

·临床研究·

基于声学分析和机器学习构建咳嗽和清嗓分类模型*

许 维¹ 吕倩倩² 江钟立^{1,2} 林 枫^{1,2,3}

摘要

目的:建立气道廓清动作的声学分类器,为实施肺康复的咳嗽训练提供监测工具。

方法:健康男性11例和女性15例,分别在平卧位、45°靠坐位和90°端坐位,根据随机视觉指令执行咳嗽和清嗓动作各10次,并同时录制声音,分析声音片段的时域、频域和信息域特征,由此构建声音的特征矢量用于机器学习。采用的模型包括:线性判别分析、分类回归决策树、随机森林和线性分类器型支持向量机。

结果:模型间比较显示随机森林方法所建分类器具有更高的准确度(0.9162)和一致性(Kappa值为0.8323)。验证结果显示该模型无论在区分体位因素或不区分体位情况下,对咳嗽音有较高的准确度、一致性、敏感度和特异度。

结论:咳嗽和清嗓动作具有声学差异,并且这种差异可以由随机森林方法构建机器学习模型加以分类,由此为肺康复治疗中采用声学手段辅助判断气道廓清动作类型提供了依据。

关键词 咳嗽音;清嗓音;机器学习;声学分析

中图分类号:R493.042 文献标识码:A 文章编号:1001-1242(2020)-12-1434-05

Classifying cough and throat clearing behaviors based on acoustic analysis and machine learning/XU Wei, LYU Qianqian, JIANG Zhongli, et al./Chinese Journal of Rehabilitation Medicine, 2020, 35(12): 1434—1438

Abstract

Objective: To establish an acoustic classifier of airway clearance related behaviors and to provide a monitoring tool for cough training in pulmonary rehabilitation.

Method: Eleven healthy males and fifteen females were seated in flat, 45-degree and 90-degree positions, respectively. Cough and throat clearing behaviors were performed 10 times each according to the random visual instruction, and the sound was recorded at the same time. For each sound segment, acoustic features of time-domain, frequency-domain and information-domain were analyzed, then extracted the feature vectors of sound for machine learning. The models contain linear discriminant analysis, classification regression decision tree, random forest and linear classifier type support vector machine.

Result: The comparison between the models showed that the classifiers established by random forest method had higher accuracy (0.9162) and the Kappa value (0.8323). Whether or not the influences of postural factors are included, the verification results showed that random forest model had the highest accuracy, the Kappa value, sensitivity and specificity for classification.

Conclusion: There are acoustic differences between cough and throat clearing, and the differences can be classified by machine learning model. Our study provides a promising strategy for developing airway clearance behavior related classifier based on acoustic information.

Author's address Sir Run Run Hospital, Nanjing Medical University, Jiangsu, Nanjing, 211100

Key word cough sound; throat clearing sound; machine learning; acoustic analysis

DOI:10.3969/j.issn.1001-1242.2020.12.005

*基金项目:南京医科大学科技发展基金(NMUB2018295);江苏省高校哲学社会科学优秀创新团队建设项目(2017STD006)

1 南京医科大学附属逸夫医院,江苏省南京市,211100; 2 南京医科大学第一附属医院; 3 通讯作者

第一作者简介:许维,女,主管治疗师; 收稿日期:2019-07-29

肺部疾病或因疾病长时间卧床患者,常继发呼吸功能不全并伴有气道廓清障碍,通常会出现呼吸困难、疲劳和运动不耐受等致残症状。此外,患者常有肺外表现,如骨骼肌功能障碍,以及焦虑、抑郁等心理障碍。较低的身体活动水平,导致患者生活质量明显受损,给患者造成巨大的负担^[1]。因此,针对气道廓清障碍的肺康复治疗,逐渐受到重视。

气道廓清技术(airway clearance therapy, ACT)是肺康复中的重要内容,是利用呼吸训练、体位引流等物理方式或机械振动方式影响气道壁本身和气道内的气流,使附着于气道壁的痰液松动,并通过咳嗽动作将痰液排出体外^[2]。但是,目前在气道廓清训练过程中,主要依靠治疗师人为主观判断患者的呼吸和咳嗽行为,缺少客观的人体活动检测指标以及记录手段。对每一次主动发起的咳嗽是否起到了有效的气道廓清效应,目前缺少能够实时监测、信息反馈、适应性调整和持续记录结果的方法。此外,患者并不总能清晰地区分咳嗽和清嗓这两种动作^[3],从而在咳嗽训练中夹杂了清嗓动作。因此,虽然气道廓清训练已经是临床常用肺康复技术,但是整个训练过程缺少可持续记录的客观依据,影响了该训练技术的质量控制。

肺部听诊是一种成熟的查体技术。有经验的临床医师可以通过听诊辨别咳嗽音和清嗓音。这种经验是通过反复的临床实践在检查者脑中形成的辨识机制。本研究旨在模仿这一过程,记录并标注咳嗽训练过程中产生的咳嗽音和清嗓音,形成可供计算机深度学习的数据集,最终建立实时记录和区分咳嗽和清嗓动作的模型。

1 资料与方法

1.1 一般资料

健康志愿者26例(11例男性,15例女性,年龄19—29岁)。纳入标准:①肺功能评估,激励式肺量计吸气量>2500ml者;②无慢性肺部疾病史者;③近半年来无呼吸道感染病史者;④无吸烟史者。排除标准:①心肺疾病,不能耐受反复的咳嗽和清嗓动作者;②认知功能障碍,不能理解并正确执行呼吸训练者;③有精神障碍,不能配合者;④其他经临床判断不能接受呼吸和咳嗽训练者,如肋骨骨折等;⑤不能

签署知情同意书者。本试验已通过南京医科大学第一附属医院和附属逸夫医院伦理委员会批准(2018-SR-098, 2019-SR-005)。

1.2 研究方法

对入组的志愿者,首先进行认知、年龄、BMI、胸围、颈围和腰围等评估。试验中,将微型声音采集器以弹力胶带贴附于受试者胸骨柄位置。在安静环境下,采集健康志愿者平卧位自然呼吸音5min,缓慢深吸气和呼气各10次,每次间隔20s。然后,志愿者分别在仰卧位、45°半卧位、90°端坐位,根据随机视觉指令执行10次咳嗽和10次清嗓动作。采用Raiselong Health Acoustics(V2.5.0.4)声学平台制备和管理声学样本库。声音录制探头为Raiselong—HA1,以16bit的“.wav”格式保存。

1.3 数据处理

采用RavenPro(V1.5)对声音片段进行标注和声学特征提取。使用R软件(V3.5.1)分析声音片段的时域、频域和信息域特征,从而建立携带声学矢量的样本库。对录制的咳嗽音和清嗓音进行数据清洗。为了均衡两种任务的次数,采用R软件的“ROSE”软件包实施重采样和平滑自助。然后,从已制备的平衡数据集中,随机选择80%作为训练集,20%作为测试集,用机器学习算法构建分类器,方法为十折交叉验证。采用的模型包括:线性判别分析(linear discriminant analysis, LDA)、分类回归决策树(classification and regression tree, CART)、随机森林(random forest, RF)和线性分类器型支持向量机(support vector machine, SVM)。采用准确度(accuracy)和Kappa值进行模型间比较。前者表示的是分类正确的样本数占总样本数的比例,其计算公式为:(真阳性率+真阴性率)/(真阳性率+假阳性率+真阴性率+假阴性率);后者度量的是同一模型在多次分类结果之间的一致性。Kappa值的取值范围在[-1, 1],数值越大说明分类的一致性越好^[4-5]。此外,采用敏感度(即真阳性率)和特异度(即真阴性率)对模型进行分析^[6]。声学特征在不同体位间或不同任务间的比较,采用随机区组方差分析。正态分布采用Shapiro-Wilk W检验,非正态分布经Box-Cox变换后进入方差分析。事后检验采用TukeyHSD,以 $\alpha=0.05$ 为显著性界值。

2 结果

研究对象平均年龄为22岁,颈围、胸围、腰围分别为 $33.58\pm 2.57\text{cm}$ 、 $88.23\pm 7.15\text{cm}$ 和 $77.92\pm 9.95\text{cm}$,体质指数(BMI) $22.42\pm 2.58\text{kg/m}^2$,激励式肺量计吸气量为 $3417.31\pm 447.42\text{ml}$ 。另测得蒙特利尔认知评估量表(MOCA,正常值 ≥ 26 分)和简易精神状态评价量表(MMSE,正常值 ≥ 27 分)分别为 29.12 ± 0.99 分和 29.92 ± 0.27 分。

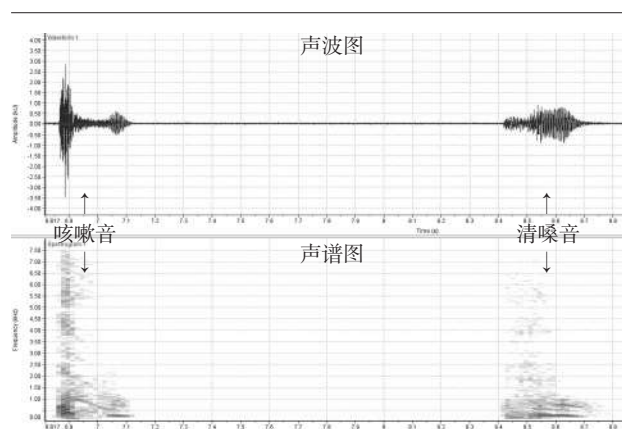
经切分和人工检查,弃去受干扰(如受吞咽动作干扰)或不慎未能按指令执行(如在要求咳嗽时做了清嗓并纠正的咳嗽动作)的片段,最终得到732次咳嗽音和843次清嗓音。见图1。

选取表1中的8个参数构建特征矢量。以下分析为得到同时满足两个条件的指标集,即:①指标个数尽可能少;②尽可能准确地预测是咳嗽还是清

嗓。变量选择方法使用全子集回归,即所有可能的模型都会被检验^[7]。首先,把8个量化指标全部作为自变量,以咳嗽和清嗓任务作为应变量,进行全子集回归,从而筛选进入到后续机器学习模型的指标。此处采用R软件的“leaps”包计算,子集量设置为4个。全子集回归图(图2)显示,在adj-R²最大值0.58时,达峰时间、总能耗、峰值功率、强度达峰时的最高频率和聚合熵构成了最少且最能解释咳嗽和清嗓的5个指标。因此,选取上述指标建立模型。模型间比较显示(表2)随机森林方法所建分类器具有更高的准确度(0.9162)和一致性(Kappa值为0.8323)。

用20%测试集进一步对随机森林模型进行分析。在不区分体位情况下,准确度为0.9026,一致性(Kappa值)为0.8049,敏感度为0.8800,特异度为0.9241。对平卧位数据验证,准确度为0.9192,一致性为0.8385,敏感度为0.8868,特异度为0.9565。半卧位数据验证,准确度为0.8614,一致性为0.7147,敏感度为0.8333,特异度为0.8814。端坐位数据验证,准确度为0.9259,一致性为0.8519,敏感度为0.9091,特异度为0.9434。受试者工作特征曲线(receiver operating characteristic curve, ROC)是具有不同阈值的可能真阳性和假阳性的全部集合,而ROC曲线下的面积(area under ROC curve, AUC)可作为模型选择的标准^[7]。AUC取值范围一般为[0.5, 1],AUC值越大的分类模型,正确率越高^[8]。将不分体位的20%数据验证后绘制随机森林模型的ROC曲线(图3),得到AUC为0.967。

图1 咳嗽音和清嗓音的声波图和声谱图



注:声波图的垂直轴为声波的振幅(kU),声谱图的垂直轴为声音的频率(kHz)。各图水平轴均为时间(s)。此处为截图,时间轴起点不为0。

表1 参数含义

类别及参数	单位	意义
时域特征		
音长	s	从声音开始至结束所持续的时间
半耗能时长	s	声音开始至能量消耗一半时所持续的时间
达峰时间	s	从峰值的10%上升至峰值的90%所用时长
频域特征		
峰值频率	Hz	在声音片段的全范围内的最高频率
强度达峰时最高频率	Hz	强度达到峰值时所含声音片段的最高频率
时频域特征		
总耗能	dB	声音片段的总能量
峰值功率	dB	声音片段的最大功率值
信息域特征		
聚合熵	bit	声音片段中的能量分布,可衡量其无序性。

图2 全子集回归图



①截距;②音长;③半耗能时长;④达峰时间;⑤总耗能;⑥峰值功率;⑦峰值频率;⑧强度达峰时的最高频率;⑨聚合熵

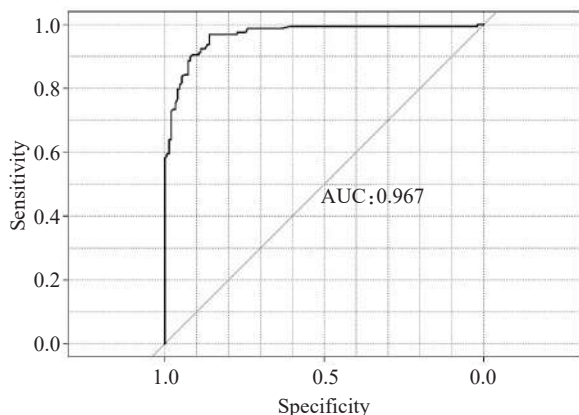
注:横轴为各声音指标,纵轴为校正后的(adj-R²)

表2 4个模型间准确度与Kappa值比较

	LDA	CART	RF	SVM
准确度	0.8770(0.8188—0.9353)	0.8453(0.7698—0.8913)	0.9162(0.8561—0.9420) ^①	0.8734(0.8188—0.9568)
Kappa值	0.7540(0.6377—0.8704)	0.6905(0.5395—0.7826)	0.8323(0.7123—0.8840) ^①	0.7468(0.6377—0.9136)

①模型间比较,RF具有最高准确度及最高Kappa值;LDA:线性判别分析;CART:分类回归决策树;RF:随机森林;SVM:线性分类器型支持向量机

图3 受试者工作特征(ROC)曲线



3 讨论

咳嗽训练是气道廓清技术的重要组成部分,通常由治疗师手法辅助,如胸廓按压法^[9-10]、腹部推挤按压法、前胸壁按压法等,或者由患者自助进行咳嗽加强训练^[11]。但在咳嗽训练中,治疗师手法辅助的力度是否合适,患者是否产生了有效咳嗽,是需要治疗师当场进行主观判断的。但是,这事实上又是实施治疗训练的重点与难点。咳嗽能够清除气道深部异物,而清嗓动作仅清除喉部刺激物。然而,当患者咳嗽力量减弱时,常与清嗓动作混淆,仅凭人耳难以准确判断^[3]。目前关于咳嗽音的研究主要集中于将咳嗽作为症状用作诊断疾病^[12],而不是把咳嗽和清嗓当作气道保护行为进行管理,也未涉及与康复治疗技术的结合。因此到目前为止,尚未建立成熟且广泛应用的咳嗽音检测、记录和评价体系。

本研究采用声学平台制备声学样本库后,使用RavenPro1.5初步分析声音片段的声学特征参数,并得到声音的声波和声谱图(图1)。咳嗽通常分为三个连续阶段:①吸气阶段:声带外展,吸气肌收缩,肺内气体体积增加^[13]。②压缩阶段:声门闭合,为形成足够的胸腔压力做准备,腹肌、肋间肌等肌肉主动收缩增加腹内压和胸腔内压^[14];③呼气阶段:声门快速

开放,气体高速冲出,并振动声带产生典型的咳嗽音。图1显示,在一次咳嗽过程中,声波出现两个能量峰值,而清嗓时仅产生一个能量峰值。因此,通过咳嗽音和清嗓音的声波、声谱图可以对这两种动作进行初步区分。

机器学习作为一种人工智能技术,可构建具有数据学习能力的算法。通过提取咳嗽和呼吸音中的特征可用来描述呼吸道的特征。再把声音数据库与已知的临床诊断进行特征匹配后,通过机器学习可找到这些特征的最佳组合,以创建精确的诊断测试或严重程度测量,并构建分类器^[15]。分类过程中的一个重要组成部分是选择适量的特征来表示数据。本文从时域、频域、时频域和信息域方面研究了咳嗽音的特征。有研究表明,咳嗽音的功率、能耗与咳嗽流速和咳嗽食管压力的相关性较好^[16]。此外,持续时间^[16]、峰值频率^[16]、强度达峰时的最高频率^[17]、强度达峰时间^[18]、半耗能时长^[19]、熵^[20]等参数也常作为人体声学的特征指标。本研究选取这些参数构建特征矢量,并用全子集回归对特征指标进行筛选。结果显示达峰时间、总能耗、峰值功率、强度达峰时的最高频率和聚合熵适于进一步探索建立分类模型。

为获得高准确度和最佳效率,数据处理方法和分类算法的选择至关重要。通过总结以往咳嗽检测和分类工作所采用的方法,可以发现随机森林和线性分类器型支持向量机具有较高的敏感度和特异性^[15]。线性判别分析由于求解简单,也常被用于声学特征分析^[21]。此外,分类回归决策树对语音检测的正确率也较高^[22]。因此,本研究选取上述4个模型进行模型间比较。结果显示,随机森林方法所建分类器具有更高的准确度和一致性。因此,最终选取随机森林方法进行建模。最后经过20%测试集的检验,得到ROC曲线下面积为0.967,提示模型具有较好的敏感度和特异度。

有效咳嗽的训练体位,当前普遍采取上身前倾位^[23]。而肺康复训练过程中,治疗师根据患者实际

情况,则会在卧位、半卧位和坐位等多个体位进行咳嗽手法辅助^[11,24]。如何判断在这些体位和操作下能否产生有效咳嗽,目前仍缺乏客观的评价方法。为进一步为手法辅助操作提供体位对照数据,本研究在咳嗽音与清嗓音的采集中,选取了平卧位、45°半卧位和90°端坐位。通过数据验证发现,不同体位的准确度和一致性相近,平卧位和端坐位的准确度和一致性略高于半卧位。因此,不同体位对咳嗽音的影响需要进一步探索。此外,咳嗽音的影响因素较多,如性别、年龄、体重、疾病等^[25],均有待进一步的研究。

本研究的意义,主要在于为临床肺康复训练提出一个新的问题:如何实时并持续记录、客观评估和反馈判断治疗师干预下患者的气道廓清行为?可以在本研究基础上开拓新思路:用声学分析来改造传统听诊技术,实现基于机器学习算法的计算机辅助气道廓清肺康复训练。例如,在咳嗽训练的同时,通过采集声音并呈现声波图像或参数,判断每一次尝试咳嗽是否真正实施的是咳嗽动作,开发基于咳嗽音的生物反馈训练技术。本研究的不足之处在于样本量较小,试验对象仅为青年健康志愿者。因此,本研究目前仅为探索性分析,其所得分类器模型并不能外推用于中年或老年人的气道廓清行为识别。后续研究将在本研究的基础上增大样本量以及丰富试验对象来源,进一步完善咳嗽音与清嗓音声学常模样本库,以形成供计算机深度学习的数据集,最终建立可实时记录和区分咳嗽和清嗓动作的模型。从本研究最终所用的80%分析数据集来看,采用了500次以上的咳嗽或清嗓数据,已经可以为初步建立分类模型提供依据。在今后应用中,可将咳嗽音声信号与常人数据比对,判断主动咳嗽的动作是否具有气道廓清功能的有效咳嗽,并由此提供听觉或视觉信息给患者或肺康复治疗师,由患者对主动咳嗽行为进行调整,或由治疗师对辅助咳嗽手法等康复操作加以控制和改进。最终通过反馈循环的形成,达到实施并记录有效气道廓清动作的目的。

参考文献

- [1] Rochester CL, Vogiatzis I, Holland AE, et al. An Official American Thoracic Society/European Respiratory Society Policy Statement: enhancing implementation, use, and delivery of pulmonary rehabilitation[J]. *Am J Respir Crit Care Med*, 2015, 192(11):1373—1386.
- [2] Volsko TA. Airway clearance therapy: finding the evidence[J]. *Respir Care*, 2013, 58(10):1669—1678.
- [3] Xiao Y, Carson D, Boris L, et al. The acoustic cough monitoring and manometric profile of cough and throat clearing: differentiate cough and throat clearing[J]. *Dis Esophagus*, 2014, 27(1):5—12.
- [4] Ladds MA, Thompson AP, Kadar JP, et al. Super machine learning: improving accuracy and reducing variance of behaviour classification from accelerometry[J]. *Animal Biotelemetry*, 2017, 5(1):8.
- [5] Viera AJ, Garrett JM. Understanding interobserver agreement: the Kappa statistic[J]. *Fam Med*, 2005, 37(5):360—3.
- [6] Maroco J, Silva D, Rodrigues A, et al. Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests[J]. *BMC Res Notes*, 2011, 4:299.
- [7] Wang Y, Chen H, Schwartz T, et al. Assessment of a disease screener by hierarchical all-subset selection using area under the receiver operating characteristic curves[J]. *Stat Med*, 2011, 30(14):1751—1760.
- [8] 汪云云, 陈松灿. 基于AUC的分类器评价和设计综述[J]. *模式识别与人工智能*, 2011, 24(1):64—71.
- [9] Branson RD. Secretion management in the mechanically ventilated patient[J]. *Respiratory Care*, 2007, 52(10):1328—1342.
- [10] Nozoe M, Mase K, Ogino T, et al. Effects of chest wall compression on expiratory flow rates in patients with chronic obstructive pulmonary disease[J]. *Braz J Phys Ther*, 2016, 20(2):158—165.
- [11] Frownfelter D, Dean E. Cardiovascular and pulmonary physical therapy-E-Book: evidence to practice[M]. Elsevier Health Sciences, 2014.
- [12] Schmit KM, Coeytaux RR, Goode AP, et al. Evaluating cough assessment tools: a systematic review[J]. *Chest*, 2013, 144(6):1819—1826.
- [13] Mellies U, Goebel C. Optimum insufflation capacity and peak cough flow in neuromuscular disorders[J]. *Ann Am Thorac Soc*, 2014, 11(10): 1560—1568.
- [14] Watchie J. Cardiovascular and pulmonary physical therapy: a clinical manual[M]. St. Louis, Mo: Saunders/Elsevier, 2010.

(下转第1443页)