

UCSF

UC San Francisco Previously Published Works

Title

The diagnostic role of high-speed vocal fold vibratory imaging.

Permalink

<https://escholarship.org/uc/item/2q0748r1>

Journal

Journal of voice : official journal of the Voice Foundation, 27(5)

ISSN

0892-1997

Authors

Mendelsohn, Abie H
Remacle, Marc
Courey, Mark S
[et al.](#)

Publication Date

2013-09-01

DOI

10.1016/j.jvoice.2013.04.011

Peer reviewed

The Diagnostic Role of High-Speed Vocal Fold Vibratory Imaging

*Abie H. Mendelsohn, †Marc Remacle, ‡Mark S. Courey, §Friedrich Gerhard, and ||Gregory N. Postma, *Los Angeles, and ‡San Francisco, California, †Yvoir, Belgium, §Graz, Austria, and ||Augusta, Georgia

Summary: Objective. Although high-speed imaging (HSI) has been identified as a valuable tool in phonatory biomechanics research, to date, there have only been a selected number of reports investigating the clinical utility of HSI. We aim to elucidate the role of HSI in the diagnosis of the dysphonic patient.

Methods. The video files from 28 consecutive dysphonic patients with concurrently acquired videostroboscopy and HSI were retrospectively collected. Stroboscopy video files were edited to include vibratory motion only. Videos were then anonymously and randomly presented to four academic laryngologists. Experts were asked to assign a single best diagnosis for each video file. Assigned diagnoses were then compared with treatment diagnoses conferred based on medical history, phonatory evaluation, laryngeal examination, and response to treatment.

Results. Interrater analysis for the four laryngologists demonstrated significant and meaningful correlations for the diagnoses of polyps, cysts, nodules, and normal examination using stroboscopy ($\kappa > 0.40$, $P < 0.001$). The experts demonstrated significant and meaningful correlations for the diagnoses of polyps, presbyphonia, and normal examination using HSI ($\kappa > 0.40$, $P < 0.001$). Combined intrarater analysis performed by comparing single rater's diagnosis for single patient across both modalities resulted in poor correlation without statistical significance ($\kappa = 0.30$, $P = 0.07$). Both stroboscopy- and HSI-assigned diagnoses matched the treatment diagnoses at equal predicted frequencies (32.3%), as demonstrated through multivariate logistic regression analysis ($P < 0.001$).

Conclusion. Overall, HSI did not improve the diagnostic accuracy above stroboscopy alone. Although specific laryngeal states such as presbyphonia may be better diagnosed with HSI, further studies are required to define HSI's precise role in the clinical setting.

Key Words: Stroboscopy–Laryngeal examination–High speed–Voice diagnostics–Interrater reliability–Vocal fold vibration–Visual subjective evaluation–Laryngoscopy–Diagnosis.

INTRODUCTION

Vocal fold mucosal vibration is the basis for voice production.¹ Based on this principle, the investigation² and the examination³ of those mucosal vibrations, or waves, have been at the cornerstone of phonatory evaluation. The standard of practice of the examination of mucosal waves is videolaryngostroboscopy. The clinical utility of stroboscopy was established with initial studies of clinical utility⁴ and diagnostic concordance.⁵ Since that time, stroboscopy has been confirmed to be the recommended examination modality in the specialized evaluation of the dysphonic patient.⁶

Notwithstanding, laryngeal stroboscopy contains many inherent weaknesses, including a dependence on regular phonation and minimum requisite phonation time of 2 seconds.⁷ Such limitations ultimately restrict its universal applicability in the clinical setting, specifically toward patients presenting with irregular phonatory cycles and limited phonatory dura-

tions. To address such technologic shortcomings, high-speed imaging (HSI) was applied to laryngoscopy to more clearly visualize mucosal wave mechanics. By capturing at least 2000 images per second, HSI can capture at least 10–20 frames per vibratory cycle depending on the fundamental frequency. By capturing multiple frames within each cycle, HSI is not dependent on periodic vibratory motion and can describe vibratory behavior beyond what was possible with videostroboscopy.⁸ Since its inception, lighting capacity, image quality, and image capture have advanced which allowed for increased clinical applications of laryngeal HSI. Because of these improvements, HSI laryngeal examinations have become more practical in the clinical setting. Current clinical recommendations include offering patients HSI examination as a clinical adjunctive analytic tool,⁹ specifically with severely dysphonic voices.¹⁰ However, there is currently little information guiding the specific clinical scenarios or indications, in which HSI may offer clinical benefit above the standard of laryngovideostroboscopy. It is therefore our objective to perform an exploratory study to identify in which specific laryngeal pathologic states HSI may improve the diagnostic yield above the current standard of videostroboscopy.

METHODS

Institutional Review Board approval was obtained before the initiation of this study.

Dysphonic patients presenting to the participating academic laryngology center are routinely evaluated and examined by clinical voice assessment and indirect laryngovideostroboscopy. Patients are systematically offered high-speed imaging

Accepted for publication April 23, 2013.

Presenting Information: This research was presented at The Fall Voice Conference, NYU School of Medicine; October 4–6, 2012; New York, New York.

From the *Department of Head and Neck Surgery, David Geffen School of Medicine at UCLA, Los Angeles, California; †Department of Otolaryngology-Head and Neck Surgery, Louvain University Hospital of Mont-Godinne, Yvoir, Belgium; ‡Department of Otolaryngology-Head and Neck Surgery, University of California - San Francisco, San Francisco, California; §Department of Phoniatrics, Speech and Swallowing, ENT University Hospital, Medical University of Graz, Graz, Austria; and the ||Department of Otolaryngology, Georgia Health Sciences University, Augusta, Georgia.

Address correspondence and reprint requests to Abie H. Mendelsohn, Department of Head and Neck Surgery, UCLA Voice Center for Medicine and the Arts, 924 Westwood Blvd, Suite 515, Los Angeles CA 90024. E-mail: Amendelsohn@mednet.ucla.edu

Journal of Voice, Vol. 27, No. 5, pp. 627–631

0892-1997/\$36.00

© 2013 The Voice Foundation

<http://dx.doi.org/10.1016/j.jvoice.2013.04.011>

(HSI) using a Color High-Speed Video System (model 9710; KayPENTAX, Montvale, NJ). HSI examination is performed during the same clinical encounter as the stroboscopic examination. All digital video examinations are stored on archived data drives.

We reviewed our HSI experience during an 18-month period (1/1/2011 to 6/30/2012). Twenty-eight unique patients underwent videostroboscopy and HSI examinations during the same clinical encounter. To provide a fully inclusive study, no patient was excluded based on video quality or pathologic diagnosis. With each patient providing two laryngeal examinations, a total of 56 video files were collected. Clinical data were subsequently retrieved from the medical charts. The final treatment diagnosis was also recorded from the medical records. Treatment diagnosis was determined by history, laryngoscopic examination, voice assessment, operative findings (when applicable), and response to treatment.

Digital video files from both examinations were transferred to a video-editing system (iMovie 11 v.9.0.4; Apple, Inc., Cupertino, CA) for further optimization. A single laryngologist (A.H.M.) performed all video-editing responsibilities. Videos were trimmed to include only images demonstrating vibratory assessment. As HSI does not routinely capture abduction, stroboscopy videos were edited to remove abduction views. The videos were further edited to include the best representative 5-second clip as judged by the editing author. From the 5-second video clip, a 3-second sample was copied and replayed at the end of the video, slowed to half-time (additional 6-second). This resulted in a single 11-second video for both the stroboscopy and HSI examinations. Because HSI video clips did not include sound files, sound was removed from

the stroboscopy videos as has been performed in previous studies.¹¹

Examination videos were de-identified, coded, and randomized. The videos were then uploaded to an online rating platform (<http://www.classmarker.com>; ClassMarker Pty Ltd., Sydney, Australia). Each video was presented separately. Raters had the option to replay videos in full screen mode and *ad libitum*. Raters were presented with a list of 10 clinical diagnoses and were asked to choose a single best diagnosis for each examination (Figure 1).

Four academic laryngologists, each with a minimum of 15 years clinical experience, were asked to participate in the study by performing the video interpretations. Three of the four experts had never seen the examinations before the study. One expert had previously viewed the examinations; however, no significant difference was demonstrated between these raters and the other raters' diagnostic concordance.

Diagnostic assignments were obtained in spreadsheet format for each rater. Matching single patient video codes allowed re-labeling of single patient laryngeal examination results for stroboscopy and HSI. Standard descriptive statistics were applied. Subsequently, Cohen's kappa was configured to measure interrater and intrarater correlations. Cohen's kappa coefficient is a statistical measure of interrater agreement for categorical items and is a more robust measure than simple percent agreement calculation because it takes into account the agreement occurring by chance. Multivariate logistic regression models were configured for each examination technique controlling for predicted probabilities of random matching for both stroboscopy and HSI. Statistical significance was defined as *P* value less than 0.05. Statistical analysis was performed in *Stata Statistics/Data Analysis* v11.2 (StataCorp LP, College Station, TX).

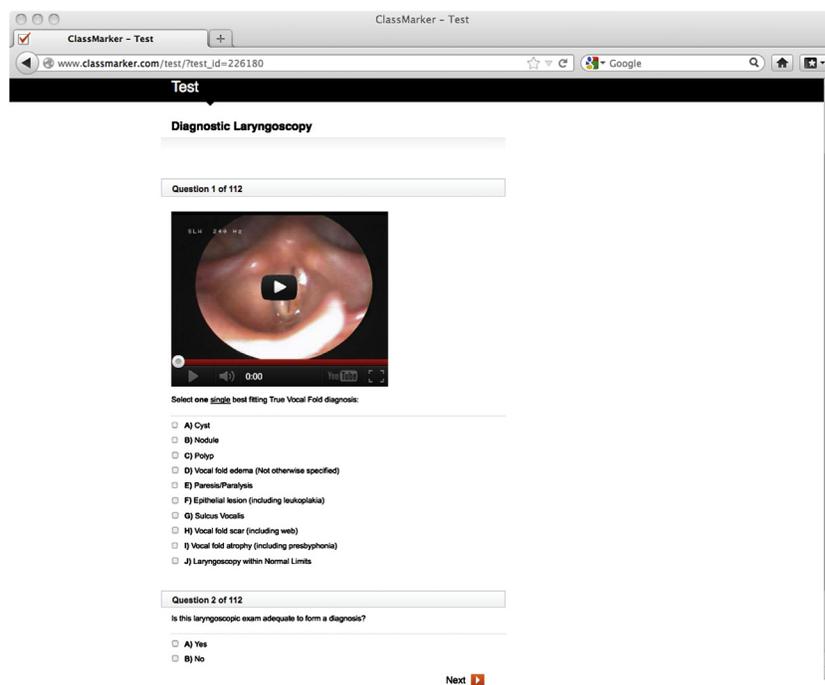


FIGURE 1. Laryngeal examination rating online platform. Video could be played as presented or in full screen mode. Raters were asked to choose the single best diagnosis for the examination.

RESULTS

Twenty-eight unique patients underwent laryngeal videostroboscopy and HSI examination at the same clinical encounter. The study group comprised 17 females and 11 males. The average age at examination was 48.7 years (range 18–79 years). The most common symptomatic complaint seen in 14 (50%) of the patients was a chronic nonfluctuating decrease in voice quality. The most common treatment diagnosis given following complete history, examinations, and response to therapy was normal vibratory examination (n = 7, 25.0%), including such etiologies as muscle tension dysphonia and functional dysphonia. The most common pathologic treatment diagnosis was benign vocal fold lesions (cyst = 2, nodules = 3, polyp = 2; 25.0%). Each of the 28 patients had a single stroboscopy and single HSI video file resulting in 56 video files. Four expert raters viewed each video file resulting in 224 rated laryngoscopic examinations. Figure 2 displays the comparative incidence of each examination diagnosis for both the stroboscopy examinations (Figure 2A) and the HSI examinations (Figure 2B) as compared with the incidence of each treatment diagnosis.

To specifically evaluate each laryngeal state by examination technique, interrater correlations using Cohen’s kappa analysis were calculated for each diagnostic category and are displayed in Table 1. Although multiple laryngeal pathologies demonstrated statistically significant interrater correlation values for both techniques, most of these correlations failed to reach a substantive magnitude as previously defined by Landis and Koch.¹² They described that kappa values must fall between 0.21 and

0.40 to be considered *fair* level of agreement, 0.41 and 0.60 for *moderate* agreement, and 0.61 and 0.80 for *substantial* agreement. Based on this definition of correlation, stroboscopy diagnoses resulted in only “Polyps” reaching significant and *substantial* interrater correlation. “Cysts,” “Nodules,” and “Normal Exams” reached significant and *moderate* correlations. For HSI, none of the correlated diagnoses with significant *P* values reached a *substantial* magnitude of agreement. “Polyps,” “Presbyphonia,” and “Normal Exams” reached significant and *moderate* correlations. Notably, as a combined modality across all diagnoses, neither stroboscopy nor HSI could reach above a *fair* level of correlation. Additionally, when HSI was used, “Paralysis/Paresis,” “Epithelial lesions,” and “Sulcus vocalis” demonstrated negative kappa values. This finding suggests a worse-than-random guessing correlation for these three diagnostic categories.

In assessing which laryngeal examination technique matched the treatment diagnosis more frequently, a logistic regression analysis model was formulated with random effects to consider individual patient and rater variability. The results of the logistic model are displayed in Table 2. The calculated predicted probabilities for each laryngoscopic examination matching the single treatment diagnosis for that examination was exactly equal at 32.3%. The exactly identical values of association between the two imaging techniques was confirmed through repeated statistical modeling. Both techniques reached statistical significance.

Overall diagnostic concordance between stroboscopy and HSI was 41.1%. Statistically, this was analyzed using

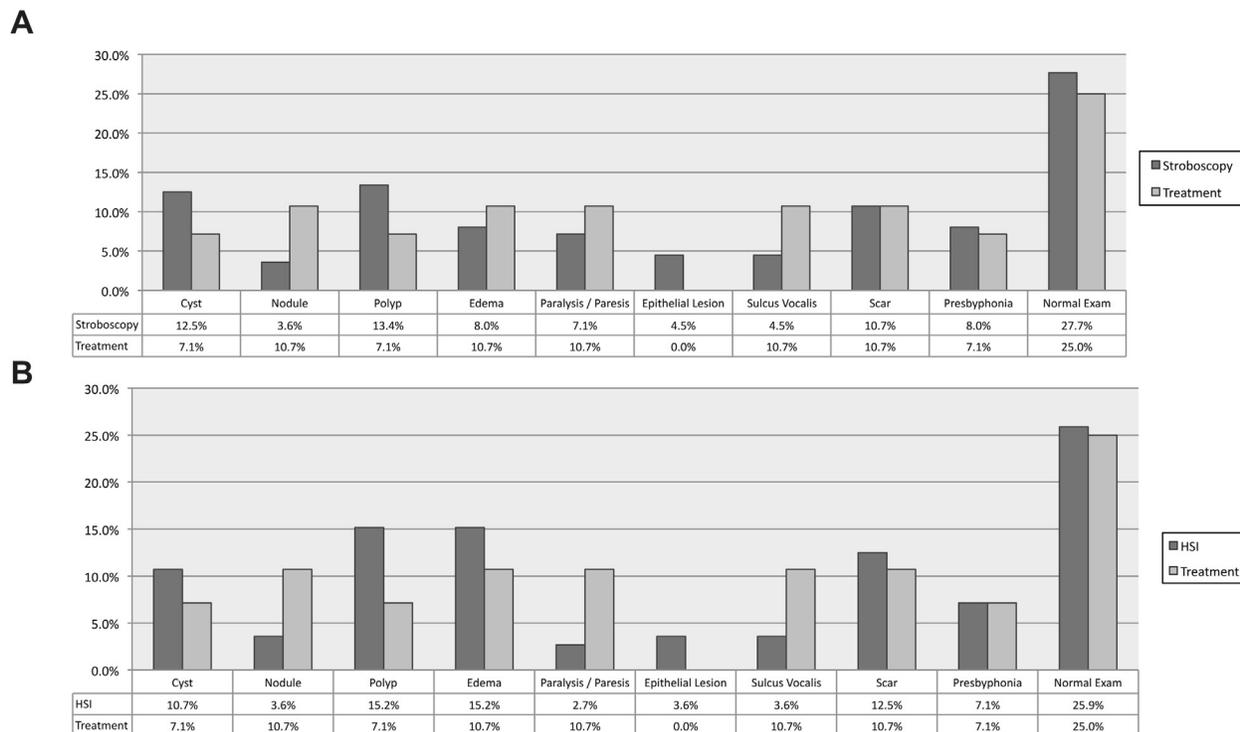


FIGURE 2. Incident frequency of diagnostic assignment. The figure displays the incidence of each diagnostic category by stroboscopy (A) and HSI (B) as compared with the treatment diagnostic incidence.

TABLE 1.
Interrater Reliability (Cohen's kappa) for Laryngeal Stroboscopy and HSI by Diagnostic Categories

Laryngeal Diagnosis	Stroboscopy			HSI		
	Kappa	Z	P Value	Kappa	Z	P Value
Cyst	0.4014	5.20	<0.001*	0.1911	2.48	0.01*
Nodule	0.4815	6.24	<0.001*	0.1358	1.76	0.04*
Polyp	0.6151	7.97	<0.001*	0.4683	6.07	<0.001*
Edema	0.1543	2.00	0.02*	0.1447	1.88	0.03*
Paralysis/paresis	0.1923	2.49	0.01*	-0.0275	-0.36	0.64
Epithelial lesion	0.0928	1.20	0.11	-0.0370	-0.48	0.68
Sulcus vocalis	0.0928	1.20	0.11	-0.0370	-0.48	0.68
Scar	0.1289	1.67	0.05*	0.1293	1.68	0.05*
Presbyphonia	0.2348	3.04	0.00*	0.5513	7.15	<0.001*
Normal exam	0.5094	6.60	<0.001*	0.4571	5.93	<0.001*
Combined effects	0.3456	11.60	<0.001*	0.2788	9.01	<0.001*

* Statistically significant correlations.

a combined intrarater correlation kappa value. Across all raters, the diagnostic correlation between the two techniques was 0.303 ($P = 0.07$), reaching a *fair* level of agreement without statistical significance. Although the overall agreement frequency of 41.1% exceeded the expected random agreement of one of 10 diagnostic choices or 10%, the statistical correlation calculation failed to demonstrate this level of correlation as significant beyond the chance of a random occurrence.

DISCUSSION

The evaluation of the dysphonic patient requires a specialized approach. Although medical history and voice analysis are critical for an accurate diagnosis, close laryngeal examination is paramount. The accepted standard for the specialized laryngeal evaluation is the performance of laryngovideostroboscopy.⁶ The synchronized flashing light of stroboscopy provides sampling images of the propagating vocal fold mucosal wave, presenting valuable information not otherwise obtained during halogen light indirect laryngoscopy. This synchronization of light flashes requires real-time assessment of fundamental frequency, which is not consistently possible in dysphonic voices. As such, there are clinical scenarios in which stroboscopy is of diminished clinical utility. In distinction, HSI obtains real-time visualization of vocal fold vibrations and, as such, is not limited by dysphonic or aperiodic voices.

The present study set out to explore the clinical scenarios in which HSI may improve the diagnostic potential beyond the limitations of stroboscopy. To answer this challenge, this study presented stroboscopic and HSI examinations to four academic

laryngologists specifically for diagnostic assignment. The study was designed to mimic the clinical demand of typical laryngology practice and, therefore, no etiology of dysphonia was excluded from this study group. The goal of the study was designed to analyze the diagnostic capacity of the video examinations alone; therefore, no auditory and no clinical information was provided with the presented laryngeal examinations. Previous investigations have examined the auditory characteristics that are more indicated for HSI examination, mainly those with moderate-to-severe dysphonia.¹⁰

HSI demonstrated significant and meaningful interrater correlations for specific diagnostic categories, including vocal fold polyps and presbyphonia. Expert raters were also able to reliably identify normal laryngeal examinations with HSI. Using stroboscopy, significant correlations were demonstrated with vocal fold polyps, cysts, and nodules. This finding correlates to previous reports demonstrating reliable stroboscopic evaluation of subepithelial lesions without benefit from HSI.¹⁰ Though vocal fold polyps were reliably diagnosed using HSI in the present study, the magnitude of correlation was greater using stroboscopy, and the data do not support the indication of HSI to improve the diagnostic capacity of polyps. The present study did identify presbyphonia as a clinical pathology in which the diagnostic concordance of HSI was greater than stroboscopy. That HSI diagnoses presbyphonia with improved correlation can be explained by more subtle laryngoscopic findings in absence of mass effect. The diagnosis of presbyphonia requires visualization of exaggerated mucosal wave excursions seen with vocal fold atrophy, which may be suboptimally visualized using stroboscopy. Interestingly, the diagnostic

TABLE 2.
Multivariate Logistic Regression Analysis Modeling Predicted Diagnostic Margin by Laryngoscopic Technique

Laryngoscopy Technique	Margin	Standard Error	Z	P Value	95% Confidence Interval
Stroboscopy	0.3227	0.06	5.68	<0.001	0.21–0.43
HSI	0.3227	0.06	5.68	<0.001	0.21–0.43

identification of sulcus vocalis, which classically presents diagnosis challenges, was not readily enhanced by HSI. This finding may be explained by the absence of abductory views and the limitations of medial edge visualization provided during HSI.

Overall, as individual laryngeal examination techniques, stroboscopy and HSI provided exactly the same likelihood of predicting the treatment diagnosis. The exact identical values of predicted frequency on multivariate analysis between the two imaging techniques was a surprising finding that was confirmed through repeated statistical modeling. However, it is not surprising that HSI was not identified as a significantly more efficacious modality as this is in agreement with previous comparative studies that noted HSI should function only in a complementary manner to stroboscopy. Olthoff et al noted that while HSI was more reliable and less prone to interpretive variances than stroboscopy, their comparative study was unable to demonstrate any specific clinical scenarios in which HSI is clearly indicated.¹¹ Patel et al found HSI a useful adjunct to stroboscopy specifically with neuromuscular etiologies of dysphonia; however, it is unclear as to what extent clinical diagnosis and treatment would differ without HSI.¹⁰ Mortensen and Woo described a single case report in which HSI did change the clinical course of a dysphonia patient. Using HSI, a new finding of unilateral mucosal wave asymmetry was seen which was not found on stroboscopy.¹³ Based on the results of the present study, HSI appears to be without improved diagnostic capacity as a stand-alone technique apart from stroboscopy. The authors agree with previous recommendations that HSI be used in challenging diagnostic cases. Specifically, in the case of vocal fold scar. Although the present study did not identify vocal fold scar to be better examined by HSI, it has been the authors' clinical experience that this entity is identified more readily using HSI. Additionally, HSI should be used when specific vibratory factors may enhance operative or treatment planning.

The present study was prospectively designed but still contains many limitations. As a diagnostic modality, laryngeal imaging is never used independent of critical factors including voice quality, medical history, along with laryngeal abductory views and mobility. The present study purposely removed all these factors to highlight the imaging differences between the two modalities; however, in doing so, limited the diagnostic capacity and reliability of the imaging techniques. Additionally, this study sought to find diagnostic concordance for conditions such as benign vocal fold lesions in which there is much disagreement regarding the basic diagnostic criteria.^{14,15} For instance, it is unclear whether the diagnostic variability in the present study with vocal fold nodules represents limitations

of the imaging modality or the agreement as to the physical appearance of those nodules. Finally, the wide range of diagnostic possibilities engaged multiple levels of variability, which ultimately limits the statistical power of the findings. However, as an initial exploratory study, the wide incorporation was a planned design characteristic so that future work may have a basis in which to limit additional comparative studies to diagnoses of interest such as presbyphonia.

CONCLUSION

Overall, HSI did not improve the diagnostic accuracy above stroboscopy alone. Although specific laryngeal states such as presbyphonia may be better diagnosed with HSI, further studies are required to define HSI's precise role in the clinical setting.

REFERENCES

1. Timcke R, Von Leden H, Moore P. Laryngeal vibrations: measurement of the glottic wave. Part I. The normal vibratory cycle. *Arch Otolaryngol.* 1958;68:1-9.
2. Berke GS, Gerratt BR. Laryngeal biomechanics: an overview of mucosal wave mechanics. *J Voice.* 1993;7:123-128.
3. Titze IR, Jiang JJ, Tzu-Yu H. Measurements of mucosal wave propagation and vertical phase difference in vocal fold vibration. *J Acoust Soc Am.* 1992; 91:2420.
4. Woo P, Colton R, Casper J, Brewer D. Diagnostic value of stroboscopic examination in hoarse patients. *J Voice.* 1991;5:231-238.
5. Sataloff RT, Spiegel JR, Hawkshaw MJ. Strobovideolaryngoscopy: results and clinical value. *Ann Otol Rhinol Laryngol.* 1991;100(9 pt 1):725-727.
6. Paul BC, Chen S, Sridharan S, Fang Y, Amin MR, Branski RC. Diagnostic accuracy of history, laryngoscopy, and stroboscopy. *Laryngoscope.* 2013; 123:215-219.
7. Kitzing P. Stroboscopy—a pertinent laryngological examination. *J Otolaryngol.* 1985;14:151-157.
8. Jiang JJ, Yumoto E, Lin SJ, et al. Quantitative measurement of mucosal wave by high-speed photography in excised larynges. *Ann Otol Rhinol Laryngol.* 1998;107:98-103.
9. Inwald EC, Döllinger M, Schuster M, Eysholdt U, Bohr C. Multiparametric analysis of vocal fold vibrations in healthy and disordered voices in high-speed imaging. *J Voice.* 2011;25:576-590.
10. Patel R, Dailey S, Bless D. Comparison of high-speed digital imaging with stroboscopy for laryngeal imaging of glottal disorders. *Ann Otol Rhinol Laryngol.* 2008;117:413-424.
11. Olthoff A, Woywod C, Kruse E. Stroboscopy versus high-speed glottography: a comparative study. *Laryngoscope.* 2007;117:1123-1126.
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159-174.
13. Mortensen M, Woo P. High-speed imaging used to detect vocal fold paresis: a case report. *Ann Otol Rhinol Laryngol.* 2008;117:684-687.
14. Hantzakos A, Remacle M, Dikkers FG, et al. Exudative lesions of Reinke's space: a terminology proposal. *Eur Arch Otorhinolaryngol.* 2009;266: 869-878.
15. Rosen CA, Gartner-Schmidt J, Hathaway B, Simpson CB, Postma GN, Courey M, Sataloff RT. A nomenclature paradigm for benign midmembranous vocal fold lesions. *Laryngoscope.* 2012;122:1335-1341.