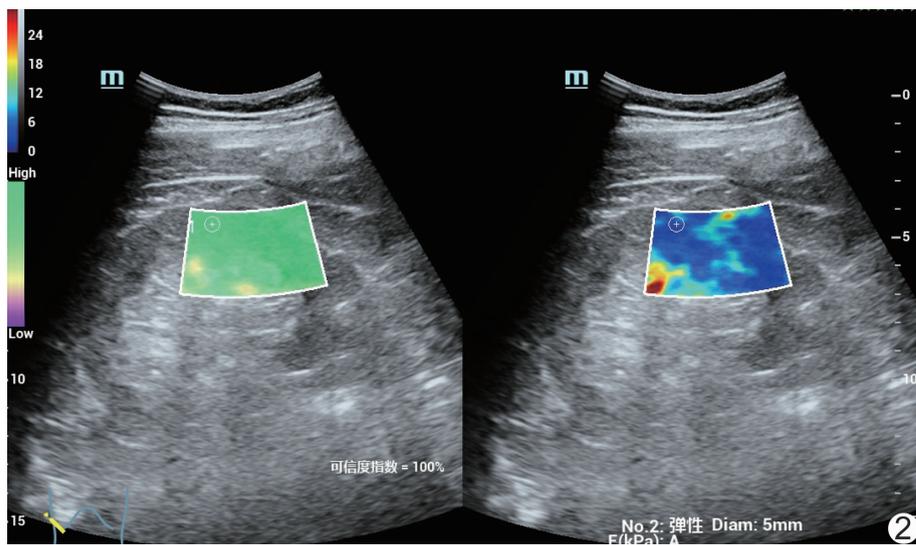


图 2 男, 52 岁, 不典型膜性肾病, 弹性成像图, 肾皮质弹性值 4.96 kPa

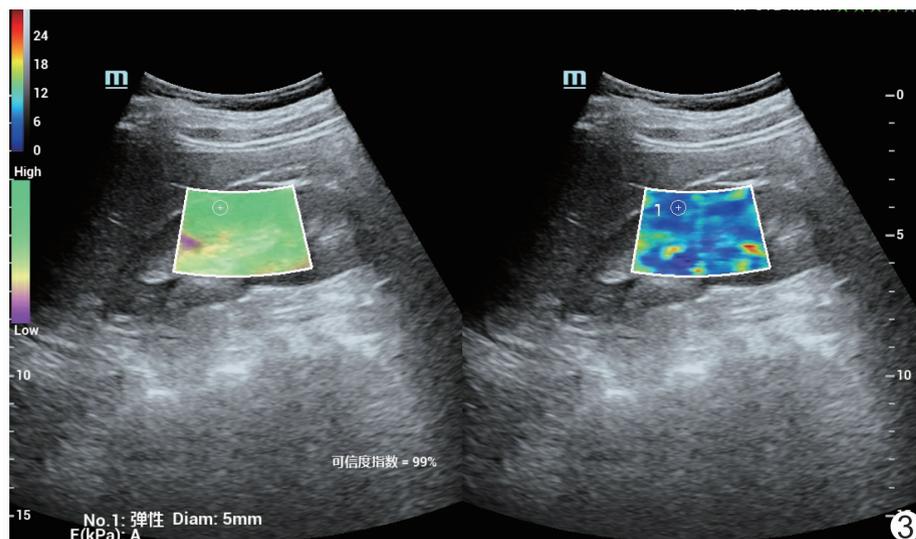


图 3 男, 31 岁, IgA 肾病伴缺血性肾损伤, 弹性成像图, 肾皮质弹性 5.77 kPa

表 3 SVM 与 Logistic 回归模型预测效果比较

预测模型	敏感度(%)	特异度(%)	准确率(%)	AUC
SVM	83.23	86.97	85.10	0.852
Logistic 回归	80.85	86.23	83.74	0.835
χ^2 值	0.55	0.09	0.21	—
P 值	0.46	0.76	0.65	—

3 讨论

支持向量机(support vector machine, SVM)是一种基于结构最小化理论的机器学习方法,适用于处理样本量少的高维度的线性及非线性医学相关数据^[4]。本研究应用 SVM 和 Logistic 回归构建模型判断不同肾脏疾病,观察联合应用常规超声、彩色多普勒超声和弹性成像诊断肾脏病变的效能,比较不同模型诊断能力的差异。

肾脏具有维持机体水和电解质稳定的重要作用,也是多种疾病的靶器官。不同类型肾损伤可导致肾小球硬化、肾小管萎缩及肾间质病变等多种不同组织改变。肾脏代偿能力较强,常规血尿检验提示异常时,通常肾脏损伤已比较严重^[5],而此时常规超声和彩色多普勒超声可无特异性发现,仅能对肾脏大小和血流改变进行观测,使得无创且能定量评估肾脏疾病的影像学检查手段成为临床迫切需要。实时剪切波弹性成像是近年发展起来的新型超声检查技术^[6],通过向目标组织施加激励使之产生相应组织位移,以超声结合数字信号后处理技术评价目标组织的弹性硬度,能无创、定量地反映组织弹性,现已逐步用于临床诊断肾脏及肾脏相关疾病^[7-8]。目前相关研究多单独分析剪切波弹性成像与肾损伤的关系,联合应用常规超声、彩色超声和弹性成像诊断肾脏疾病者较少,且缺乏联合诊断建立肾损伤的模型观察。

本研究采用传统统计学方法为 Logistic 回归建模,需要较多样本,对于分析小样本高维度数据存在一定缺陷,故引入 SVM 思想建立诊断模型。首先将原始样本数据映射到高维特征空间,然后识别一个超平面,将数据项分为 2 类;该超平面通过最大化两类的距离和最小化分类错误得到,具有较好的泛化能力^[9]。既往研究^[10-11]已证明 SVM 是解决医疗问题的有效方法。通过 10 次随机分组方式,本研究建立的 SVM 模型计算 10 次检验结果平均的准确率,可最大程度地减少抽样不均匀对训练样本和测试样本造成的诊断结果的随机误差,从而更真实有效地反映 SVM 模型对肾脏疾病的诊断效能。

单因素分析结果显示,常规二维超声所示右肾宽

度、右肾实质厚度、右肾皮质厚度及左肾实质厚度在肾病组与对照组之间的差异具有统计学意义,且肾病组大于对照组。本研究肾病组患者均接受肾穿刺活检,其中慢性肾病急性发作及急性肾病占比相对较高,急性肾病时肾脏存在肿胀和充血等改变,导致肾脏不同程度增大,故部分肾脏常规超声测值增大。弹性成像测值中,双侧肾皮质硬度组间差异有统计学意义,且肾病组低于对照组,可能与肾脏疾病患者肾血流减少,肾脏周围纤维被膜形成,弹性受血液压力影响大有关,与既往研究^[12]相符。本研究 Logistic 回归模型和 SVM 模型均将超声弹性列为最终计算参数,提示剪切波弹性数值对于判断肾病具有重要参考价值;但肾脏彩色多普勒血流参数在肾病组与对照组之间差异无明显统计学意义,可能与肾病组仅包含 1 例高血压肾病有关。本研究中 Logistic 回归模型的预测准确率为 83.74%,SVM 模型为 85.10%,提示 2 种模型均具较好诊断效能,而 SVM 模型的可拓展性更高;Logistic 模型主要在数据层面建立,而 SVM 可进一步分析包括医疗图像、文字描述等数据在内的更多医疗信息,提高对肾脏疾病的判断能力。

综上所述,多模态超声是无创诊断肾脏疾病的较好手段;SVM 模型和 Logistic 模型均有较高诊断效能,尚需要更多数据加以进一步优化。

[参考文献]

- [1] 冯杰,刘茂东.特发性膜性肾病治疗时机与进展[J].临床荟萃,2019,34(3):284-288.
- [2] 喻敏,张伟文,刘春雅,等.431 例肾活检病理类型及流行病学特点[J].中国现代医生,2015,53(35):24-27.
- [3] 高峰,蒋苏齐.实时剪切波弹性成像技术在儿童慢性肾病中的应用价值及临床相关性研究[J].中国临床医学影像杂志,2019,30(8):574-577.
- [4] KOUROU K, EXARCHOS T P, EXARCHOS K P, et al. Machine learning applications in cancer prognosis and prediction [J]. Computat Struct Biotechnol J, 2015,13:8-17.
- [5] 杨茜,谢红浪.急性肾损伤进展为慢性肾脏病的机制[J].肾脏病与透析肾移植杂志,2016,25(3):274-278.
- [6] 秦敏.肝脏超声实时剪切波弹性在早期诊断胆道闭锁中的价值探讨[J].肝脏,2019,24(7):815-816.
- [7] GOERTZ R S, AMANN K, HEIDE R, et al. An abdominal and thyroid status with acoustic radiation force impulse elastometry-a feasibility study: Acoustic radiation force impulse elastometry of human organs[J]. Eur J Radiol, 2011,80(3):226-230.
- [8] 赵一冰,陈涛,郭稳,等.应用实时剪切波弹性成像技术评价三种良性软组织肿瘤硬度的初步研究[J].中国超声医学杂志,2019,

35(7):643-646.

[9] MATIVO J M, HUANG S. Prediction of students' academic performance: Adapt a methodology of predictive modeling for a small sample size[C]. Proc of 44th Annual Frontiers in Education Conference, Madrid: IEEE, 2014:1-3.

[10] 饶飘雪,叶枫. 基于 Logistic 回归、ANN、SVM 的乳腺癌复发影响因素研究[J]. 计算机系统应用, 2016, 25(7):259-263.

[11] 杨志燕,黄天宝,王树山,等. 基于支持向量机的急性百草枯中毒预后模型的建立与评价[J]. 南京医科大学学报(自然科学版), 2018, 38(10):143-147.

[12] 李立红,朱立松,武敬平,等. 声辐射力脉冲弹性成像技术在老年急慢性肾病患者中的应用价值[J]. 中华医学超声杂志(电子版), 2017, 14(6):462-467.