Otolaryngology–Head and Neck Surgery (2009) 141, 522-526

ORIGINAL RESEARCH–PEDIATRIC OTOLARYNGOLOGY

# Voice outcomes following repeated surgical resection of laryngeal papillomata in children

**Theresa Holler, MD, Jennifer Allegro, MSc, SLP (C), Neil K. Chadha, MBChB (Hons), FRCS, Michael Hawkes, MD, FRCPC, Robert V. Harrison, PhD, Vito Forte, MD, FRCSC,** and **Paolo Campisi, MD, MSc, FRCSC, FAAP,** Toronto, Ontario, Canada

*No sponsorships or competing interests have been disclosed for this article.*

## ABSTRACT

**OBJECTIVES:** 1) To apply perceptual and acoustic voice assessments to children treated for juvenile-onset recurrent respiratory papillomatosis (JORRP); 2) to compare voice outcomes following treatment for JORRP using microdebrider versus carbon dioxide ($CO_2$) laser.

**STUDY DESIGN:** Prospective cohort study.

**SETTING:** This study was conducted at a tertiary pediatric academic center (March 2008-March 2009).

**SUBJECTS AND METHODS:** Children with active JORRP were assessed using perceptual and acoustic voice analysis following treatment with either $CO_2$ laser or microdebrider. Outcome measures included overall severity rating, jitter, shimmer, and noise-to-harmonic ratio (NHR). The unpaired Student *t* test and Pearson correlation tests were used to explore the statistical significance of hypothesis tests.

**RESULTS:** Eleven patients (8 male, 3 female) aged three to 17 years were enrolled. There were six children in the $CO_2$ laser cohort and five children in the microdebrider cohort. The immediate postoperative scores were significantly lower in the microdebrider cohort (vs the $CO_2$ cohort) for jitter, shimmer, NHR, and perceptual scores ($P < 0.05$), indicating a better voice quality in the microdebrider group. Jitter, shimmer, and NHR showed a significant positive correlation with the proportion of $CO_2$ laser procedures ($P < 0.05$).

**CONCLUSION:** This is the first study to use perceptual and objective acoustic evaluations to compare voice outcomes following microdebrider or $CO_2$ laser treatment of JORRP. The results of this study suggest that treatment with the microdebrider results in a better immediate and early postoperative voice outcome. Moreover, the data demonstrate a correlation of worsening voice quality with increased exposure to the $CO_2$ laser.



An often overlooked consequence of human papillomavirus (HPV) infection is the transmission of virus from mother to child, presumably at the time of birth. This transmission may present clinically in the first years of life in the form of juvenile-onset recurrent respiratory papillomatosis (JORRP). JORRP is an uncommon condition, with a reported prevalence of 1.7 to 2.6 per 100,000 children in the United States.[1] It manifests as recurrent, wart-like growths that can occur anywhere along the respiratory tract, but it most commonly involves the laryngeal mucosa, resulting in dysphonia and airway compromise. JORRP causes marked physical and emotional suffering for affected patients because of the need for repeated surgical resection of the airway lesions.

The results of a 2004 Web-based survey indicated that 52.7 percent of American Society of Pediatric Otolaryngology members preferred using a microdebrider to resect HPV-induced laryngeal papillomata.[2] The survey also revealed that 41.9 percent of members preferred the carbon dioxide ($CO_2$) laser, and the remaining 5.4 percent used a combination of both modalities. Although there have been many arguments proposed in favor of one treatment modality over the other with respect to issues such as surgical times, postoperative pain, and safety, there are minimal data in the literature comparing voice outcomes. This is surprising given that poor voice quality is a major morbidity associated with repeated surgical debulking.

Several animal studies have demonstrated significant deep thermal injury to vocal folds treated with the $CO_2$ laser, resulting in delayed healing, necrosis, and scarring.[3-7] However, evidence in support of the microdebrider, from a voice perspective, is limited and derived from anecdote[8,9] or studies with a small sample size and subjective questionnaires.[10]

The aim of this study was to compare voice outcomes between JORRP children treated using the $CO_2$ laser versus those treated using the microdebrider. We hypothesized that

Received March 26, 2009; revised May 14, 2009; accepted June 17, 2009.

0194-5998/$36.00 
doi:10.1016/j.otohns.2009.06.080

the microdebrider provides a more favorable voice outcome than the $CO_2$ laser in the treatment of recurrent laryngeal papillomas. To test our hypothesis, a prospective cohort study was performed to evaluate the voice quality of children with JORRP preoperatively and postoperatively, using a standardized perceptual rating scale and computer-assisted acoustic analysis.

## METHODS

The study was approved by the Research Ethics Board of the Hospital for Sick Children, Toronto, Canada. Participation in the study was offered to all children (3 to 18 years) undergoing active treatment for JORRP. The only exclusion criterion was an inability to complete the voicing tasks, usually due to young age. The patients were divided into two cohorts depending on the technique used for surgical papilloma debulking ($CO_2$ laser vs microdebrider). The patients received treatment by the modality of choice of their responsible surgeon.

The study was conducted prospectively. Enrollment in the study was initiated during clinic visits when the decision was made to proceed with a surgical intervention. Voice analyses were completed at three time points: 1) on the day of enrollment ("pre-op"); 2) four to six hours postoperatively ("immediate post-op"); and 3) one week postoperatively ("early post-op"). Post-treatment evaluations of voice were limited to the immediate and early post-op periods to avoid confounding by papillomata regrowth. At the time of surgery, the extent of papilloma growth was documented using the Derkay scoring system.[11]

All voice samples were analyzed perceptually and by computerized acoustic analysis. Voice samples were collected using the Multi-Dimensional Voice Program and Real Time Pitch program of the Computerized Speech Laboratory (CSL) Model 4500.[12] Standardized procedures were used for all recordings. Patients were seated in a quiet environment. A microphone was held in an off-axis position and at a constant mouth-to-microphone fixed distance of 10 cm.

Perceptual voice analyses were performed using the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) rating scale.[13] Perceptual analyses were completed preoperatively and postoperatively on recorded voice samples and were conducted by a speech-language pathologist trained in the management of pediatric voice disorders. The speech-language pathologist was blinded to the treatment modality. The CAPE-V is an instrument used to describe the severity of perceptual attributes of voice disorders. The instrument consists of six main attributes: overall severity, roughness, breathiness, strain, pitch, and loudness. The attributes are rated using a visual analog scale by placing a mark along a 100-mm line; each attribute is then scored by measuring the distance to this mark along the 100-mm line. Descriptive labels can also be used to generally describe the deviant voice attribute (eg, mild, moderate, severe). Consistency of the perceptual attribute can also be categorized as "consistent" or "intermittent." The voice stimuli used for the CAPE-V include sustained vowel production, six sentences, and conversational speech. The sentence stimuli elicit various laryngeal behaviors, such as hard glottal attack, nasal phonemes, voiced phonemes, and voiceless phonemes. Sentences and spontaneous speech were recorded using the Real Time Pitch program of the CSL.

Acoustic voice analysis was performed using the Multi-Dimensional Voice Program (MDVP) of the CSL.[12] Patients were instructed to produce the vowel /a/ for three seconds using a comfortable pitch and volume. Three repetitions of a sustained /a/ vowel production were collected and their average was used for analysis. The phoneme /a/ was chosen because it is a steady-state vowel and a relatively easy phoneme for children to reproduce. It also allows for extraction of frequency and amplitude measures, which were compared to previously published pediatric norms.[14]

The MDVP program can extract up to 33 acoustic variables from each voice sample. For the present study, five parameters were selected for analysis: short-term frequency variation ("jitter"), long-term frequency perturbation ("vFO"), short-term amplitude perturbation (shimmer), long-term amplitude perturbation ("vAM"), and noise-to-harmonic ratio ("NHR"). The parameters of jitter, vFO, shimmer, and vAM are all measures of variability in the frequency or amplitude in the voice signal. The noise-to-harmonic ratio is a general measure of the noise present in the analyzed voice signal. Short- and long-term variation measures and signal-to-noise ratios are common parameters used for evaluation of deviant voice qualities.[15,16]

The sample population included children who had previously undergone a significant proportion of their surgeries by $CO_2$ laser debulking, exclusively microdebrider debulking, or a mixture of the two modalities. The children were defined as being in a $CO_2$ cohort if they had undergone more than 25 percent of their procedures with the $CO_2$ laser, and in the microdebrider cohort if they had undergone 75 percent or more of their procedures by the microdebrider. This allowed comparison between the two cohorts using the Mann-Whitney $U$ test, with alpha $= 0.05$ for exploring the statistical significance of hypothesis tests. Spearman correlation coefficients were also determined for the proportion of procedures performed by the $CO_2$ laser versus each acoustic voice parameter and perceptual score.

## RESULTS

We recruited 11 children (8 male, 3 female) aged three to 17 years (mean 8.6 years). All children were infected with HPV subtypes 6 or 11. Before study enrollment, the children had undergone between five and 66 (median 13) surgical procedures for debulking of their laryngeal papillomata (Table 1). Acoustic voice parameters and blinded

**Table 1**
**Demographics and surgical history of included children**

| Case | Age (years) | Sex | No. of previous surgical debulking procedures | | |
|---|---|---|---|---|---|
| | | | $CO_2$ laser | Micro-debrider | $CO_2$ laser proportion |
| 1 | 3 | Female | 0 | 10 | 0 |
| 2 | 4 | Male | 0 | 5 | 0 |
| 3 | 7 | Female | 0 | 8 | 0 |
| 4 | 5 | Male | 0 | 6 | 0 |
| 5 | 4 | Female | 0 | 6 | 0 |
| 6 | 8 | Male | 2 | 6 | 25 |
| **7** | **13** | **Male** | **15** | **28** | **30** |
| **8** | **11** | **Male** | **14** | **26** | **31** |
| **9** | **17** | **Male** | **40** | **26** | **61** |
| **10** | **8** | **Male** | **18** | **0** | **100** |
| **11** | **15** | **Male** | **13** | **0** | **100** |

Bold cells represent the $CO_2$ cohort; unbold cells represent the microdebrider cohort.

perceptual voice analysis were obtained for all participants, as per the study protocol. There were no significant differences ($P > 0.05$) in the measured acoustic parameters between the microdebrider and $CO_2$ laser groups at study baseline (preoperatively).

Overall immediate post-op and early post-op acoustic voice parameters were compared, and no statistically significant difference was demonstrated ($P > 0.05$). Therefore, immediate post-op voice outcomes were used for comparisons between the surgical technique cohorts, as these outcomes were collected most consistently, with no missing data for the 11 children. The immediate post-op acoustic voice parameters and perceptual overall severity outcomes were compared between the $CO_2$ laser and microdebrider cohorts (Table 2). The postoperative scores were significantly lower in the microdebrider cohort (compared to the $CO_2$ cohort) for jitter, shimmer, NHR, and perceptual scores ($P < 0.05$), indicating a better voice quality in the microdebrider group. Figures 1, 2, and 3 show scatter plots of the proportion (percentage) of procedures performed by $CO_2$ laser versus jitter, shimmer, and NHR, respectively. All of these parameters showed a significant positive correlation (and therefore increased) with the proportion of $CO_2$ procedures ($P < 0.05$). This finding suggests that voice quality deteriorated with increased exposure to $CO_2$ laser.

**Table 2**
**Comparison of acoustic and perceptual voice parameters between $CO_2$ laser and microdebrider cohorts**

| Voice outcome measure | Immediate post-op score (median) | | $CO_2$ laser cohort vs microdebrider cohort ($P$ value) |
|---|---|---|---|
| | $CO_2$ laser | Micro-debrider | |
| Jitter % | 4.57 | 1.27 | 0.009 |
| Shimmer % | 14.66 | 5.54 | 0.017 |
| vFO % | 8.36 | 7.66 | 0.177 |
| vAM % | 24.99 | 23.40 | 0.662 |
| NHR % | 0.31 | 0.14 | 0.004 |
| CAPE-V "overall severity" perceptual score/100 | 60 | 9 | 0.030 |

Normative values are as follows: jitter = 1.24% ± 0.07 (SEM); shimmer = 3.35% ± 0.12; vFO = 1.75% ± 0.08; vAM = 15.10% ± 0.77; and NHR = 0.11% ± 0.002 (Campisi et al 2002). Normal value for CAPE-V "overall severity" perceptual score is 0.

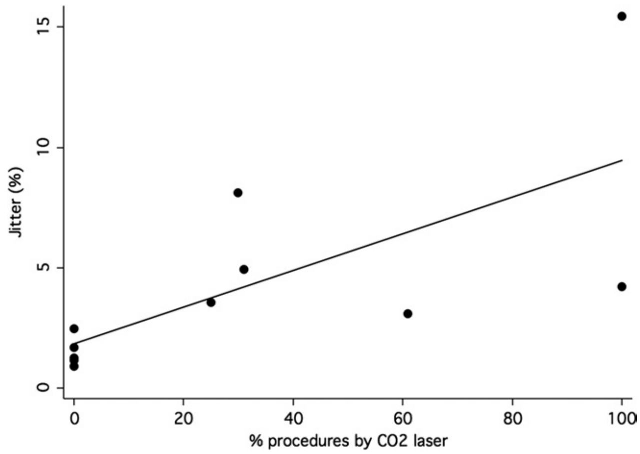


**Figure 1** Immediate post-op jitter acoustic score versus proportion of surgical procedures by $CO_2$ laser. Regression line shown on scatter plot; Spearman correlation coefficient rho = 0.827; $P = 0.002$.

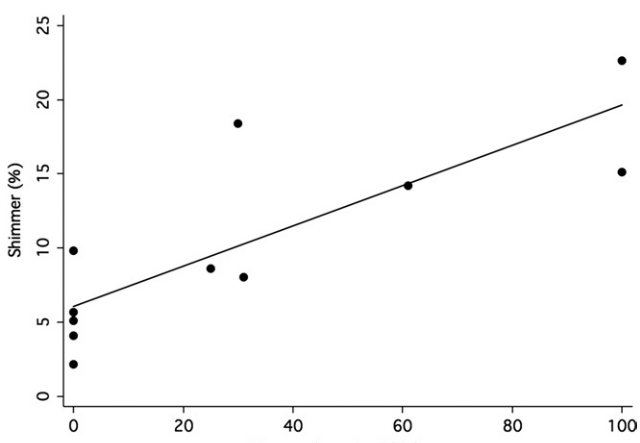


**Figure 2** Immediate post-op shimmer acoustic score versus proportion of surgical procedures by $CO_2$ laser. Regression line shown on scatter plot; Spearman correlation coefficient rho = 0.793; $P = 0.004$.

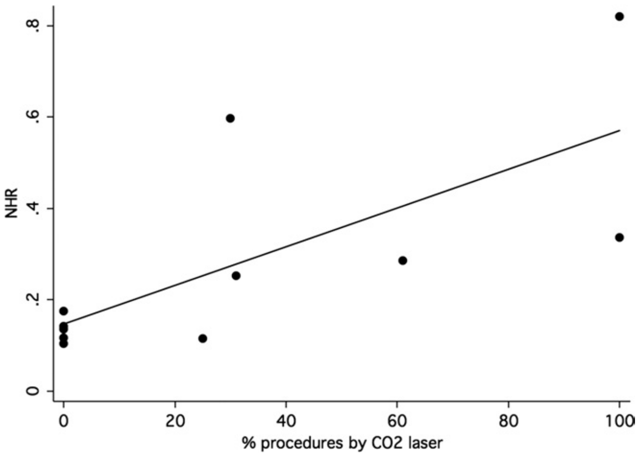


**Figure 3** Immediate post-op NHR acoustic score versus proportion of surgical procedures by $CO_2$ laser. Regression line shown on scatter plot; Spearman correlation coefficient rho = 0.774; $P = 0.005$.

## DISCUSSION

Studies assessing voice outcome in children treated for JORRP with the microdebrider or $CO_2$ laser are limited. Anecdotal parental reports suggest an improved voice quality and faster return to clear phonation following surgery with the microdebrider compared to their previous experience with $CO_2$ laser treatment.[9] In 2003, Pasquale et al[10] randomized 19 patients with JORRP (aged 2-20 years) to treatment with either the $CO_2$ laser or the microdebrider and assessed voice quality using a 10-point Likert scale. When stratified for disease severity, patients treated with the microdebrider self-rated their voice quality superior to subjects in the $CO_2$ laser group. Finally, Parsons and Bothwell[8] demonstrated faster recovery of voice using the microdebrider. The present study represents the first in the literature to use both perceptual and objective measures to compare voice outcomes following microdebrider or $CO_2$ laser treatment of JORRP.

The microdebrider uses suction to pull laryngeal papillomas into the shaft of the instrument where they are resected by a rotating or oscillating blade. When used appropriately, tissue injury is limited to the mucosal layer of the vocal folds. In contrast, the $CO_2$ laser uses light energy that is absorbed and converted to heat energy, causing water molecules within the targeted tissues to vaporize. This process inevitably results in collateral thermal injury to surrounding and deeper tissues.[4-7,17] Injury to the deeper structures may significantly alter the vibratory characteristics of the mucosa and lamina propria of the vocal folds, resulting in permanent voice changes. This may be particularly important in children with JORRP because of the need for repeated treatments.

The perceptual evaluation of voice is a pivotal component of an overall voice assessment. The evaluation of perceptual voice features preoperatively and postoperatively provides salient feedback regarding voice outcome and severity to both the experienced clinician and naïve listener. Recent studies have been undertaken to investigate rater reliability using perceptual protocols, including the CAPE-V.[18,19] Inter-rater reliability has been shown to be strong for the overall severity domain of the CAPE-V.[18] In the design of this study, overall severity rating was therefore selected to be representative of the perceptual assessment. Perceptual voice analysis in this study was performed by one rater; therefore, conclusions about inter-rater reliability were not possible. Patients in the microdebrider cohort had significantly lower postoperative overall severity ratings than those in the $CO_2$ laser cohort.

In contrast to perceptual judgments of voice, acoustic analysis provides an objective quantification of voice characteristics. The CSL and MDVP were used to measure five acoustic parameters: jitter, vFO, shimmer, vAM, and NHR. Jitter, shimmer, and NHR were statistically significantly lower in the microdebrider cohort in the immediate post-op period. Furthermore, the values for jitter, shimmer, and NHR in the microdebrider cohort fell within the 95 percent confidence interval of the normative values.[14] These results support our hypothesis that the microdebrider provides a more favorable voice outcome than the $CO_2$ laser in the treatment of JORRP. The acoustic results were consistent with the perceptual evaluations.

Due to the evolution in surgical practices at our institution and surgeon preferences, a proportion of the study subjects were exposed to both of the surgical modalities. As such, the subjects were stratified according to a predefined threshold of exposure to the $CO_2$ laser. Additionally, due to the rarity of this disease, the number of subjects treated during the study period at this institution was small. These factors have prevented the comparison of cohorts treated exclusively by one modality or the other and have precluded a multivariate analysis to eliminate other possible cofounders such as age, sex, and the number of procedures. To overcome these factors, a multicenter prospective cohort study is required to definitely confirm the findings of this preliminary study.

## CONCLUSIONS

To our knowledge, this is the first study to have used perceptual and objective acoustic evaluations to compare voice outcomes following microdebrider or $CO_2$ laser treatment of JORRP. The results of this study suggest that treatment with the microdebrider results in a better immediate and early postoperative voice outcome. Moreover, the data demonstrate a correlation of worsening voice quality with increased exposure to the $CO_2$ laser.

## AUTHOR INFORMATION

From the Departments of Otolaryngology–Head and Neck Surgery (Drs Holler, Chadha, Harrison, Forte, and Campisi) and Communication Disorders (Dr Allegro); the Division of Infectious Diseases (Dr Hawkes); and the Centre for Paediatric Voice and Laryngeal Function (Drs Allegro and Campisi), Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada.

Corresponding author: P. Campisi, MSc, MD, FRCSC, FAAP, Assistant Professor, Director, Centre for Paediatric Voice and Laryngeal Function, Department of Otolaryngology-Head and Neck Surgery, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada.

E-mail address: paolo.campisi@sickkids.ca.

## AUTHOR CONTRIBUTIONS

**Theresa Holler,** aided in study design and REB approval, primary data collection, assisted in data analysis, manuscript writing and editing; **Jennifer Allegro,** aided in study design, secondary data collection, perceptual data analysis, assisted in manuscript writing and editing; **Neil Chadha,** aided in study design, secondary data collection, statistical analysis, assisted in writing and editing manuscript and preparing figures; **Michael Hawkes,** aided in study design, secondary data collection, statistical analysis, assisted in preparing figures and editing final manuscript; **Robert Harrison,** aided in study design and data collection, guidance for statistical analysis, assisted in preparing and editing final manuscript; **Vito Forte,** aided in study design, primary surgeon and assisted in data collection and analysis, assisted in writing and editing final manuscript; **Paolo Campisi,** study conception, primary surgeon and assisted in data collection and analysis, assisted in writing and editing the final manuscript.

## DISCLOSURES

**Competing interests:** None.

**Sponsorships:** None.

## REFERENCES

1. Reeves WC, Ruparelia SS, Swanson KI, et al. National registry for juvenile-onset recurrent respiratory papillomatosis. Arch Otolaryngol Head Neck Surg 2003;129:976–82.
2. Schraff S, Derkay CS, Burke B, et al. American Society of Pediatric Otolaryngology members' experience with recurrent respiratory papillomatosis and the use of adjuvant therapy. Arch Otolaryngol Head Neck Surg 2004;130:1039–42.
3. Durkin GE, Duncavage JA, Toohill RJ, et al. Wound healing of true vocal cord squamous epithelium after CO2 laser ablation and cut forceps stripping. Otolaryngol Head Neck Surg 1986;95(3 Pt 1):273–7.
4. Speyer M, Joe J, Davidson JM, et al. Thermal injury patterns and tensile strength of canine oral mucosa after carbon dioxide laser incisions. Laryngoscope 1996;106:845–50.
5. Sullivan CA, Rader A, Abdul-Karim FW, et al. Dose-related tissue effects of the $CO_2$ and noncontact Nd:YAG lasers in the canine glottis. Laryngoscope 1998;108:1284–90.
6. Fortune DS, Huang S, Soto J, et al. Effect of pulse duration on wound healing using a $CO_2$ laser. Laryngoscope 1998;108:843–8.
7. Torkain BA, Guo S, Jahng AW, et al. Noninvasive measurement of ablation crater size and thermal injury after CO2 laser in the vocal cord with optical coherence tomography. Otolaryngol Head Neck Surg 2006;134:86–91.
8. Parsons DS, Bothwell MR. Powered instrument papilloma excision: an alternative to laser therapy for recurrent respiratory papilloma. Laryngoscope 2001;111:1494–6.
9. El-Bitar MA, Zalzal GH. Powered instrumentation in the treatment of recurrent respiratory papillomatosis. Arch Otolaryngol Head Neck Surg 2002;128:425–8.
10. Pasquale K, Wiatrak B, Woolley A, et al. Microdebrider versus CO2 laser removal of recurrent respiratory papillomatosis: a prospective analysis. Laryngoscope 2003;113:139–43.
11. Derkay C, Malis D, Zalzal G, et al. A staging system for assessing severity of disease and response to therapy in recurrent respiratory papillomatosis. Laryngoscope 1998;108:935–7.
12. MultiDimensional Voice Program user manual. Lincoln Park, NJ: Kay Elemetrics, 1993.
13. Kempster GB, Gerrat BR, Verdolini Abbott K, et al. Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. Am J Speech Lang Pathol 2008 Oct 16 [Epub ahead of print].
14. Campisi P, Tewfik TL, Manoukian JJ, et al. Computer assisted voice analysis: Establishing a pediatric database. Arch Otolaryngol Head Neck Surg 2002;128:156–60.
15. Colton RH, Casper JK, Leonard R. Differential diagnosis of voice problems. In: Understanding voice problems. A physiological perspective for diagnosis and treatment. 3rd ed. Lippincott Williams & Wilkins; 2006. p. 12–63.
16. Campisi P, Tewfik TL, Pelland-Blais E, et al. MultiDimensional Voice Program analysis in children with vocal cord nodules. J Otolaryngol 2000;29:302–8.
17. Reinisch L. Laser physics and tissue interactions. Otolaryngol Clin North Am 1996;29(6):893–914.
18. Kelchner LN, Brehm SB, Weinrich B, et al. Perceptual evaluation of severe pediatric voice disorders: reliability using the consensus auditory perceptual evaluation of voice. J Voice 2009 Jan 8 [Epub ahead of print].
19. Karnell MP, Melton SD, Childes JM, et al. Reliability of clinician-based (GRBAS and CAPE-V) and patient-based (V-RQOL and IPVI) documentation of voice disorders. J Voice 2007;21:576–90.