

# 基于机器学习算法评估心脏超声对医护人员队列中慢性疲劳综合征诊断价值的初探

## Preliminary study on diagnostic value of cardiac ultrasound on chronic fatigue syndrome in medical staff queue based on machine learning algorithms

朱鸿<sup>1</sup>, 方彦鹏<sup>1</sup>, 刘洋<sup>1</sup>, 陈驰<sup>2</sup>

(1. 贵州中医药大学第二附属医院超声科, 贵州 贵阳 550003; 2. 贵州中医药大学, 贵州 贵阳 550002)

**摘要:** 对我院2016年1月至2018年1月127例健康体检医护人员通过双盲法进行慢性疲劳综合征(chronic fatigue syndrome, CFS)相关问卷采集并收集体检及心脏超声数据, 建立回归模型, 采用回归模型或交叉验证进行内部验证, 并根据机器学习XGBoost提供的诊断模型增益值评估心脏超声指标在CFS诊断中的价值。结果显示, 127例医务人员体检数据进行机器学习预计算, 获得的人工智能模型受试者95%CI为0.881 6 (0.864 1, 0.912 7), 工作特征曲线(AUC)为0.902 1; 非-CFS与CFS病人敏感性(0.618 9、0.892 1)、准确性(0.892 1、0.618 9)、阳性预测值(0.924 2、0.835 6)、阴性预测值(0.835 6、0.924 2)、流行率(0.856 4、0.143 6)、检测率(0.841 8、0.074 5)、检测流行率(0.910 8、0.089 2)。提示基于心脏超声等客观体检指标的快速诊断模型有助于在临床诊疗及体检过程中优化检查项目、快速区分CFS人群, 更好地进行健康指导及后续医疗干预。

**关键词:** 机器学习算法; 心脏超声; 慢性疲劳综合征; 诊断价值

**中图分类号:** R311; R747.9 **文献标识码:** B

**文章编号:** 1002-221X(2019)03-0185-02

**DOI:** 10.13631/j.cnki.zggyyx.2019.03.006

慢性疲劳综合征(chronic fatigue syndrome, CFS)于20世纪90年代中期由日本学者Fukuda首次提出<sup>[1]</sup>。鉴于CFS的病因、危险因素和病理生理学仍然未知, 因此CFS的诊断、治疗和预防策略制定存在重大缺陷。有研究表明CFS与心血管系统功能的降低密切相关<sup>[2-4]</sup>; 有学者通过心脏磁共振标记的方法, 进一步证实CFS患者的心脏容积、心输出量明显降低<sup>[5]</sup>。由此可见, CFS患者的心功能可出现明显异常, 甚至在CFS早期即可出现。eXtreme Gradient Boosting(XGBoost)是一种专为诊断而设计的机器学习算法, 它被认为是梯度推进决策树(GBDT)的实现<sup>[6]</sup>。XGBoost可构建一系列的弱分类器, 并进行迭代组合, 目的是将这些单独预诊断效果很弱的分类器组合成为稳健的估计器。因此, 本研究通过回顾收集127例健康体检者基础检查与心脏超声结果, 利用XGBoost分析初步建立诊断模型, 现将结果报告如下。

### 1 对象与方法

**收稿日期:** 2019-03-28; **修回日期:** 2019-04-30

**作者简介:** 朱鸿(1968—), 女, 主任医师, 主要研究方向: 心血管、妇产科超声。

**通信作者:** 陈驰, 讲师, E-mail: 2303870722@qq.com。

### 1.1 对象

通过我院2016年1月至2018年1月医护人员健康体检数据库, 收集127例心脏超声、血清学、影像学及其他体检项目指标用于构建模型, 变量来源包括(1)一般情况: 身高、体重、既往病史询问、生活习惯, 问卷填写;(2)血清学资料: 血脂代谢、甲状腺功能指标、血常规、尿常规、便常规、肝肾功能、尿酸等生化指标;(3)影像学指标: 腹部彩超、颈部血管彩超、甲状腺彩超、心电图等, 包含数字信息及非数字信息, 非数字信息后期通过数据提取变为二分类哑变量或多分变量进入数据库中存储;(4)心脏超声: 右室流出道内径(RVOT)、升主动脉内径(AAO)、左房内径(LA)、室间隔厚度(IVSd)、左室后壁厚度(LVPWd)、左室舒张期内径(LVed)、左室收缩期内径(LVes)、右房内径(RA)、左室流出道内径(LVOT)、肺动脉内径(PA)、心率(HR)、左室舒张早期血流峰值(E)、左室舒张晚期血流峰值(A)、E/A比值、射血分数(EF)、左室短轴缩短率(FS)、心排出量(CO)、收缩末期容积(ESV)、舒张末期容积(EDV)。

### 1.2 方法

参照2005年美国CDC发布的CFS诊断标准。收集127例医务人员的体检数据, 使用双盲法对医务人员进行CFS相关问卷采集、数据录入, 并存放至临床试验电子数据采集管理系统(EDC)。通过XGBoost建立CFS快速诊断模型, 分析工作特征曲线(AUC)、敏感性和特异性。根据XGBoost提供的诊断模型增益值评估心脏超声指标在CFS的诊断价值, 并采用 $\alpha$ 信度系数对量表中各选项及总量表的信度进行检测。

### 1.3 统计分析

数据录入易佰dataweb管理软件, 使用Hosmer-Lemeshow(H-L)测试和校准图评估XGBoost模型的诊断性能。整个机器学习过程由XGBoost R软件包执行(版本: 0.6.4.1.tar.gz, <https://cran.r-project.org/web/packages/xgboost/index.html>); 采用Kaiser-Meyer-Olkin检验对效度进行分析。

## 2 结果

### 2.1 体检数据的准确率与AUC模型分辨率

对127例医务人员体检数据进行机器学习预计算, 获得的人工智能模型受试者95%CI为0.881 6 (0.864 1, 0.912 7), 工作特征曲线(AUC)为0.902 1。

### 2.2 非-CFS与CFS的敏感性、特异性比较

通过计算KMO值为0.748; 采用易佰软件进一步分析,

非-CFS与CFS病人敏感性(0.618 9、0.892 1)、准确性(0.892 1、0.618 9)、阳性预测值(0.924 2、0.835 6)、阴性预测值(0.835 6、0.924 2)、流行率(0.856 4、0.143 6)、检测率(0.841 8、0.074 5)、流行检测率(0.910 8、0.089 2)。

### 2.3 暴露变量特征分析

各指标 $\alpha$ 信度系数均 $>0.7$ ,采用Kaiser-Meyer-Olkin检验对效度进行分析,根据XGBoost增益值进行评估。见表1。模型受试者非-CFS与CFS病人敏感性、准确性、流行率、检测率、检测流行率等有明显的差异。

表1 暴露变量特征分析

指标	Gain	Cover	频率
性别	0.328 316 564	0.138 558 438 8	0.102 955 195
年龄	0.116 108 984	0.083 805 312 1	0.075 309 819
身高	0.111 655 113	0.063 686 436 8	0.055 290 753
体重指数	0.068 717 540	0.062 681 855 2	0.059 103 908
左心室壁厚度	0.039 092 945	0.039 606 445 8	0.042 897 998
右心室壁厚度	0.030 224 814	0.045 558 859 3	0.045 757 865
EV	0.029 092 525	0.053 155 816 3	0.055 290 753
RA	0.026 971 489	0.050 022 399 0	0.054 337 464
ESV	0.024 942 311	0.039 252 085 0	0.040 038 132
AAO	0.023 903 343	0.029 292 591 4	0.048 617 731
eGFR	0.022 035 051	0.049 761 633 4	0.041 944 709
HR	0.021 739 969	0.043 481 180 1	0.046 711 153
HDL-C	0.021 535 929	0.042 195 575 0	0.040 991 420
TG	0.021 271 439	0.058 958 013 0	0.049 571 020
UA	0.020 694 528	0.049 765 355 3	0.046 711 153

### 3 讨论

CFS是指至少6个月的持续性或复发性疲劳,休息时不会减轻,活动大量减少,疲劳不能通过医疗或精神疾病来解释,并且必须伴有不常见的劳累后疲劳、记忆力或注意力不集中、睡眠不足、头痛、肌肉疼痛、关节疼痛、喉咙痛和宫颈压痛典型症状中的至少4种<sup>[7]</sup>。不同的国家和研究机构对CFS的诊断制定了不同的标准,比较公认的标准有加拿大临床共识标准(CCC)和美国国家医学研究所(IOM)制定的SEID指南。这些标准均有其优缺点,CCC诊断标准主要通过7种症状,包括运动后不适和神经认知障碍等对CFS进行临床诊断<sup>[8,9]</sup>,涵盖的症状更多,但在提高特异性(降低误诊率)的同时增加了漏诊率,不利于临床筛查使用<sup>[10,11]</sup>。我国目前尚无针对中国人群的CFS诊治指南,且现有诊断主要基于主观问卷调查,大多为患者的主观感受,较少有客观实验室指标,诊断的准确性易受人为因素干扰。CSF和非-CSF判定并无金标准,这是目前的相关临床现状决定的,并非主观研究设计失误,也是今后研究需要弥补的不足。我们本次采用Kaiser-Meyer-Olkin检验对效度进行分析,KMO值为0.748( $>0.7$ ),说明问卷的结构效度较好。

机器学习和医学领域的结合在疾病的诊断、预后中具有较高的准确性。利用机器学习找到收集的临床、生化等各个指标之间的相关性,可以提高诊断的准确性。CFS与心血管系

统疾病的研究相对较少,Hollingsworth等通过对128例患者心脏超声检查发现,CFS患者的左室射血分数、舒张末期容积和心输出量均明显降低,且这种心功能的降低与CFS的病情严重程度呈正相关<sup>[2]</sup>。

本研究以临床医护人员作为对象并建立相关队列,通过常见的临床体检客观指标,应用机器学习算法构建诊断模型,筛查试验强调的是低漏诊率,确诊试验强调的是低误诊率。从结果可见模型受试者95%CI为0.881 6(0.864 1,0.912 7)、工作特征曲线(AUC)为0.902 1,非-CFS与CFS病人敏感性、准确性、流行率、检测率、检测流行率均有明显的差异,符合实际临床应用场景。若进一步加大样本量,模型的精准性将会越来越高。

### 参考文献:

- [1] Fukuda K, Straus SE, Hickie I, *et al.* The chronic fatigue syndrome: A comprehensive approach to its definition and study. International chronic fatigue syndrome study group [J]. *Ann Intern Med*, 1994, 121 (12): 953-959.
- [2] Hollingsworth KG, Jones DE, Taylor R, *et al.* Impaired cardiovascular response to standing in chronic fatigue syndrome [J]. *Eur J Clin Invest*, 2010, 40 (7): 608-615.
- [3] Miwa K, Fujita M. Cardiac function fluctuates during exacerbation and remission in young adults with chronic fatigue syndrome and "small heart" [J]. *J Cardiol*, 2009, 54 (1): 29-35.
- [4] Miwa K, Fujita M. Cardiovascular dysfunction with low cardiac output due to a small heart in patients with chronic fatigue syndrome [J]. *Intern Med*, 2009, 48 (21): 1849-1854.
- [5] Hollingsworth KG, Hodgson T, Macgowan GA, *et al.* Impaired cardiac function in chronic fatigue syndrome measured using magnetic resonance cardiac tagging [J]. *J Intern Med*, 2012, 271 (3): 264-270.
- [6] Alevizos M, Karagkouni A, Panagiotidou S, *et al.* Stress triggers coronary mast cells leading to cardiac events [J]. *Ann Allergy Asthma Immunol*, 2014, 112 (4): 309-316.
- [7] Fenouillet E, Vigouroux A, Steinberg JG, *et al.* Association of biomarkers with health-related quality of life and history of stressors in myalgic encephalomyelitis/chronic fatigue syndrome patients [J]. *J Transl Med*, 2016, 14: 251.
- [8] Asprusten TT, Fagermoen E, Sulheim D, *et al.* Study findings challenge the content validity of the Canadian Consensus Criteria for adolescent chronic fatigue syndrome [J]. *Acta Paediatr*, 2015, 104 (5): 498-503.
- [9] Jason LA, Sunnquist M, Brown A, *et al.* Chronic fatigue syndrome versus systemic exertion intolerance disease [J]. *Fatigue*, 2015, 3 (3): 127-141.
- [10] Morris G, Maes M. Case definitions and diagnostic criteria for myalgic encephalomyelitis and chronic fatigue syndrome: From clinical-consensus to evidence-based case definitions [J]. *Neuro Endocrinol Lett*, 2013, 34 (3): 85-99.
- [11] Jason LA, Brown A, Evans M, *et al.* Contrasting chronic fatigue syndrome versus myalgic encephalomyelitis/chronic fatigue syndrome [J]. *Fatigue*, 2013, 1 (3): 168-183.