

• 论著——临床研究 •

基于发声与言语障碍分析参数对病理嗓音的检测^{*}魏梅¹ 杜建群¹ 耿磊² 王巍¹

[摘要] 目的:利用发声与言语嗓音障碍分析(ADSV)软件对病理嗓音进行检测,明确 ADSV 参数在识别病理嗓音中的作用。方法:分别采集 113 例病理嗓音患者(声带息肉患者 93 例、声门型喉癌患者 20 例)以及 47 例嗓音正常志愿者的持续元音和连续言语的声学信号,采用 ADSV 分析各组在持续元音和连续言语下的嗓音参数:倒频谱峰值(CPP)、倒频谱峰值标准差(CPP SD)、低/高频谱比(L/H ratio)、低/高频谱比标准差(L/H ratio SD)及嗓音障碍倒频谱/频谱指数(CSID),探究这些参数在识别病理嗓音中的作用。结果:在持续元音声学信号中,正常组的 CPP、L/H ratio 值均大于病理嗓音组($P < 0.001$),CPP SD、CSID 则小于病理嗓音组($P < 0.001$),其中 CPP、CSID 的 ROC 曲线下面积分别为 0.95、0.99,是诊断病理嗓音的重要声学参数。在连续言语声学信号中,正常组的 CPP、CPP SD、L/H ratio 均大于病理嗓音组($P < 0.001$),其中 CPP SD 的曲线下面积为 0.90,对于诊断病理嗓音具有较高的准确性。同时 ADSV 嗓音分析参数 CPP、CPP SD、CSID、L/H ratio 在声带息肉组和声门型喉癌组之间的差异也有统计学意义($P < 0.05$)。判别分析模型结果显示利用 ADSV 嗓音参数不仅能够区别病理嗓音,而且能够较好地区分声带息肉和喉癌。结论:ADSV 嗓音分析参数不仅能够区分正常与病理嗓音信号,而且还能区分不同类型的病理嗓音,其在诊断病理嗓音上具有较高的敏感性和特异性。

[关键词] 发声与言语嗓音障碍分析;病理嗓音;嗓音参数

DOI:10.13201/j.issn.2096-7993.2022.07.002

[中图分类号] R767.92 [文献标志码] A

Detection of speech pathology based on parameters of analysis of dysphonia in speech and voice

WEI Mei¹ DU Jianqun¹ GENG Lei² WANG Wei¹

(¹Department of Otorhinolaryngology Head and Neck Surgery, Tianjin First Central Hospital, Institute of Otolaryngology of Tianjin, China Key Laboratory of Auditory Speech and Balance Medicine, Key Medical Discipline of Tianjin[Otolaryngology], China Quality Control Centre of Otolaryngology, Tianjin, 300192, China; ²School of Life Sciences, Tianjin University of Technology, Tianjin Key Laboratory of Photoelectric Detection Technology and System)

Corresponding author: WANG Wei, E-mail: wwei1106@hotmail.com

Abstract Objective: To analysis speech pathology based on dysphonia in speech and voice(ADSV). **Methods:** The acoustic signals of continuous vowels and continuous speech of one-hundred and thirteen individuals were collected, including 93 vocal cord polyps cases, 20 glottis laryngeal carcinoma cases and 47 volunteers without speech sound disorders. Cepstral peak prominence(CPP), CPP standard deviation(CPP SD), L/H spectral ratio(L/H ratio), L/H ratio standard deviation(L/H ratio SD) and cepstral/spectral index of dysphonia(CSID) were analyzed by ADSV to explore the role of these parameters in the recognition of speech pathology. **Results:** In the acoustic signal of continuous vowels, CPP and L/H ratio in normal group were higher than those in pathological voice group($P < 0.001$), while CPP SD and CSID were lower than those in pathological voice group($P < 0.001$), CPP and CSID areas under ROC curve were 0.95 and 0.99, respectively, which were important acoustic parameters for diagnosing pathological voice. In continuous speech acoustic signals, CPP, CPP SD and L/H ratio in the normal group were all higher than those in the speech disorders group($P < 0.001$), and the area under the curve of CPP SD was 0.90, which showed high accuracy in diagnosing pathological voice. The ADSV voice analysis parameters CPP, CPP SD, CSID, and L/H ratio also showed significant differences between the vocal cord polyp group

^{*}基金项目:天津市医学重点学科建设项目;天津市卫生健康委员会科技人才培养项目(No:KJ20136)

¹天津市第一中心医院耳鼻咽喉头颈外科 天津市耳鼻喉科研究所 天津市听觉言语与平衡医学重点实验室 天津市医学重点学科(耳鼻咽喉科学) 天津市耳鼻喉质量控制中心(天津,300192)

²天津工业大学生命科学院 天津光电检测技术与系统重点实验室

通信作者:王巍,E-mail:wwei1106@hotmail.com

and the glottic laryngeal cancer group. The results of the discriminant analysis model show that the use of ADSV voice parameters can distinguish vocal cord polyps and laryngeal cancers. **Conclusion:** The ADSV voice analysis parameters can not only distinguish the voice signals of the normal group and the pathological group, but also distinguish different types of pathological voices. It has high sensitivity and specificity in diagnosing pathological voices.

Key words analysis of dysphonia in speech and voice; pathological voice; voice parameters

随着社会经济的进步,嗓音健康状况被越来越多的人所关注,通过早期筛查诊断及时发现嗓音问题,是保证嗓音健康的重要方法。由于嗓音具有多维性特点,临床上多利用主观评估法和客观评估法对嗓音进行多方面的评估以确定其质量及障碍的严重程度^[1],其中客观评估法主要包括频闪喉镜、声学分析等^[2]。嗓音声学分析作为无创性的嗓音检测方法,在临床上广泛用于评估嗓音质量及嗓音疾病治疗效果^[3]。目前声学分析方法主要包括基于时间的扰动分析和基于频谱二次分析的倒频谱分析。基于时间的扰动分析一般用于分析持续元音,但由于周期变化检测的不可靠和不准确性,扰动分析可能难以全面评估发声困难。发声与言语嗓音障碍分析(analysis of dysphonia in speech and voice, ADSV)则是基于倒频谱理论,针对持续元音及连续语音的一种声学分析方法,相较于仅针对元音的声学分析方法,ADSV能够对持续的元音和言语状态下的元音成分进行计算,更能反映患者的嗓音状态,是评估发声困难和语音质量严重程度强有力的预测因子。该分析的主要参数包括倒频谱峰值(cepstral peak prominence, CPP),倒频谱峰值标准差(CPP standard deviation, CPP SD),低/高频谱比(L/H spectral ratio, L/H ratio),低/高频谱比标准差(L/H ratio standard deviation, L/H ratio SD),嗓音障碍倒频谱/频谱指数(cepstral/spectral index of dysphonia, CSID)。本研究通过ADSV软件分析正常组与病理嗓音组及不同类型病理嗓音组间在持续元音和连续言语下的声学特征,明确ADSV参数是否能够区分病理嗓音,验证利用嗓音参数预测病理嗓音的准确性,并综合评估ADSV参数在辨别病理嗓音中的应用价值。

1 资料与方法

1.1 研究对象

病理嗓音组选取2020年1月—2021年12月由天津市第一中心医院耳鼻咽喉头颈外科经频闪喉镜检查具有典型特征并最终确诊为声带息肉、声门型喉癌患者113例,男66例,女47例;年龄36~62岁;其中声带息肉患者93例(男46例、女47例),声门型喉癌患者20例(均为男性)。正常组选取志愿受试者47例,男23例,女24例;年龄31~

65岁;均无咽喉、呼吸系统等影响呼吸功能的疾病史,口咽部结构正常,无鼾症等症状,不吸烟,没有受过专业发声或者歌唱训练,频闪喉镜检查喉部无异常,声带黏膜波正常。病理嗓音组与正常组的年龄($\chi^2=1.552, P=0.213$)和性别($\chi^2=1.206, P=0.272$)差异无统计学意义。

1.2 ADSV分析

检测设备采用美国KAYPENTAX公司生产的4500语音工作站的ADSV软件,由同一位嗓音医师采用ADSV对受试者进行持续元音和连续言语声学分析。

持续元音声学信号采集:受试者测试前可以先做练习,测试在环境噪声小于45 dB的隔声室内完成,受试者口距麦克风约15 cm,保持自然舒适坐姿,以平稳舒适音调发长元音/a/,持续时间至少4 s,连续测3次。

连续言语声学信号采集:两组受试者在上述环境中以相同的发声方法诵读句子“床前明月光,疑是地上霜,举头望明月,低头思故乡”,收集声学信号。

利用ADSV软件对持续元音声学信号和连续言语声学信号进行分析,并得到嗓音声学参数:CPP, CPP SD, L/H ratio, L/H ratio SD, CSID。

1.3 统计学方法

通过SPSS 19.0软件进行统计分析,卡方检验用于检测病理嗓音组与正常组间年龄和性别的分布差异,不同组间声学参数的差异比较采用独立样本 t 检验。将差异有统计学意义的声学参数进行受试者工作特征曲线(ROC)分析,利用ROC曲线确定最佳临界值、敏感性、特异性、曲线下面积(AUC),当ROC曲线下面积值为0.5~0.7时有较低准确性,>0.7~0.9有一定的准确性,>0.9则有较高的准确性。利用Fisher判别法构建判别模型,综合评估ADSV嗓音参数区分病理嗓音的能力。

2 结果

2.1 正常组与病理嗓音组间持续元音和连续言语下ADSV分析结果

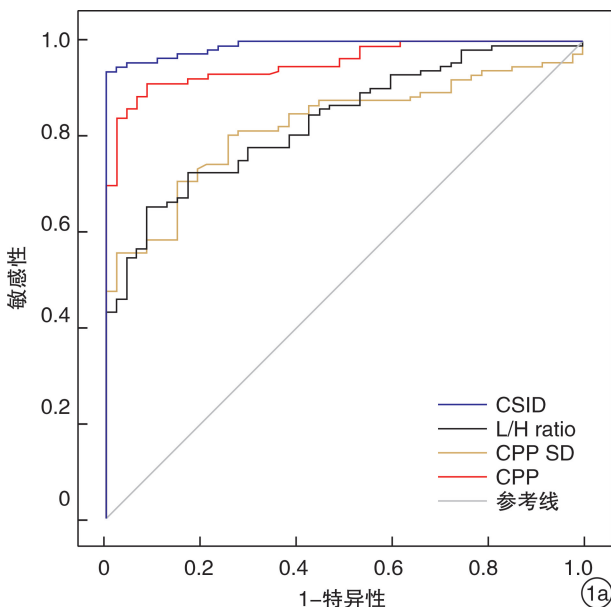
正常组与病理嗓音组持续元音和连续言语下ADSV分析参数比较见表1。利用ADSV软件对

持续元音声学信号进行分析发现,正常组中 CPP、L/H ratio 值大于病理噪音组, CPP SD、CSID 值小于病理噪音组, 两组比较差异均有统计学意义 ($P < 0.001$)。对两组连续言语声学信号进行分析发现,正常组中 CPP、CPP SD、L/H ratio 值均大于病理噪音组, 两组比较差异均有统计学意义 ($P < 0.001$)。

表 1 正常组与病理噪音组持续元音和连续言语下 ADSV 分析参数比较

参数	正常组	病理噪音组	P 值
持续元音下			
CPP	13.98±1.73	8.89±2.67	<0.001
CPP SD	0.73±0.21	1.30±0.62	<0.001
L/H ratio	32.64±4.19	26.40±5.61	<0.001
L/H ratio SD	1.79±0.64	1.99±0.77	0.130
CSID	4.43±8.53	42.31±18.81	<0.001
连续言语下			
CPP	5.89±0.79	0.79±1.34	<0.001
CPP SD	3.67±0.34	2.79±0.67	<0.001
L/H ratio	27.89±2.47	24.39±3.97	<0.001
L/H ratio SD	10.11±1.32	10.38±1.68	0.190

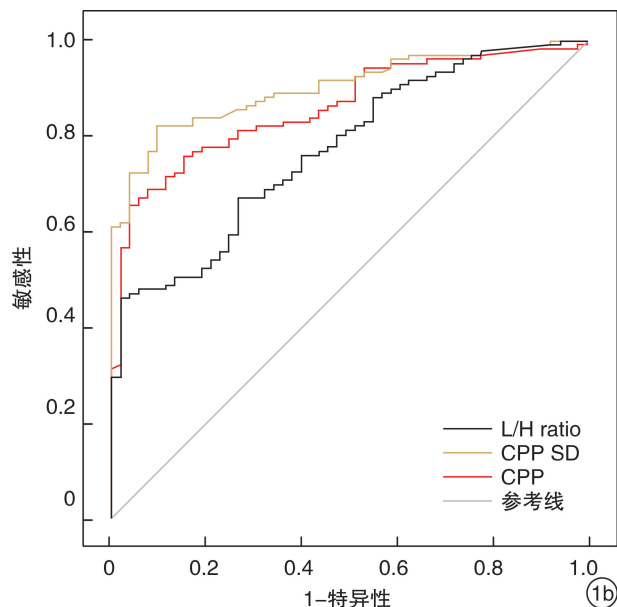
2.2 正常组与病理噪音组间持续元音和连续言语下 ADSV 分析参数的 ROC 曲线分析结果



利用 ROC 曲线分析正常组与病理噪音组之间有统计学差异的 ADSV 噪音分析参数发现,在持续元音声学信号中 CPP、CPP SD、L/H ratio、CSID 的曲线下面积均大于 0.7,对检测病理噪音均有一定的准确性,其中 CPP、CSID 的曲线下面积大于 0.9,对诊断病理噪音具有较高的准确性。在连续言语声学信号中 CPP、CPP SD、L/H ratio 的曲线下面积也大于 0.7,对检测病理噪音有一定的准确性,其中 CPP SD 的曲线下面积为 0.9,对诊断病理噪音具有较高的准确性(表 2,图 1)。

表 2 持续元音和连续言语下 ADSV 分析参数的 ROC 曲线结果

参数	曲线下面积	最佳临界值	敏感性	特异性	约登指数
持续元音下					
CPP	0.95	11.84	0.94	0.88	0.82
CPP SD	0.82	0.94	0.71	0.85	0.56
L/H ratio	0.83	28.21	0.92	0.66	0.57
CSID	0.99	19.59	0.94	1.00	0.94
连续言语下					
CPP	0.86	4.79	0.96	0.66	0.62
CPP SD	0.90	3.36	0.89	0.83	0.73
L/H ratio	0.76	24.30	0.98	0.47	0.45



1a:持续元音下 ADSV 分析参数的 ROC 曲线;1b:连续言语下 ADSV 分析参数的 ROC 曲线;与参考面积 0.5 的差异均有统计学意义。

图 1 持续元音和连续言语下 ADSV 分析参数的 ROC 曲线图

2.3 ADSV 分析参数识别正常与病理噪音的能力

利用判别分析方法评估 ADSV 噪音参数识别正常与病理噪音的能力。持续元音下, CPP、CPP

SD、L/H ratio、CSID 在两组中均有统计学差异且具有较高的诊断价值,因此将这 4 个参数一起纳入判别分析,分析发现,正常组 100% (47/47) 分类正

确,而病理嗓音组中 88.5%(100/113)的患者分类正确,有 13 例患者被分类到了正常组。在连续言语下,我们纳入了有统计学差异且有一定诊断价值的 CPP、CPP SD、L/H ratio 参数进行判别分析,正常组 95.7%(45/47)被正确分类,2 个正常人被判别为病理嗓音组,而病理嗓音组中 73.5%(83/113)的患者被正确分类,30 例患者被误判到正常组。

2.4 ADSV 分析参数在不同病理嗓音间的差异

不同病理嗓音组持续元音和连续言语下 ADSV 分析参数比较见表 3。在持续元音声学信号中,声带息肉组 CPP、L/H ratio 大于声门型喉癌组,CSID 值则小于声门型喉癌组,差异均有统计学意义($P < 0.001$)。在连续言语声学信号中,声带息肉组 CPP、CPP SD、L/H ratio 均大于声门型喉癌组($P < 0.001$)。

表 3 不同病理嗓音组持续元音和连续言语下 ADSV 分析参数比较

参数	声带息肉组	声门型喉癌组	P 值
持续元音下			
CPP	9.56±2.02	5.79±3.17	<0.001
CPP SD	1.33±0.65	1.21±0.52	0.420
L/H ratio	27.16±4.64	21.03±6.94	<0.001
L/H ratio SD	2.02±0.78	1.87±0.66	0.440
CSID	37.72±14.79	63.63±21.07	<0.001
连续言语下			
CPP	4.56±1.14	2.94±1.42	<0.001
CPP SD	2.97±0.51	1.98±0.74	<0.001
L/H ratio	25.36±2.79	21.06±6.26	<0.001
L/H SD	10.73±1.52	11.22±1.97	0.212

2.5 ADSV 分析参数识别不同病理嗓音的能力

利用判别分析方法评估 ADSV 嗓音参数识别不同病理嗓音的能力。持续元音信号中的 CPP、L/H ratio、CSID 在两组中的差异有统计学意义,因此将这 3 个参数纳入判别分析模型中,通过判别分析发现,声带息肉组 87.1%(81/93)的患者分类正确,声门型喉癌组 65.0%(13/20)的患者分类正确。在连续言语中,纳入了有统计学差异的 CPP、CPP SD、L/H ratio 参数进行判别分析,声带息肉组 88.2%(82/93)的患者被正确分类,声门型喉癌组 75.0%(15/20)的患者可被正确分类。

3 讨论

ADSV 软件得到的嗓音分析参数包括 CPP、CPP SD、L/H ratio、L/H ratio SD 和 CSID。CPP 能够反映信号的周期性,是一种有价值的临床嗓音参数,它不受录音技术、声音幅度/响度差异的影

响,是衡量整体发声障碍严重程度及语音质量的最佳嗓音参数^[4-6]。有研究表明 CPP 与声音严重程度和气息音有一定的相关性,但与声音的粗糙或紧张程度只有微弱的相关性,较低的 CPP 值表明发声困难程度高^[7-10]。本研究发现,不管是在连续元音还是在持续言语下,正常组的 CPP 值均高于病理嗓音组,与其他学者的研究结果一致^[7,11-13]。声带息肉组中 CPP 值也高于声门型喉癌组。

CPP SD 可用于评估整个发声持续过程中声音信号的可变性。本研究发现在持续元音信号中,相比正常组,病理嗓音组的 CPP SD 值更高,说明病理嗓音组患者在发元音时存在不稳定性。而在连续言语声学信号中,正常组的 CPP SD 值更高,这是由于连续言语存在的浊音、辅音、音高和音量的变化使 CPP SD 的幅值增加^[14],而病理嗓音患者则会降低。对比声带息肉组与声门型喉癌组,结果显示持续元音下两组间的差异无统计学意义,在连续言语中声带息肉组的 CPP SD 大于声门型喉癌组。

L/H ratio 指嗓音频谱中低频区与高频区间能量的比值,在 ADSV 中,低、高频的界限划分为 4000 Hz,有研究表明正常的嗓音信号主要集中在 4000 Hz 以下,正常语音信号往往具有较高的 L/H ratio。有发声障碍的个体含有较多的气息音,气息音的能量多集中在 4000 Hz 以下,因此发声障碍患者 L/H ratio 降低^[4]。本研究结果也发现正常组比病理嗓音组有更高的 L/H ratio。病理嗓音患者中声带息肉组的 L/H ratio 高于声门型喉癌组。

CSID 是声音的多维表示,是将 CPP、L/H ratio、CPP SD 和 L/H ratio SD 四个声学参数相结合的一个多元评估^[15]。研究发现,CSID 是一种合理准确评估发声困难严重程度的方法^[16],也是一种可灵敏检测治疗前后声音样本变化的方法^[2]。CSID 值的增加反映了嗓音严重性的增加^[17],但目前并没有明确的标准表明在何范围内的 CSID 值为正常。CSID 估计值通常介于 0 和 100 之间,但在某些情况下,它可以出现 0 或高于 100 的分数^[18]。有研究表明,CSID 能够很好地区分正常与病理嗓音,可作为潜在的嗓音筛查参数^[13,19]。本研究结果显示在持续元音下正常组的 CSID 值小于病理嗓音组,并且对于识别病理嗓音具有较高的准确性,在病理嗓音患者中,声带息肉组的 CSID 值也小于声门型喉癌组。

ROC 分析结果显示,持续元音和连续言语下的 ADSV 嗓音分析参数在区别正常嗓音和病理嗓音上均有一定的敏感性和特异性。为更好地评估

ADSV 嗓音分析参数区别病理嗓音的能力,我们建立了判别分析模型进行分析,研究表明,在持续元音下的 ADSV 嗓音分析参数识别正常嗓音和病理嗓音的准确率分别为 100% 和 88.5%,识别声带息肉和声门型喉癌的准确率分别为 87.1% 和 65.0%;在连续言语下的 ADSV 嗓音分析参数识别正常嗓音和病理嗓音的准确率分别为 95.7% 和 73.5%,识别声带息肉和声门型喉癌的准确率分别为 88.2% 和 75.0%。

本研究显示,ADSV 嗓音分析参数在正常组和病理嗓音组之间,以及不同病理嗓音组之间均有显著性差异,其识别病理嗓音具有较高的敏感性和特异性,而且 ADSV 嗓音分析参数在识别不同病理嗓音方面也有很好的区分能力。ADSV 作为既能分析持续元音又能分析连续言语的嗓音分析参数是传统声学分析方法的一个重要补充,在嗓音疾病筛查诊断中具有一定的应用前景。由于本研究病理嗓音患者有限,以及连续言语分析中句子的合理性等问题,ADSV 的临床价值还需要大量的样本进行深入的研究。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 刘杰,肖翠,邓利,等. 48 例双侧广基型声带息肉患者不同术式疗效的主客观评估[J]. 临床耳鼻咽喉头颈外科杂志,2021,35(2):160-163.
- [2] Awan SN, Roy N, Zhang D, et al. Validation of the Cepstral Spectral Index of Dysphonia (CSID) as a Screening Tool for Voice Disorders: Development of Clinical Cutoff Scores[J]. J Voice, 2016, 30(2): 130-144.
- [3] Lechien JR, Huet K, Finck C, et al. Are the Acoustic Measurements Reliable in the Assessment of Voice Quality? A Methodological Prospective Study[J]. J Voice, 2021, 35(2): 203-215.
- [4] Heman-Ackah YD, Sataloff RT, Laureyns G, et al. Quantifying the cepstral peak prominence, a measure of dysphonia[J]. J Voice, 2014, 28(6): 783-788.
- [5] Brockmann M, Drinnan MJ, Storck C, et al. Reliable jitter and shimmer measurements in voice clinics: the relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task[J]. J Voice, 2011, 25(1): 44-53.
- [6] Aghaei F, Khoramshahi H, Zamani P, et al. A Cepstral Peak Prominence(CPP) Voice Analysis in Iranian Post-lingual Deaf Adult Cochlear Implant Users[J]. J Voice, 2021.
- [7] Solomon NP, Awan SN, Helou LB, et al. Acoustic analyses of thyroidectomy-related changes in vowel phonation[J]. J Voice, 2012, 26(6): 711-720.
- [8] Awan SN, Roy N. Acoustic prediction of voice type in women with functional dysphonia[J]. J Voice, 2005, 19(2): 268-282.
- [9] Awan SN, Awan JA. A Two-Stage Cepstral Analysis Procedure for the Classification of Rough Voices[J]. J Voice, 2020, 34(1): 9-19.
- [10] Barsties V, Latoszek B, Ulozaitė-Stanienė N, Maryn Y, et al. The Influence of Gender and Age on the Acoustic Voice Quality Index and Dysphonia Severity Index: A Normative Study[J]. J Voice, 2019, 33(3): 340-345.
- [11] Awan SN, Giovinco A, Owens J. Effects of vocal intensity and vowel type on cepstral analysis of voice[J]. J Voice, 2012, 26(5): 670. e15-20.
- [12] 余明强,周莉,徐新林,等. 声带息肉患者持续元音及连贯言语声的倒频谱声学分析[J]. 听力学及言语疾病杂志,2016,24(1):15-19.
- [13] Demirci AN, Köse A, Aydinli FE, et al. Investigating the cepstral acoustic characteristics of voice in healthy children[J]. Int J Pediatr Otorhinolaryngol, 2021, 148: 110815.
- [14] Franca MC. Acoustic comparison of vowel sounds among adult females[J]. J Voice, 2012, 26(5): 671. e9-17.
- [15] Sauder C, Bretl M, Eadie T. Predicting Voice Disorder Status From Smoothed Measures of Cepstral Peak Prominence Using Praat and Analysis of Dysphonia in Speech and Voice(ADSV)[J]. J Voice, 2017, 31(5): 557-566.
- [16] Peterson EA, Roy N, Awan SN, et al. Toward validation of the cepstral spectral index of dysphonia(CSID) as an objective treatment outcomes measure[J]. J Voice, 2013, 27(4): 401-410.
- [17] Lee Y, Kim G, Sohn K, et al. The Usefulness of Auditory Perceptual Assessment and Acoustic Analysis as a Screening Test for Voice Problems[J]. Folia Phoniatr Logop, 2021, 73(1): 34-41.
- [18] Lee JM, Roy N, Peterson E, et al. Comparison of Two Multiparameter Acoustic Indices of Dysphonia Severity: The Acoustic Voice Quality Index and Cepstral Spectral Index of Dysphonia[J]. J Voice, 2018, 32(4): 515. e1-515. e13.
- [19] Naunheim MR, Garneau J, Park C, et al. Voice Outcomes After Radiation for Early-Stage Laryngeal Cancer[J]. J Voice, 2020, 34(3): 460-464.

(收稿日期:2022-04-06)