

· 临床研究 ·

数字化音频分析在吞咽障碍患者误吸风险评估中的应用

张耀文 陈华玉 唐志明 窦祖林 赵妃 林依秋 武惠香 史静 温红梅 万桂芳

中山大学附属第三医院康复医学科, 广州 510630

通信作者: 唐志明, Email: tangzhm3@mail.sysu.edu.cn

【摘要】 目的 基于 Praat 语音分析软件探讨声学数字化分析方法在吞咽障碍患者的误吸风险评估中的应用。**方法** 选取 46 例脑卒中吞咽障碍患者, 每例患者均进行标准化的吞咽造影检查, 要求每例患者在进食每口食物前、后分别发“衣/i/”音 3 次, 采集声音样本。根据渗漏-误吸分级 (PAS) 将所有患者分成非误吸组 (16 例) 和误吸组 (30 例), 检测和比较 2 组患者进食前和进食后的组内及组间嗓音的基频、相对平均扰动度 (RAP)、相对频率微扰 (Jitter)、相对振幅微扰 (Shimmer)、振幅微扰商 (APQ) 和谐波噪声比 (HNR) 等声学参数的变化情况。**结果** 非误吸组患者的各项声学参数指标进食前后比较, 数据差异均无统计学意义 ($P>0.05$); 误吸组患者的基频、RAP 和 Jitter 数据进食前后差异有统计学意义 ($P<0.01$)。进食前, 非误吸组患者的 RAP [$(0.31\pm0.22)\%$] 和 Jitter [$(0.49\pm0.32)\%$] 分别与误吸组的 RAP [$(1.02\pm0.53)\%$] 和 Jitter [$(1.71\pm0.84)\%$] 比较, 组间差异有统计学意义 ($P<0.01$), 且非误吸组优于误吸组; 进食后, 仅 Jitter 值 [非误吸组 $(0.49\pm0.32)\%$; 误吸组 $(0.74\pm0.43)\%$] 的组间差异有统计学意义 ($P=0.048$)。**结论** 可在进食评估前, 通过对声学参数 RAP 和 Jitter 的分析, 识别吞咽障碍患者误吸的风险。

【关键词】 吞咽障碍; 误吸; 数字化音频分析

基金项目: 国家自然科学基金面上项目 (81472153); 广州市科技计划项目 (202007030001)

DOI: 10.3760/cma.j.issn.0254-1424.2021.12.003

Application of digital acoustic analysis in assessing aspiration risk among patients with dysphagia

Zhang Yaowen, Chen Huayu, Tang Zhiming, Dou Zulin, Zhao Fei, Lin Yiqiu, Wu Huixiang, Shi Jing, Wen Hongmei, Wan Guifang

Department of Rehabilitation Medicine, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou 510630, China

Corresponding author: Tang Zhiming, Email: tangzhm3@mail.sysu.edu.cn

【Abstract】 Objective To explore the application of digital acoustic analysis in assessing the risk of aspiration among persons with dysphagia using the Praat speech analysis software. **Methods** The swallowing of 46 stroke survivors with dysphagia was studied using video fluoroscopy. Each patient was required to pronounce "yi/i/" 3 times before and after taking each mouthful of food, and their voice samples were collected. The subjects were divided into a non-aspiration group of 16 and an aspiration group of 30 based on their penetration-aspiration scale scores. Fundamental frequency, relative average perturbation (RAP), jitter, shimmer, amplitude perturbation quotient and harmonic-to-noise ratio were compared between the two groups before and after taking food. **Results** For the non-aspiration group, there was no significant difference in the acoustic data before and after eating. For the aspiration group there were significant differences in fundamental frequency, relative average perturbation and jitter before and after taking food. The average RAP and jitter of the non-aspiration group were significantly better than the aspiration group's averages before eating. After eating, however, significant differences were observed only in average jitter. **Conclusions** Analyzing perturbation and jitter can help to identify persons at risk of aspiration.

【Key words】 Dysphagia; Aspiration; Digital acoustic analysis

Funding: A General Project of the National Natural Science Foundation of China (81472153); Science and Technology Program of Guangzhou (202007030001)

DOI: 10.3760/cma.j.issn.0254-1424.2021.12.003

调查研究显示,我国脑卒中急性期和恢复期患者患有吞咽功能障碍的比例分别为 46.3% 和 59.0%^[1]。吞咽障碍会引起渗漏或误吸导致吸入性肺炎,病程在一年内的脑卒中患者中约有 12.6% 存在吸入性肺炎^[2],严重影响患者的康复进程和生活质量。临床上常使用吞咽造影检查 (videofluoroscopic swallow study, VFSS) 作为吞咽功能评估的金标准,但是受限于仪器设备、人员配置和放射暴露等因素^[2],临床对于患者是否存在误吸的评估常依据患者进食评估前后是否发生声音的改变^[4]。其原理在于渗漏或误吸可能会导致患者声带局部有食物聚积^[5],改变声带的弹性和体积,导致声带振动发生变化^[6],影响患者声音的质量^[7]。因此,本研究采用荷兰阿姆斯特丹大学语音科学研究所开发的开源软件 Praat 对患者进食前后的声音变化进行分析,探究声学测量方法能否反映患者误吸的风险,为吞咽障碍的临床评估提供更加高效且便捷的检测手段。

对象与方法

一、研究对象及分组

入选标准:①经 CT 或 MRI 检查诊断为脑卒中或脑肿瘤患者^[8];②年龄 18~80 岁;③经改良洼田饮水试验筛查存在吞咽障碍^[9];④患者对本研究知情同意并签署相关文件。

排除标准:①气管切开或声带麻痹导致的失声;②严重听理解障碍(无法遵嘱完成一步指令);③既往有鼻咽癌或其它头颈部肿瘤放化疗病史;④经喉镜检查发现声带存在结节或肿物。

选取 2020 年 10 月至 2021 年 5 月在中山大学附属第三医院康复医学科住院及门诊治疗且符合上述标准的吞咽障碍患者 46 例,其中男 27 例,女 19 例;年龄 27~93 岁,平均(59.71±16.90)岁;脑出血 17 例,脑梗死 22 例,脑肿瘤 7 例;脑干病变 16 例,大脑半球病变 30 例。

渗漏-误吸分级 (penetration-aspiration scale, PAS)^[10]:共分为 1~8 级。①1 级,指无渗漏和误吸;②2 级,指食物进入气道,存留在声带以上,并被清除出气道;③3 级,指食物进入气道,存留在声带以上,未被清除出气道;④4 级,指食物进入气道,附着在声带上,并被清除出气道;⑤5 级,指食物进入气道,附着在

声带上,未被清除出气道进入声带下;⑥6 级,指食物进入气道并附着在声带以下,但可被清除出气道或清除入喉部;⑦7 级,指食物进入气道声带以下,虽用力亦不能清除出气管;⑧8 级,指食物进入气道声带以下,无用力清除表现。

本研究仅考虑 1 级和 5~8 级,即无误吸或有不同程度的误吸,不分析渗漏的情况。根据患者是否误吸,将患者分成非误吸组 (PAS 分级 1 级,共纳入 16 例) 和误吸组 (PAS 分级>4 级,共纳入 30 例),2 组患者的性别、平均年龄、疾病类别和病变部位等一般临床资料经统计学分析比较,差异均无统计学意义 ($P>0.05$),详见表 1。本研究获中山大学附属第三医院伦理委员会审核批准 (中大附三医伦 2020_02-258-01)。

二、研究方法

46 例患者均完成标准化的 VFSS 评估,即依次进食中稠度、低稠度、稀流质和高稠度食物 (各 3、5 和 10 ml),以最大的程度避免患者产生大量误吸的风险。如果患者进食过程中出现误吸,则即刻嘱患者完成进食后的 3 次发音,完成后马上给予口腔清洁及吸痰处理,完成后进行胸部透视,确认无对比剂在气管或肺内残留。

造影过程全程录音录像,采用三星 Galaxy S6 连接 Sennheiser 麦克风完成音视频采集;采集过程全程在吞咽造影室内进行 (非进食状态下无辐射),造影室采用铅门密闭,除不可避免造影机产生的环境杂音外,无其它杂音。采集内容包括患者进食前发元音“衣/i/”3 次、患者造影评估的全过程、患者每次进食后发元音“衣/i/”3 次。采集到的视频使用 Adobe Premiere Pro CC 2019 软件进行音轨分离,使用造影室内无人状态下的环境音进行降噪处理,并生成 MP3 格式文件 (音频采样速率为 48 kHz,比特深度为 24),将文件导入 Praat 软件,采用软件预置插件和脚本进行分析。

三、测量指标

Praat 声学参数^[11]:共观察 6 项声学指标。①嗓音的基频 (fundamental frequency)——反映声带每秒钟的振动次数,可以反映声带是否发生病理性改变,对声学参数的信度进行校对,与患者发音时的用力程度成正相关;②相对平均扰动度 (relative average perturbation, RAP)——即基音周期相邻 3 点扰动,是某一帧基音周期与其相邻 3 帧的平均基音周期之差的平均

表 1 2 组患者的一般临床资料

组别	例数	性别 (例)		平均年龄 (岁, $\bar{x}\pm s$)	疾病类别 (例)			病变部位 (例)	
		男	女		脑出血	脑梗死	脑肿瘤	脑干	大脑半球
非误吸组	16	8	8	60.62±15.16	7	7	2	7	9
误吸组	30	19	11	59.23±17.99	10	15	5	9	21

绝对值,一定程度上反映一段时间受试者对声带振动的控制能力,RAP 值越大,声音稳定性越差;③相对频率微扰(Jitter)——即基音周期的相对扰动,是相邻基音周期差的平均绝对值,与基音周期平均值的比值,反映整体受试者对声带振动的相对控制能力,Jitter 值越大,声带振动的控制能力越差;④相对振幅微扰(Shimmer)——即振幅扰动,是相邻振幅差的平均绝对值与振幅平均值的比值,反映受试者对语音幅度的相对控制能力,Shimmer 值越大,声音嘶哑程度越高;⑤振幅微扰商(amplitude perturbation quotient, APQ)——即振幅相邻扰动,是某一帧振幅与其相邻 5 帧平均振幅之差的平均绝对值,一定程度上反映一段时间受试者对语音幅度的控制能力,APQ 越高,声带振动的规律、振幅和声门闭合状态越好;⑥谐波噪声比(harmonic noise ratio, HNR)——用以分析噪音中的噪声成分,比值越高则噪声越少。

四、统计学方法

使用 SPSS 22.0 版统计学软件包进行数据分析,所有声学参数指标采用 Shapiro-Wilk 检验是否符合正态分布,组内资料采用配对样本 t 检验、组间资料采用独立样本 t 检验进行差异性检验。 $P<0.05$ 认为差异有统计学意义。

结 果

非误吸组患者基频、RAP、Jitter、Shimmer、APQ 和 HNR 等各项声学参数指标的进食前后差异均无统计学意义($P>0.05$),但误吸组患者基频、RAP、Jitter 指标的进食前后差异有统计学意义($P<0.01$)。组间比较,进食前,2 组间的 RAP 和 Jitter 组间差异有明显统计学意义($P<0.01$),且两项指标非误吸组均优于误吸组;进食后 2 组间仅 Jitter 组间差异有统计学意义($P=0.048$),RAP 组间差异无统计学意义。详见表 2。

讨 论

语音分析是一种基于对语音信号数字化处理的无创且客观的评估技术,是基于声学数字信号波的周期性早期检测和诊断声音异常的有效方法^[12]。误吸常

常会导致患者声音嘶哑或出现湿性嗓音^[13],这可能是误吸导致声带上有食物残留,使声带的振动发生变化,进而导致患者的语音发生改变,因此声学参数可以客观地反映这种变化。本研究结果显示,进食前,误吸组 RAP 和 Jitter 的平均值分别为 $(1.02\pm0.53)\%$ 和 $(1.71\pm0.83)\%$,而非误吸组则为 $(0.31\pm0.22)\%$ 和 $(0.49\pm0.32)\%$ 。RAP 是计算 3 个连续声带振动周期的平均周期时长得出的抖动度量值,Jitter 是利用声带振动周期的时长计算得出,声带振动频率变异性越大,两者数值越大,直接反映声带振动频率的稳定性。这两项指标的差异提示,在进食前可以通过声学参数分析对患者是否存在误吸风险进行预测,这与 Alvares 等^[14]的研究结果相一致,即使用部分声学参数(如 RAP 和 NHR)评估吞咽障碍的敏感度高达 100%。进食后,误吸组患者 Jitter 的平均值为 $(0.74\pm0.43)\%$,而非误吸组则为 $(0.49\pm0.32)\%$,两组间差异有统计学意义($t=2.033, P=0.048$),提示进食后误吸组患者的声带控制能力仍较非误吸组患者差。因此,在进行吞咽造影评估之前完成声学参数分析,可以有效降低进食评估过程中发生大量误吸的风险。

本研究中,误吸组组内比较结果显示,进食后误吸组患者的基频、RAP 和 Jitter 均与进食前有明显的差异,说明患者在进食后发音过程中声带振动的稳定性和规律性有所提高,即进食后的声学表现优于进食前,这与以往临床研究中的部分发现不符,主要存在 2 种可能:①患者在进食后基频的均值为 (188.66 ± 40.43) Hz,较进食前 (172.61 ± 44.06) Hz 有显著提高($t=-6.92, P<0.01$),提示患者进食后发音时的用力程度更高,增强了对声带振动稳定度的控制能力;而 Santos 等^[15]研究发现,存在误吸风险的患者会尝试更加用力地进行发音以清除气道内残留的食物,患者在误吸后发音时会更加用力,这种用力发音可以减少患者发声过程中声带振动的不规则性,从而提高声音的表现,与本研究结果一致;②部分患者发生误吸后(特指存在显性误吸的患者),会产生咳嗽反射,这种咳嗽除可将误吸物清除出气道外,在咳嗽过程中也会将咽喉部原有的分泌物清除出气道,可以使患者的声带振动幅度恢复到正常

表 2 2 组患者进食前后的声学参数比较($\bar{x}\pm s$)

组别	例数	基频(Hz)	RAP(%)	Jitter(%)	Shimmer(%)	APQ(%)	HNR(%)
非误吸组							
进食前	16	193.20±38.82	0.31±0.22	0.49±0.32	9.12±4.48	7.89±4.59	14.97±3.30
进食后	16	192.92±38.79	0.31±0.21	0.49±0.32	9.11±4.46	7.89±4.60	14.77±3.36
误吸组							
进食前	30	172.61±44.06	1.02±0.53 ^a	1.71±0.84 ^a	9.69±4.57	7.14±3.05	12.99±5.81
进食后	30	188.66±40.43 ^b	0.32±0.18 ^b	0.74±0.43 ^{bc}	9.10±4.58	7.96±4.42	14.36±4.70

注:与非误吸组进食前比较,^a $P<0.01$;与误吸组进食前比较,^b $P<0.01$;与非误吸组进食后比较,^c $P<0.05$

水平,从而减少进食后声带振动时的不规则扰动,表现为 RAP($t=9.12, P<0.01$) 和 Jitter($t=5.73, P<0.01$) 均较进食前显著降低。

综上所述,采用通用的声学数字化分析软件对吞咽障碍患者的声学参数进行分析发现,误吸症状发生时进食前后的声学参数有明显的特征学改变,通过分析患者语音声学参数(特别是 RAP 和 Jitter)的改变可以识别吞咽障碍患者误吸的风险,为客观识别临床误吸体征提供客观依据。但由于本研究中纳入的误吸组患者数量较少,没有根据误吸食物的性状进行分类分析,在未来的研究中可以考虑纳入更多的病例,分析误吸不同性状的食物对患者嗓音的影响。

参 考 文 献

- [1] 李超,张梦清,窦祖林,等.中国特定人群吞咽功能障碍的流行病学调查报告[J].中华物理医学与康复杂志,2017,39(12):937-943. DOI:10.3760/cma.j.issn.0254-1424.2017.12.014.
- [2] Feng MC, Lin YC, Chang YH. The mortality and the risk of aspiration pneumonia related with dysphagia in stroke patients[J]. J Stroke Cerebrovasc Dis, 2019, 28(5):1381-1387. DOI:10.1016/j.jstrokecerebrovasdis.2019.02.011.
- [3] 邱莉芳,陈晓锋,金欣.脑卒中后吞咽障碍的研究与展望[J].黑龙江中医药,2018,47(2):110-112.
- [4] 万桂芳,张耀文,史静,等.改良容积粘度测试在吞咽障碍评估中的灵敏性及特异性研究[J].中华物理医学与康复杂志,2019,41(12):900-904. DOI:10.3760/cma.j.issn.0254-1424.2019.12.004.
- [5] 牛丽,李彦杰,秦合伟,等.脑卒中后吞咽障碍康复评估和治疗研究进展[J].中国医药导报,2020,17(28):48-51.
- [6] Santamato A, Panza F, Solfrizzi V, et al. Acoustic analysis of swallowing

sounds: a new technique for assessing dysphagia[J]. J Rehabil Med, 2009, 41(8):639-645. DOI:10.2340/16501977-0384.

- [7] Groves-Wright KJ, Boyce S, Kelchner L. Perception of wet vocal quality in identifying penetration/aspiration during swallowing[J]. J Speech Lang Hear Res, 2010, 53(3):620-632. DOI:10.1044/1092-4388(2009/08-0246).
- [8] 刘鸣,刘峻峰,吴波.脑血管病分类分型进展与解读[J].中华神经科杂志,2017,50(3):163-167. DOI:10.3760/cma.j.issn.1006-7876.2017.03.002.
- [9] 温水群,邓振兴,徐秋萍,等.改良洼田饮水试验在脑卒中吞咽障碍患者中的研究应用[J].中国当代医药,2018,25(16):47-49. DOI:10.3969/j.issn.1674-4721.2018.16.013.
- [10] Rosenbek JC, Robbins JA, Roecker EB, et al. A penetration-aspiration scale[J]. Dysphagia, 1996, 11(2):93-98. DOI:10.1007/BF00417897.
- [11] 叶志腾.应用 Praat 软件分析成人嗓音声学参数[D].福建:福建医科大学,2009. DOI:10.7666/d.y1487533.
- [12] 刘恒鑫,李朝阳,郝建萍.远程嗓音评估与治疗研究进展[J].中国听力语言康复科学杂志,2021,19(3):195-199. DOI:10.3969/j.issn.1672-4933.2021.03.009.
- [13] 中国吞咽障碍康复评估与治疗专家共识组.中国吞咽障碍评估与治疗专家共识(2017年版)[J].中华物理医学与康复杂志,2018,40(1):1-10. DOI:10.3760/cma.j.issn.0254-1424.2017.12.001.
- [14] Alvares CM, Devadiga D, Jeevan M, et al. Analysis of cough in post-stroke individuals with dysphagia[J]. J Clin Diagn Res, 2018, 12(12):OC01-OC05. DOI:10.7860/JCDR/2018/37417.12326.
- [15] Santos KWD, Scheeren B, Maciel AC, et al. Vocal variability post swallowing in individuals with and without oropharyngeal dysphagia[J]. Int Arch Otorhinolaryngol, 2015, 19(1):61-66. DOI:10.1055/s-0034-1394129.

(修回日期:2021-10-29)

(本文编辑:汪 玲)

· 读者 · 作者 · 编者 ·

本刊对参考文献的有关要求

执行 GB/T 7714-2005《文后参考文献著录规则》。采用顺序编码制著录,依照其在文中出现的先后顺序用阿拉伯数字标出,并将序号置于方括号中,排列于文后。内部刊物、未发表资料(不包括已被接受的待发表资料)、个人通信等请勿作为文献引用。日文汉字请按日文规定书写,勿与我国汉字及简化字混淆。同一文献作者不超过 3 人全部著录;超过 3 人只著录前 3 人,后依文种加表示“等”。作者姓名一律姓氏在前、名字在后,外国人的名字采用首字母缩写形式,缩写名后不加缩写点;不同作者姓名之间用“,”隔开,不用“和”、“and”等连词。题名后请标注文献类型标志。文献类型标志代码参照 GB 3469-1983《文献类型与文献载体代码》,如参考文献类型为杂志,请于参考文献末尾标注 DOI 号。中文期刊用全名。示例如下。

- [1] 陈登原.国史旧闻[M].北京:中华书局,2000:29.
- [2] 胡永善.运动功能评定//王茂斌.康复医学[M].2 版.北京:人民卫生出版社,2002:67-78.
- [3] 刘欣,申阳,洪葵,等.心脏性猝死风险的遗传检测管理[J].中华心血管病杂志,2015,43(9):760-764. DOI:10.3760/cma.j.issn.0253-3758.2015.09.003.
- [4] Mahowald ML, Krug HE, Singh JA, et al. Intra-articular Botulinum Toxin Type A: a new approach to treat arthritis joint pain[J]. Toxicon, 2009, 54(5):658-667. DOI:10.1016/j.toxicon.2009.03.028.
- [5] 余建斌.我们的科技一直在追赶:访中国工程院院长周济[N/OL].人民日报,2013-01-12(2). [2013-03-20].http://paper.people.com.cn/rmrb/html/2013-01/12/nw.D110000renmrb_20130112_5-02.htm.