



# Acoustic parameters for the evaluation of voice quality in patients with voice disorders

Gelin Li<sup>1</sup>, Qian Hou<sup>1</sup>, Chi Zhang<sup>1</sup>, Zhen Jiang<sup>2</sup>, Shusheng Gong<sup>1</sup>

<sup>1</sup>ENT, Beijing Friendship Hospital affiliated to Capital Medical University, Beijing, China; <sup>2</sup>Voice Research Center of Central Conservatory of Music, Beijing, China

**Contributions:** (I) Conception and design: G Li, S Gong; (II) Administrative support: S Gong; (III) Provision of study materials or patients: Z Jiang, G Li; (IV) Collection and assembly of data: G Li; (V) Data analysis and interpretation: C Zhang, G Li; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

**Correspondence to:** Shusheng Gong. Beijing Friendship Hospital affiliated to Capital Medical University, No. 95 Yongan Road, Xicheng District, Beijing 100050, China. Email: gongss@ccmu.edu.cn.

**Background:** To investigate the value of the standard deviation of the fundamental frequency ( $F_0$  SD), jitter, and shimmer for the evaluation of voice quality and the description of vocal characteristics in patients with voice disorders.

**Methods:** This prospective cohort study included 4 groups: vocal cord polyps (VCP) group (n=55), early-stage (stage I–II) laryngeal carcinoma (ELC) group (n=35), mutational falsetto (MF) group (n=17), and a normal control group (n=29). The participants were asked to emit a sustained vowel /a/ and raise the pitch gradually. Acoustic parameters, including  $F_0$ ,  $F_0$  SD, jitter, and shimmer, were recorded and analyzed.

**Results:** The  $F_0$  SD was highest in the MF group. The  $F_0$  SD in the MF group and the ELC group was significantly higher than that in the VCP group and the control group ( $P < 0.05$ ), and the  $F_0$  SD in the VCP group was significantly higher than that in the control group ( $P < 0.05$ ). However, there was no significant difference in  $F_0$  SD between the MF group and the ELC group ( $P > 0.05$ ). The jitter and shimmer in the ELC group were significantly higher compared to the other groups ( $P < 0.05$ ), and the jitter and shimmer in the VCP group were significantly higher than those in the MF group and the control group ( $P < 0.05$ ). There were no significant differences in jitter or shimmer between the MF group and the control group ( $P > 0.05$ ).

**Conclusions:**  $F_0$  SD, jitter, and shimmer are important parameters for the evaluation of pitch variation during sustained phonations, and can discriminate between MF and voice disorders.

**Keywords:** Acoustic analysis; standard deviation of fundamental frequency; jitter; shimmer; voice quality; voice disorder

Submitted Sep 17, 2020. Accepted for publication Nov 18, 2020.

doi: 10.21037/apm-20-2102

**View this article at:** <http://dx.doi.org/10.21037/apm-20-2102>

## Introduction

Acoustic analysis is an objective, noninvasive modality for the evaluation of voice quality in patients with voice disorders such as laryngitis, laryngospasm, laryngeal tumors, spasmodic dysphonia, and vocal cord paralysis (1–5). The widely used parameters of acoustic analysis mainly include the standard deviation of the fundamental frequency ( $F_0$  SD), jitter, and shimmer (3,6). Jitter refers to the short-term variations in the  $F_0$  between contiguous glottal cycles,

and shimmer represents the short-term variations in the amplitude of sound waves (6).

According to the existing literature, these acoustic parameters have been extensively used for the identification of voice abnormalities. In 2016, Lopes and colleagues investigated the accuracy of acoustic parameters to discriminate between patients with different laryngeal diagnoses. They found that isolated  $F_0$  SD was the optimal parameter for distinguishing between vocal nodules

and unilateral vocal fold paralysis, vocal nodules and gastroesophageal reflux-induced voice impairment, and between a vocal polyp and *sulcus vocalis*. Shimmer exhibited high accuracy for the differential diagnosis between vocal nodules and *sulcus vocalis*, and a combination of  $F_0$  SD and jitter aided in the identification of unilateral vocal fold paralysis (6). Subsequently, based on a much larger cohort, they proposed that combined acoustic measurements can facilitate the discrimination of voice deviation intensity and predominant voice quality in patients with dysphonia (7). Ayoub *et al.* investigated the impact of smoking on voice acoustics, and demonstrated that  $F_0$  SD and jitter did not alter remarkably, although the mean  $F_0$  and speaking  $F_0$  were significantly reduced in cigarette smokers (8). Searl *et al.* compared the acoustic parameters before and after executing treatment tasks in patients with Parkinson's disease, and they noted that voice intensity was significantly increased after treatment, while there was no significant alteration in  $F_0$  SD (9). Additionally, Kang *et al.* found that some acoustic parameters (jitter, relative average perturbation, and noise-to-harmonic ratio) were associated with aspiration risk in patients with swallowing disorders (10). Jesus *et al.* found that unilateral vocal fold paralysis may cause significant alterations in various acoustic parameters including mean  $F_0$ ,  $F_0$  SD, jitter, shimmer, and mean harmonics-to-noise ratio (11).

Although these acoustic parameters have shown promise for describing voice characteristics across a number of conditions, there is still a paucity of evidence comparing the clinical values of  $F_0$  SD and perturbation parameters (jitter and shimmer). Previous studies indicated that  $F_0$  SD may have a higher sensitivity for an objective clinical voice assessment than jitter and shimmer, as a rapidly changing or shifting  $F_0$  may not alter the perturbation parameters (12), though there remains a lack of solid evidence. Therefore, this study aimed to compare the value of  $F_0$  SD, jitter, and shimmer for the evaluation of voice quality and the description of vocal characteristics in patients with voice disorders.

We present the following article in accordance with the STROBE reporting checklist (available at <http://dx.doi.org/10.21037/apm-20-2102>).

## Methods

### Subjects

All procedures performed in this study involving human participants were in accordance with the Declaration of

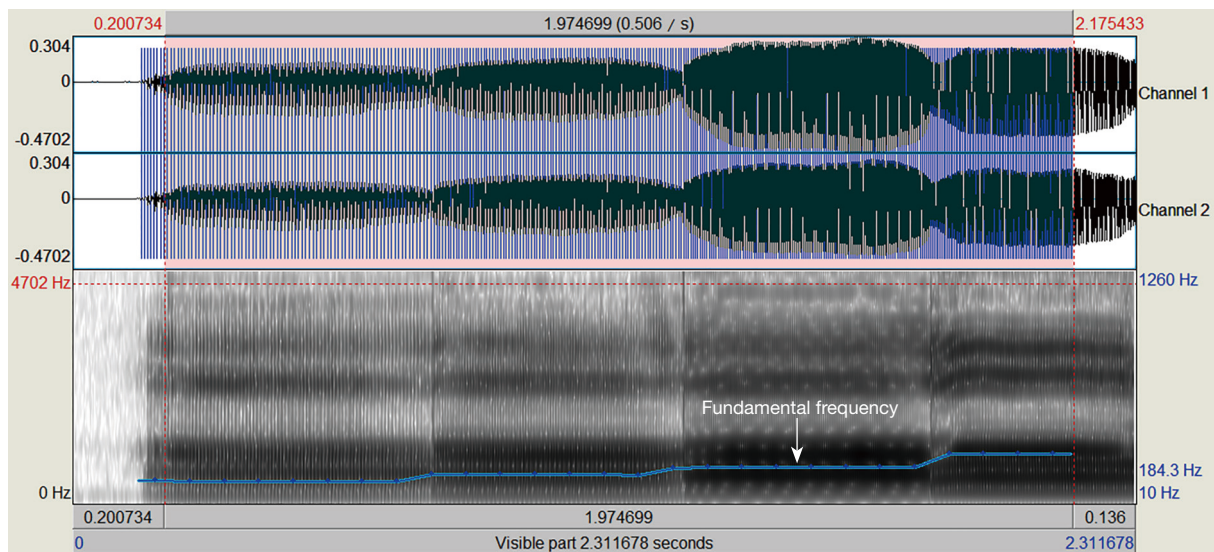
Helsinki (as revised in 2013). This study was approved by the local Ethics Committee (No. 2020-P2-074-01). Written informed consent was obtained from each participant. All the participants were male and they were enrolled from our hospital between January 2017 and December 2018. The inclusion criteria were as follows: (I) aged 18–60 years; and (II) no history of acute respiratory tract infection in the past 2 weeks. This study included 4 groups. For the vocal cord polyps (VCP) group, including 55 consecutive patients with VCP: electronic laryngoscope or stroboscopic laryngoscope demonstrated unilateral localized polyps at the edge of the vocal cord, with the base length less than 1/3 of the length of the vocal cord membranous part. For the early-stage laryngeal carcinoma (ELC) group, including 35 consecutive patients with early-stage (stage I–II) ELC: electronic laryngoscopy, stroboscopic laryngoscope, computed tomography or magnetic resonance imaging showed a tumor confined to unilateral or bilateral vocal cords, with mobility of the vocal cords. Postoperative pathology confirmed a diagnosis of squamous cell carcinoma without invasion of the vocal cord muscle layer. For the mutational falsetto (MF) group, including 17 consecutive patients with MF: the falsetto occurred during puberty, with high or unstable pitch and low volume. Stroboscopic laryngoscope showed glottal dysraphism, and falsetto was the major voice mode. In addition to traditional  $F_0$  parameters,  $F_0$  SD of mutational falsetto was significantly higher than that of normal. In addition to high pitch, another characteristic is that the unstable state of the fundamental frequency is a very important feature. The value of  $F_0$  SD can reflect the overall unstable voice characteristics of the fundamental frequency. We used the combination of  $F_0$  and  $F_0$  SD to identify mutational falsetto.

In addition, 29 healthy medical staff with no history of dysphonia or hoarseness were recruited as normal controls.

### Acoustic examinations

Acoustic examinations were performed in a quiet room with a sound field less than 45 dB, and vocal data were collected using the ZOOMH6 recorder (sampling frequency, 44,000 Hz; 16 bits). The microphone was positioned 15–20 cm away from the subject's mouth. The recorded sample was analyzed using the PRAAT software (version 3.9).

The participants were asked to emit a sustained vowel /a/, starting with a steady pronunciation for 0.5–1 seconds. Subsequently, the participants were asked to abruptly raise the pitch and keep it stable for 0.5–1 seconds, then abruptly



**Figure 1** Phonogram of a sustained vowel /a/ in a healthy control. The phonogram showed continuously increasing fundamental frequency (arrow).

raise the pitch again and maintain stable pronunciation until the pitch reached the upper limit of the natural vocal range. The shifting sections were defined as 4 continuous sections, starting from the comfort pitch and characterized by continuously increasing  $F_0$ , in which the voice signal was stable (Figure 1). Acoustic analyses were performed in the first-section stable pronunciation (vocal sample I), first 2 shifting sections (vocal sample II), and 4 continuous shifting sections (vocal sample III), respectively. Acoustic parameters, including  $F_0$ ,  $F_0$  SD, jitter, and shimmer, were recorded.

### Statistical analysis

SPSS 24.0 software (IBM Corp., Armonk, NY, USA) was used for statistical analyses. Continuous variables were expressed as the mean  $\pm$  standard deviation (SD) when data was normally distributed, or median when non-normally distributed. For univariate statistical analysis, Mauchly's test of sphericity or the Mann-Whitney U-test for continuous variables were used as appropriate. Probability (P) values  $<0.05$  were considered significant.

## Results

### Vocal samples in normal controls

The  $F_0$  ( $202.0 \pm 38.6$ ) and  $F_0$  SD ( $54.15 \pm 26.65$ ) in the voice

sample III was significantly higher than those in the voice sample II ( $P=0.000$ ;  $F_0$ ,  $145.4 \pm 14.1$ ;  $F_0$  SD,  $18.28 \pm 9.97$ ), and these parameters were also significantly higher than those in the voice sample I ( $P=0.000$ ;  $F_0$ ,  $128.2 \pm 13.9$ ;  $F_0$  SD,  $1.66 \pm 1.06$ ). There were no significant differences in jitter between any 2 groups (all  $P>0.05$ ). The shimmer in the voice sample III ( $2.76 \pm 1.11$ ) was significantly lower than that in the voice sample II ( $3.35 \pm 1.17$ ) ( $P=0.005$ ), while there was no significant difference in shimmer between the voice sample II and the voice sample I ( $P>0.05$ ). The detailed data are summarized in Table 1.

### Acoustic parameters in voice disorder cases and controls

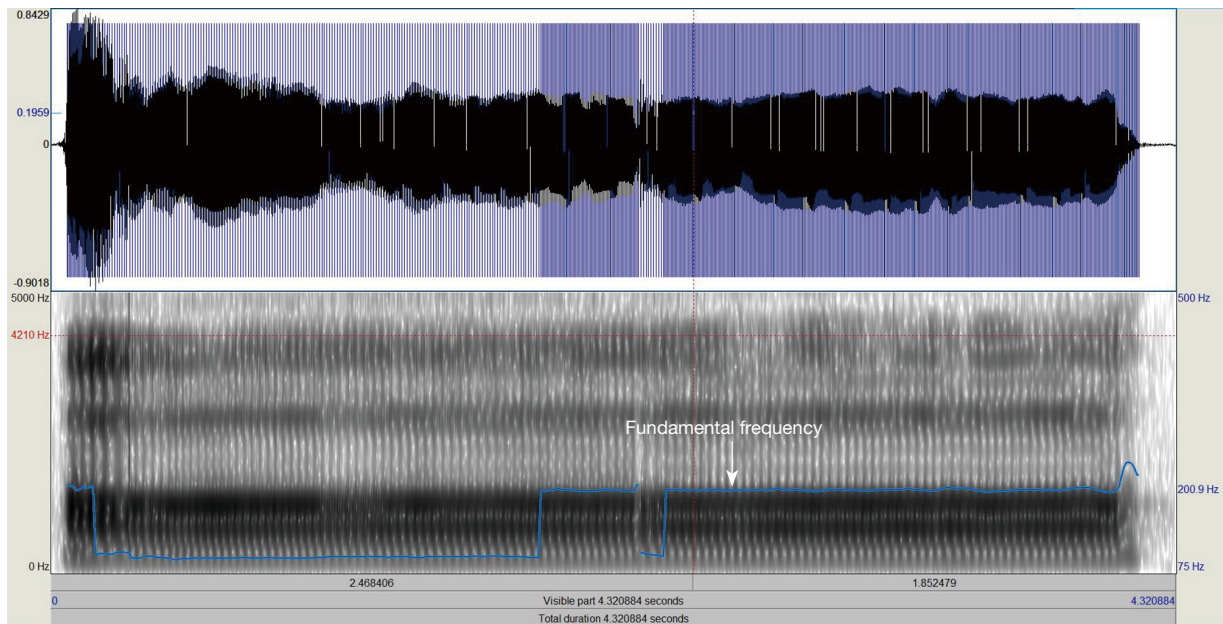
The  $F_0$  SD was highest (median, 4.00) in the MF group, and the phonogram showed abrupt alterations in  $F_0$  (Figure 2). The  $F_0$  SD in the MF group (median, 4.00) and the ELC group (median, 3.85) was significantly higher than that in the VCP group (median, 2.20) and the control group (median, 1.17) ( $P<0.05$ ), and the  $F_0$  SD in the VCP group was significantly higher than that in the control group ( $P<0.05$ ). However, there was no significant difference in  $F_0$  SD between the MF group and the ELC group ( $P>0.05$ ).

The jitter (median, 0.91) and shimmer (median, 8.49) in the ELC group were significantly higher compared to the other groups ( $P<0.05$ ), and the jitter (median, 0.59) and shimmer (median, 5.27) in the VCP group were

**Table 1** Fundamental frequency ( $F_0$ ), standard deviation of the fundamental frequency ( $F_0$  SD), jitter, and shimmer in voice samples

Voice sample	$F_0$ (Hz)	$F_0$ SD (Hz)	Jitter (%)	Shimmer (dB)
Voice sample I	128.2±13.9	1.66±1.06	0.482±0.27	3.44±1.74
Voice sample II	145.4±14.1 <sup>a</sup>	18.28±9.97 <sup>a</sup>	0.527±0.22	3.35±1.17
Voice sample III	202.0±38.6 <sup>ab</sup>	54.15±26.65 <sup>ab</sup>	0.487±0.20	2.76±1.11 <sup>ab</sup>

<sup>a</sup>, statistical significance when compared with the voice sample I; <sup>b</sup>, statistical significance when compared with the voice sample II.

**Figure 2** Phonogram of a sustained vowel/a/ in a prepubertal falsetto representative case. The phonogram showed abrupt alterations in fundamental frequency (arrow).**Table 2** Acoustic parameters in voice disorder cases and controls

Groups	N	$F_0$ SD (Hz)	Jitter (%)	Shimmer (dB)
Normal control group	29	1.17	0.33	2.62
Early-stage laryngeal carcinoma (ELC) group	35	3.85 <sup>ac</sup>	0.91 <sup>acd</sup>	8.49 <sup>acd</sup>
Vocal cord polyps (VCP) group	55	2.20 <sup>abd</sup>	0.59 <sup>abd</sup>	5.27 <sup>abd</sup>
Mutational falsetto (MF) group	17	4.00 <sup>ac</sup>	0.53 <sup>ab</sup>	3.56 <sup>ab</sup>

$F_0$  SD, standard deviation of the fundamental frequency. <sup>a</sup>, statistical significance when compared with the control group; <sup>b</sup>, statistical significance when compared with the ELC group; <sup>c</sup>, statistical significance when compared with the VCP group; <sup>d</sup>, statistical significance when compared with the MF group.

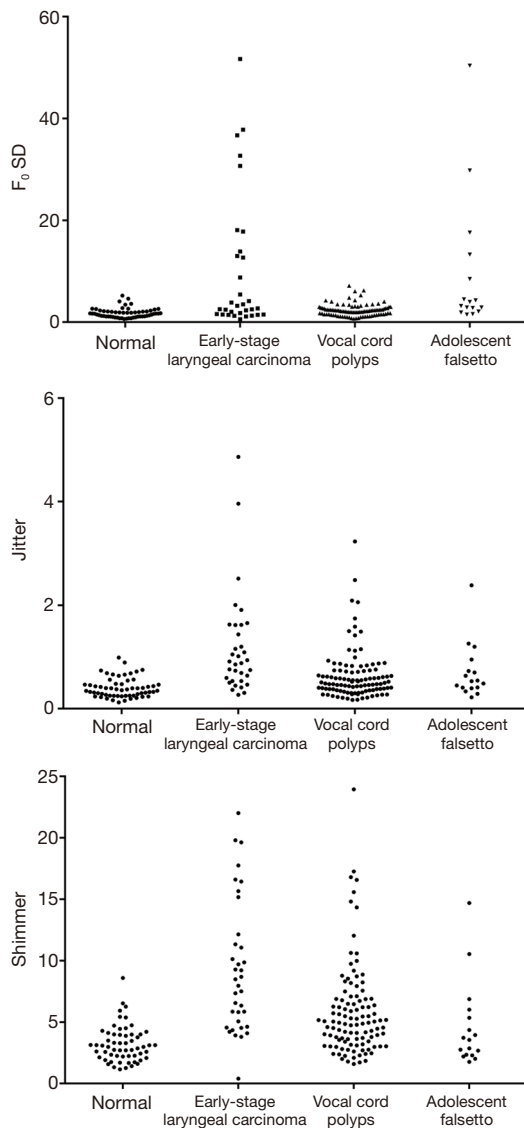
significantly higher than those in the MF group (median jitter, 0.53; median shimmer, 3.56) and the control group (median jitter, 0.33; median shimmer, 2.62) ( $P < 0.05$ ). There were no significant differences in jitter or shimmer between the MF group and the control group ( $P > 0.05$ ). The detailed

results are summarized in *Table 2* and *Figure 3*.

## Discussion

MF and voice disorders all lead to pitch changes, though the





**Figure 3** Acoustic parameters in voice disorder cases and controls.

clinical differentiation may be challenging. Voice disorders are usually characterized by phonatory instability, and acoustic analysis provides objective measures of phonatory characteristics (13). Perturbation parameters (jitter and shimmer), reflecting the perturbations in the frequency of the neighboring vibration cycles, are sensitive to phonatory deviations (14-16). In contrast to jitter and shimmer, which are perceived as voice roughness,  $F_0$  SD represents the  $F_0$  variation that is captured by measuring the SD in voice pitch. Therefore,  $F_0$  SD captures the amount of within-utterance variation in pitch, and low values of  $F_0$  SD are perceived as monotony (17).

Lim *et al.* investigated the acoustic and electroglottographic features in patients with dysphonia before and after vocal treatment, and the authors found that pitch lowered and improved in quality after treatment (18). In another study by the same authors, they found that acoustic and electroglottographic parameters, including jitter, the harmonics-noise ratio, the mean closed quotient, and the irregularity of the frequency, facilitated the objective assessment of the severity of edema and voice quality before and after surgery in patients with Reinke's edema (19). Additionally, as cleft palate can severely affect the structure and function of the vocal tract and thus impair voice quality, some researchers attempted to perform acoustic analysis of voice in children with cleft palate and velopharyngeal insufficiency. Villafuerte-Gonzalez *et al.* reported that the  $F_0$  was significantly higher in children with cleft palate when compared to normal controls, and children with velopharyngeal insufficiency had a significantly higher shimmer perturbation (20). However, in a recent study, Segura-Hernandez and colleagues also conducted acoustic analyses in children with cleft lip and palate and velopharyngeal insufficiency, and they found no significant difference in mean  $F_0$  between controls and patients. Moreover, at the onset of the treatment, jitter and shimmer were significantly increased in all patients, while at the end of the treatment, jitter and shimmer were markedly decreased (21). These findings indicate perturbation parameters are more sensitive than the mean  $F_0$ . Notably, all the above studies did not investigate  $F_0$  SD. In the current study, our findings suggest  $F_0$  SD may be a more reliable index than jitter and shimmer for the evaluation of voice quality.

MF, also known as puberphonia, is a functional voice disturbance characterized by failure of the male high-pitched preadolescent voice to transition to the lower pitch of adolescence and adulthood. Previous studies have confirmed that altered personality, neuropsychological, and social factors may contribute to the occurrence of MF, though the definitive pathological mechanism underlying the development of this disorder is not fully understood. Preliminary evidence indicates that irregular conversions of real pitch and falsetto are attributed to unstable vocal control (22). There has been no systematic acoustic analysis in patients with MF. Dagli *et al.* evaluated the outcomes of voice treatment on MF by using perceptual and acoustic analysis, and they found the  $F_0$ , jitter, and shimmer were all suppressed after treatment (23). However, the  $F_0$  SD was not included in the observational indices. In the current study, we found that there was no significant difference in jitter or

shimmer between patients with MF and controls when they were asked to emit a sustained vowel/a/, while the  $F_0$  SD showed marked alterations. Further analyses demonstrated that  $F_0$  SD, jitter, and shimmer were different between the laryngeal carcinoma group and the control group, as well as between the VCP group and controls. These findings indicate that MF is distinct from organic laryngopharyngeal diseases, and the disparate pathophysiologic processes lead to different acoustic characteristics.

Hohm *et al.* proposed that abrupt changeovers of  $F_0$  in a sustained vowel/a/ did not alter the jitter and shimmer (12), which is consistent with the results of the present study. Although the participants were asked to abruptly raise the pitch repeatedly, there was no significant difference in jitter or shimmer between different voice samples. However, the  $F_0$  SD exhibited obvious alterations with the pitch changeovers, and the variation of  $F_0$  SD was positively correlated with the degree of pitch hopping. These findings indicate that  $F_0$  SD is more sensitive to frequency hopping than jitter and shimmer, which is consistent with the acoustic results of MF.

The significant variation of  $F_0$  SD in patients with MF suggests that voice frequency is unstable and the ability to control vocal cords is weak. Moreover, the unremarkable alterations in jitter and shimmer indicate that asymmetrical vibration of the vocal cords may be involved in MF, while further studies using laryngeal high-speed photography are still warranted. In summary, acoustic analysis has the advantages of noninvasive, low price and convenience. It provides objective data for the assessment of voice disorders and becomes an indispensable detection method for voice diseases and voice disorders. In addition, it can be used to evaluate the curative effect of vocal cord polyps before and after surgery.

## Conclusions

$F_0$  SD, jitter, and shimmer are important parameters for the evaluation of pitch variation, which is of great significance in the diagnosis of MF and voice disorders.

## Acknowledgments

*Funding:* None.

## Footnote

*Reporting Checklist:* The authors have completed the

STROBE reporting checklist. Available at <http://dx.doi.org/10.21037/apm-20-2102>

*Data Sharing Statement:* Available at <http://dx.doi.org/10.21037/apm-20-2102>

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/apm-20-2102>). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All procedures performed in this study involving human participants were in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the local Ethics Committee (No. 2020-P2-074-01). Written informed consent was obtained from each participant.

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

1. Gorris C, Ricci Maccarini A, Vanoni F, et al. Acoustic Analysis of Normal Voice Patterns in Italian Adults by Using Praat. *J Voice* 2020;34:961.e9-961.e18.
2. Hanson DG, Jiang JJ, Chen J, et al. Acoustic measurement of change in voice quality with treatment for chronic posterior laryngitis. *Ann Otol Rhinol Laryngol* 1997;106:279-85.
3. Lee SH, Hong KH, Kim JS, et al. Perceptual and Acoustic Outcomes of Early-Stage Glottic Cancer After Laser Surgery or Radiotherapy: A Meta-Analysis. *Clin Exp Otorhinolaryngol* 2019;12:241-8.
4. Lu D, Chen F, Yang H, et al. Changes after voice therapy in acoustic voice analysis of Chinese patients with voice disorders. *J Voice* 2018;32:386.e1-386.e9.
5. Ramírez DAM, Jiménez VMV, López XH, et al. Acoustic

- analysis of voice and electroglottography in patients with laryngopharyngeal reflux. *J Voice* 2018;32:281-4.
6. Lopes LW, Batista Simoes L, Delfino da Silva J, et al. Accuracy of Acoustic Analysis Measurements in the Evaluation of Patients With Different Laryngeal Diagnoses. *J Voice* 2017;31:382.e15-382.e26.
  7. Lopes LW, Alves JDN, Evangelista DDS, et al. Accuracy of traditional and formant acoustic measurements in the evaluation of vocal quality. *Codas* 2018;30:e20170282.
  8. Ayoub MR, Larrouy-Maestri P, Morsomme D. The Effect of Smoking on the Fundamental Frequency of the Speaking Voice. *J Voice* 2019;33:802.e11-802.e16.
  9. Searl J, Wilson K, Haring K, et al. Feasibility of group voice therapy for individuals with Parkinson's disease. *J Commun Disord* 2011;44:719-32.
  10. Kang YA, Kim J, Jee SJ, et al. Detection of voice changes due to aspiration via acoustic voice analysis. *Auris Nasus Larynx* 2018;45:801-6.
  11. Jesus LM, Martinez J, Hall A, et al. Acoustic Correlates of Compensatory Adjustments to the Glottic and Supraglottic Structures in Patients with Unilateral Vocal Fold Paralysis. *Biomed Res Int* 2015;2015:704121.
  12. Hohm J, Dollinger M, Bohr C, et al. Influence of F0 and Sequence Length of Audio and Electroglottographic Signals on Perturbation Measures for Voice Assessment. *J Voice* 2015;29:517.e11-517.e21.
  13. Morris AE, Norris SA, Perlmutter JS, et al. Quantitative, clinically relevant acoustic measurements of focal embouchure dystonia. *Mov Disord* 2018;33:449-58.
  14. Hosseinfar S, Torabinezhad F, Ghelichi L, et al. How Do Voice Perceptual Changes Predict Acoustic Parameters in Persian Voice Patients? *J Voice* 2018;32:705-9.
  15. Torabinezhad F, Izadi F, Pourshahbaz A, et al. Acoustic Parameters in Persian-Speaking Patients with Dysphonia. *Function and Disability Journal* 2018;1:8-17.
  16. Zealouk O, Satori H, Hamidi M, et al. Vocal parameters analysis of smoker using Amazigh language. *International Journal of Speech Technology* 2018;21:85-91.
  17. Knowles KK, Little AC. Vocal fundamental and formant frequencies affect perceptions of speaker cooperativeness. *Q J Exp Psychol (Hove)* 2016;69:1657-75.
  18. Lim JY, Lim SE, Choi SH, et al. Clinical characteristics and voice analysis of patients with mutational dysphonia: clinical significance of diplophonia and closed quotients. *J Voice* 2007;21:12-9.
  19. Lim JY, Choi JN, Kim KM, et al. Voice analysis of patients with diverse types of Reinke's edema and clinical use of electroglottographic measurements. *Acta Otolaryngol* 2006;126:62-9.
  20. Villafuerte-Gonzalez R, Valadez-Jimenez VM, Hernandez-Lopez X, et al. Acoustic analysis of voice in children with cleft palate and velopharyngeal insufficiency. *Int J Pediatr Otorhinolaryngol* 2015;79:1073-6.
  21. Segura-Hernandez M, Valadez-Jimenez VM, Ysunza PA, et al. Acoustic analysis of voice in children with cleft lip and palate following vocal rehabilitation. Preliminary report. *Int J Pediatr Otorhinolaryngol* 2019;126:109618.
  22. Hodges-Simeon CR, Gurven M, Puts DA, et al. Vocal fundamental and formant frequencies are honest signals of threat potential in prepubertal males. *Behav Ecol* 2014;25:984-8.
  23. Dagli M, Sati I, Acar A, et al. Mutational falsetto: intervention outcomes in 45 patients. *J Laryngol Otol* 2008;122:277-81.
- (English Language Editor: C. Betlazar-Maseh)

**Cite this article as:** Li G, Hou Q, Zhang C, Jiang Z, Gong S. Acoustic parameters for the evaluation of voice quality in patients with voice disorders. *Ann Palliat Med* 2021;10(1):130-136. doi: 10.21037/apm-20-2102