Published in final edited form as:

J Voice. 2018 March ; 32(2): 216–220. doi:10.1016/j.jvoice.2017.05.002.

Utility of Laryngeal High-Speed Videoendoscopy in Clinical Voice Assessment

Stephanie RC Zacharias^{1,2,3,4}, Dimitar D Deliyski^{2,5,6}, Terri Treman Gerlach^{6,7}

¹Division of Speech-Language Pathology, Cincinnati Children's Hospital Medical Center, 3333 Burnet Ave, Cincinnati, OH, USA

²Division of Pediatric Otolaryngology, Cincinnati Children's Hospital Medical Center, 3333 Burnet Ave, Cincinnati, OH, USA

³Department of Otolaryngology-Head and Neck Surgery, University of Cincinnati, 3230 Eden Ave, Cincinnati, OH, USA

⁴Department of Communication Sciences and Disorders, University Cincinnati, 3202 Eden Ave, Cincinnati, OH, USA

⁵Department of Communicative Sciences and Disorders, Michigan State University, 1026 Red Cedar Rd, East Lansing, MI, USA

⁶Department of Communication Sciences and Disorders, University of South Carolina, 1300 Wheat St, Columbia, SC 29208, USA

⁷Voice and Swallowing Center, Charlotte Eye Ear Nose and Throat Associates, 6035 Fairview Rd, Charlotte, NC, USA

Abstract

Objective—To assess the utility of laryngeal high-speed videoendoscopy (HSV) as a clinical tool.

Design—Prospective study of 151 patients.

Methods—151 adult patients (52 male, 99 female) underwent both videoendoscopy with stroboscopy (videostroboscopy) and HSV examination as part of a routine clinical voice assessment. At the time of the examination, ratings for videostroboscopy were reported in the clinical report. Next, the clinicians reviewed the HSV examination and indicated the changes in ratings of HSV relative to videostroboscopy. Finally, the clinical reports were reviewed by a clinician not involved with data collection or clinical care of the patients and noted differences between videostroboscopy and HSV clinical ratings and resulting diagnoses were identified and grouped.

Corresponding author: Stephanie RC Zacharias, Ph.D., Cincinnati Children's Hospital Medical Center, 3333 Burnet Ave, MLC 15008, Cincinnati, OH 45229-3039, Tel: +1 513 803 2801, Fax: +1 513 803 1911, Stephanie.Zacharias@cchmc.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Results—Ratings of all vibratory features showed change between videostroboscopy and HSV. Mucosal wave and amplitude of vibration showed the largest percentage change, respectively, in 74% and 53% of the reports. They were followed by the features of glottal closure (36%), phase closure (32%), glottal edge (25%), and phase symmetry (21%). Ratings of supraglottic compression and vocal fold vertical level showed least change between videostroboscopy and HSV. Changes in initial diagnosis due to the inclusion of HSV were indicated in 7% of the cases.

Conclusions—HSV may be an important laryngeal imaging technique for functional assessment the pathophysiology of certain voice disorders. HSV could enable important refinements in the diagnosis and management of vocal fold pathology.

Keywords

Voice Disorders; High-Speed Videoendoscopy; Videostroboscopy; Vocal-Fold Vibration

1. Introduction

The value of direct visualization of laryngeal structures and vocal fold vibratory patterns to assess dysphonia is essential in the clinical realm. Currently, videoendoscopy with stroboscopy (videostroboscopy) is the most important clinical tool for instrumental voice assessment, allowing for the examination of vocal fold anatomy and physiology in individuals with voice disorders.¹ Videostroboscopy has improved overall assessment and diagnosis in individuals with voice disorders, yet, videostroboscopy only allows observation of periodic vocal fold vibration. In individuals with voice disorders who present with aperiodic vocal fold vibration and/or intra-cycle irregularities, the stroboscopic light is unable to synchronize and visualization and interpretation of the vibratory features of the vocal folds is lost.^{2–8} The importance of assessing vibratory behaviors of the vocal folds in refinement of therapeutic and surgical interventions has been reported.^{9,10} Therefore, laryngeal imaging techniques are essential for the assessment of vocal fold vibration and for the enhancement of the therapeutic and surgical voice outcomes in the clinical setting.

High-speed videoendoscopy (HSV) is a potential valuable component of clinical protocols for voice assessment and laryngeal evaluation. With significantly higher frame rates compared to videostroboscopy, HSV is able to provide true intra-cycle visualization within the vibratory cycle. On the contrary, videostroboscopy relies on a composite vibratory image sequence consisting of phases taken across many cycles.¹¹ Therefore, it can be expected that introducing high temporal resolution would lead to refinement of the assessment of disordered vocal fold vibration, because laryngeal imaging is not dependent on the assumption of periodicity. Other studies report such evidence.^{3,7} Additionally, clinicians report equal to higher levels of confidence when rating vibratory features using simulated stroboscopy from HSV compared to videostroboscopy, further demonstrating the efficacy of HSV-based imaging.¹² The advancement of clinical protocols and the development of commercial HSV systems highly depend on providing clinical evidence for the utility and efficacy of HSV. In a research setting, videostroboscopy and HSV have been compared in both normal and dysphonic patients, suggesting differences in ratings of vocal fold vibratory features.^{5,13–15} However, few reports systematically document the clinical value of HSV examinations, resulting in a lack of research in realistic settings on rating the differences of

vibratory features seen on videostroboscopy and HSV, and how it may affect diagnosis and treatment decision making

The purpose of this study was to provide initial clinical evidence of the utility of HSV for voice assessment in an adult clinical setting. The study was designed to evaluate the added value of HSV in the evaluation of adult patients with a voice disorder. It was hypothesized that the inclusion of HSV in the clinical protocol would result in changes of the ratings of vocal fold vibratory characteristics and the initial diagnoses. The clinical adoption of HSV could provide refinements in the assessment of vocal fold vibratory characteristics that may lead to a more accurate functional assessment and diagnosis of voice disorders.

2. Method

2.1 Research Design

The research design of this study was intended to replicate the real-life clinical settings. Instead of implementing an experimental design with controlled independent variables, a comparative study design was carried out. This allowed the participating clinicians to implement HSV as it best complements clinical voice assessment in an active voice center, The Voice and Swallowing Center at the Charlotte Eye Ear Nose and Throat Associates (CEENTA), in Charlotte, NC.

For this study, a custom-designed HSV system was implemented at the Center in 2007. An initial full-year period allowed the clinicians to be trained and become accustomed to utilizing HSV in their daily clinical practice. In the following four years, HSV was used as part of the clinical voice evaluations, subsequent to perceptual assessment, acoustic analysis, and videostroboscopy.

Three voice-specialized speech-language pathologists (voice clinicians) performed initial voice assessment of incoming patients. The clinicians were instructed to use HSV following the standard clinical protocol and consent. Clinicians were also instructed to write the voice reports based on the standard clinical protocol with the addition of noting the changes in functional ratings between videostroboscopy and HSV.

2.2 Human Data

Following voice-related quality of life self-assessment, medical history, self-reporting perceptual evaluation of voice quality, and acoustic analysis; each patient was examined by a voice clinician, using a 70-degree 10-mm Rigid Laryngoscope, Model 49-4072 (JEDMED Instrument Co, St. Louis, MO), via videostroboscopy with subsequent HSV examination. In both examinations, participants were instructed to produce multiple sustained phonations of the vowel /i/ at comfortable pitch and loudness. Comfortable pitch and loudness were elicited by carrying a conversation, followed by producing the sustained phonations at the same pitch and loudness. The digital videostroboscopy system used was KayPENTAX Model RLS 9100B (PENTAX Medical, Montvale, NJ) coupled with a 120-W xenon light source. The full videoendoscopy with stroboscopy examination recordings were several minutes in length, however the sustained-vowel videostroboscopic portions used for evaluating the studied vibratory features were 3–5 seconds in length. The custom HSV

system consisted of a Phantom v7.3 high-speed camera by Vision Research (Vision Research, Inc., Wayne, NJ), coupled with a 300-W constant xenon light source (Model 7152A, PENTAX Medical Company, Montvale, NJ). HSV examinations were recorded in color at the speed of 5,512.5 frames per second (fps) (N=122), or in monochrome at the speed of 11,025 fps (N=29). Examination recording length for the color HSV was 4.2 seconds and for the monochrome was 3.5 seconds.

Voice assessment reports of 151 voice patients, 22–84 years of age (52 male, 99 female) were used in this study. Those data were collected in the CEENTA Voice and Swallowing Center from July 2008 through January 2012. Patients were diagnosed with a variety of voice disorders (Table 1). At CEENTA, voice clinicians complete all videostroboscopy and HSV evaluations. They have been trained to rate both the anatomic and vibratory characteristics (physiology) of vocal fold phonation. These ratings of vocal fold phonation were then described and reported in the clinical report. All clinical videos and reports were subsequently reviewed by an otolaryngologist in consultation with the voice clinicians. This study was approved by the Institutional Review Board of the University of South Carolina.

2.3 Measurement

As part of the established clinical protocol, eight vocal-fold vibratory features were judged for differences between videostroboscopy and HSV: mucosal wave, amplitude of vibration, glottal closure pattern, phase closure, glottal edge, phase symmetry, supraglottic compression, and level of vocal fold approximation. Periodicity was not included due to the inherent inability of videostroboscopy to evaluate this feature.^{5,14} Voice clinicians were instructed to review steady-state phonation in both the videostroboscopy and HSV samples. For the videostroboscopy examinations, the video playback software was set at 30 fps by default, while for the HSV examinations the defaults were set at 30 and 60 fps for color and monochrome, respectively. However, clinicians had the ability to adjust the speed, pause the playback, or review frame-by-frame. In the instances in which voice clinicians noted differences, a narrative was written describing these changes. The Ear, Nose and Throat physicians also provided comments when diagnostic changes resulted from the inclusion of HSV.

A fourth voice clinician not involved in data collection or clinical care of the study patients reviewed each clinical report and identified the differences between the clinical ratings and diagnoses of HSV and videostroboscopy. The narratives written by the voice clinicians and physicians describing the differences between videostroboscopy and HSV falling within and outside of the above pre-identified vibratory features were grouped as themes. Instances where clinical judgment and diagnostic changes between videostroboscopy and HSV were noted in the clinical report, were documented.

2.4 Analysis

To identify change in clinical ratings of vibratory features between videostroboscopy and HSV, frequency counts and percentages were used. Themes were identified and grouped.

Results

Ratings of vibratory features of mucosal wave and amplitude of vibration showed the largest percentage change between videostroboscopy and HSV, followed by features of glottal closure, phase closure, glottal edge, and phase symmetry (Fig. 1). Ratings of anatomical features of supraglottic compression and vocal fold vertical level showed least change between videostroboscopy and HSV. Recurring themes emerged when vibratory feature ratings changed from videostroboscopy to HSV (Table 2).

The most significant finding from this study was the number of patients whose management changed as a direct result of the additional information HSV provided (Table 3.) Eleven patients in the study (7%) had their initial diagnosis change due to HSV. Lesion(s) of the vocal fold(s) represented the largest group with diagnostic changes (73%). Further, individuals who presented with moderate and severe dysphonia characterized 83% of the cases with diagnostic changes.

3 Discussion

HSV is emerging as a potential imaging technique for evaluating patients with voice complaints, yet usefulness in the clinic needs to be elucidated. Results from this study demonstrated there are significant clinical rating changes in all vibratory features between videostroboscopy and HSV. The greatest differences in ratings between videostroboscopy and HSV were found in the assessment of mucosal wave (74%) and amplitude of vibration (53%). These ratings require a higher temporal resolution which explains these differences. Results from previous research studies, also indicated such differences due to temporal resolution.^{16,17} However, they have not been demonstrated in the clinical setting. An unexpected result was that an increase in mucosal wave and amplitude of vibration rating was the typical change. For example, rating of severely decreased mucosal wave on videostroboscopy changed to moderately decreased on HSV. Previous reports comparing videostroboscopy to HSV did not find the same increase in mucosal wave and amplitude of vibration ratings, likely because the frame rates used in those studies were lower (2,000 – 4,000 fps) than those used in this study (5,512.5 and 11,025 fps).^{3,5,15,17} The clinical reports, in our study, stated these vibratory differences often led to clinical conclusions of refined lesion descriptions, or greater vocal fold weakness; all of which would have been missed with videostroboscopy only. These vibratory features are good indicators for vocal fold health and pliability and therefore their clinical relevance is underscored. The examples provided in Video 1 and Video 2 illustrate the differences in assessing vibratory features from videostroboscopy and HSV, respectively. These are samples of a 73-year old male diagnosed with paresis and gastroesophageal reflux. Noted differences are increased mucosal wave and asymmetric amplitude of vibration in HSV compared to videoendoscopy.

Furthermore, clinical ratings of glottal closure changed in 36% of the cases. Glottal closure is an important feature assessed during a voice evaluation, as it plays a critical role when establishing a diagnosis and clinical treatment plan. There is a range of glottic closure patterns (e.g., complete, incomplete, hourglass, anterior gap, posterior gap) representing a

variety of vocal pathology. Failure in observing the proper pattern may lead to inappropriate selection of surgical or therapy techniques.

The initial diagnostic and intervention indicators changed in a significant percentage of cases (7%), as a direct result of the inclusion of HSV. The largest diagnostic group with initial diagnostic changes was vocal fold lesions. This most likely occurred due to the improved visual vibratory characteristics, enhancing the examiners ability to accurately assess the vibratory characteristics of both the vocal fold and lesion(s). In other words, with better cycle-to-cycle visualization using HSV, there is better resolution of the vibratory characteristic. Such a high percentage of diagnostic change recognizes the fact that there are some patients in whom videostroboscopy will fail to support the correct diagnosis. These results have important clinical significance in that these individuals may undergo a variety of therapies or procedures, prior to successful intervention.

To our knowledge, this is the first study conducted in a clinical setting and with HSV rates above 4,000 fps. Results from this study differ from other research studies in the assessment of vocal fold vibratory features and diagnostic findings. This may be due to the higher temporal resolution of the HSV system used in the study. Recently it has been recommended that clinical HSV systems should allow for frame rates of 8,000 fps, or above, with a minimum frame rate of 4,000 fps.¹⁸ The study reported that HSV recordings at rates below 4,000 fps (women), and 3,200 fps (men) should be interpreted with caution and that rates of 2,000 fps and lower are inadequate for interpreting many clinical features. Therefore, previous studies may not have had high enough frame rates to properly assess the differences of vibratory features of the vocal folds and diagnostic features between videostroboscopy and HSV.

Due to the comparative research design of this study, there are inherent limitations. First, in real-life clinical settings, clinician biases may have occurred that could have affected the participants' inclusion while conducting the research. For example, some of the patients presenting with mild dysphonia may not have received the full protocol and consequently were not included in the HSV examinations. The clinicians reported that in the last two years of the study they developed a sense of the patients that can benefit from HSV. Therefore, our data may tend to over-represent the proportion of cases with moderate and severe dysphonia, where videostroboscopy failed to synchronize the image prompting the clinicians to use HSV. Nevertheless, this is added evidence of the utility of HSV demonstrating how clinicians learn to take advantage of HSV in their daily practice. This bias may have decreased the external validity of the study, specifically, the population validity, leading to poor generalization to other patient populations. Further, the descriptive nature of the study does not allow to infer the knowledge gained by demonstrating the groups of patients for whom HSV may be most appropriate. Last, inter/intra-rater reliability was not assessed due to the real-life clinical setting; however, future studies could assess reliability of ratings of VLS and HSV features, by utilizing independent clinician raters to rate endoscopic and stroboscopic features. Future studies are required to address these limitations by implementing experimental design with controlled independent variables.

4 Conclusion

This study demonstrates the utility of HSV in the daily clinical practice of a large voice center. HSV may be a more accurate method for functional assessment of the pathophysiology of patients with moderate to severe voice disorders and those with certain types of voice disorders. The clinical implementation of HSV may lead to important refinements in the diagnosis and management of vocal fold pathology.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Funding for this study was provided by the National Institutes of Health — NIDCD: R01-DC007640 “Efficacy of Laryngeal High-Speed Videoendoscopy”. Special thanks to Lori Ellen Sutton and Susan Hanks for their role in completing the voice evaluations.

References

1. Awan S, Barkmeier-Kraemer J, Courey M, Deliyski D, Eadie T, Svec J, Patel R, Hillman R, Paul D (2014). Standard Clinical Protocols for Endoscopic, Acoustic and Aerodynamic Voice Assessment: Recommendations from ASHA Expert Committee. Abstracts: Convention of the American Speech-Language-Hearing Association. Orlando, Florida, 11 20–22, 2014
2. Deliyski D Laryngeal high-speed videoendoscopy In: Kendall KA, Leonard RJ, eds. Laryngeal Evaluation: Indirect Laryngoscopy to High-Speed Digital Imaging. New York, NY: Thieme Medical Publishers; 2010: 243–270.
3. Patel R, Dailey S, Bless D (2008). Comparison of high-speed digital imaging with stroboscopy for laryngeal imaging of glottal disorders. *Ann Otol Rhinol Laryngol*, 117(6):413–424. [PubMed: 18646437]
4. Deliyski D (2008). Clinical implementation of laryngeal high-speed videoendoscopy: Challenges and evolution. *Folia Phoniatrica et Logopaedica*, 60:33–44. [PubMed: 18057909]
5. Kendall KA (2009). High-speed laryngeal imaging compared with videostroboscopy in healthy subjects. *Arch Otolaryngol Head Neck Surg*, 135(3):274–281. [PubMed: 19289706]
6. Zacharias SR, Weinrich B, Brehm SB, Kelchner L, Deliyski D, Tabangin M, de Alarcon A (2015). Assessment of Vibratory Characteristics in Children Following Airway Reconstruction Using Flexible and Rigid Endoscopy and Stroboscopy. *JAMA Otolaryngol Head Neck Surg*, 141(10):882–7. [PubMed: 26402578]
7. Zacharias SR, Myer CM 4th., Meinzen-Derr J, Kelchner L, Deliyski DD, de Alarcón A (2016). A Comparison of Videostroboscopy and High-speed Videoendoscopy in Evaluation of Supraglottic Phonation. *Ann Otol Rhinol Laryngol*, 125(10):829–37. [PubMed: 27411363]
8. Zacharias SR, Brehm SB, Weinrich B, Kelchner L, Tabangin M, de Alarcon A (2016). A Feasibility of Clinical Endoscopy and Stroboscopy in Children with Bilateral Vocal Fold Lesions. *Am J Speech Lang Pathol*, 25(4):598–604. [PubMed: 27893084]
9. Hartnick C, Zeitels S (2005). Pediatric video laryngo-stroboscopy. *International Journal of Pediatric Otorhinolaryngology*, 69(2), 215–219. [PubMed: 15656955]
10. Sataloff RT, Spiegel JR, Carroll LM, Schiebel BR, Darby KS, & Rulnick R (1988). Stroboscopy in professional voice users: results and clinical value. *Journal of Voice*, 1(4), 359–364.
11. Mehta DD, Deliyski DD, & Hillman RE (2010). Commentary on why laryngeal stroboscopy really works: Clarifying misconceptions surrounding Talbot’s law and the persistence of vision. *Journal of Speech, Language, and Hearing Research*, 53(5), 1263–1267.

12. Powell ME, Deliyski DD, Hillman RE, Zeitels SM, Burns JA, Mehta DD (2016). Comparison of Videostroboscopy to Stroboscopy Derived From High-Speed Videoendoscopy for Evaluating Patients With Vocal Fold Mass Lesions. *American journal of speech-language pathology*, 25(4): 576–589. [PubMed: 27716854]
13. Bonilha HS, Deliyski DD (2008). Period and glottal width irregularities in vocally normal speakers. *Journal of Voice*, 22:699–708. [PubMed: 18031989]
14. Bonilha H, Deliyski D, Gerlach T (2008). Phase asymmetries in normophonic speakers: visual judgments and objective findings. *Am L Speech-Language Pathol*, 17:367–376.
15. Yan Y, Damrose E, Bless D (2007). Functional analysis of voice using simultaneous high-speed imaging and acoustic recordings. *Journal of Voice*, 21, 604–616. [PubMed: 16968665]
16. Shaw HS, Deliyski DD (2008). Mucosal wave: a normophonic study across visualization techniques. *J of Voice*, 22(1):23–33. [PubMed: 17014988]
17. Patel. R, Dailey. S, Bless. D (2008). Comparison of high-speed digital imaging with stroboscopy for laryngeal imaging of glottal disorders. *Ann Otol Rhinol Laryngol*, 117, 413–424. [PubMed: 18646437]
18. Deliyski D, Powell M, Zacharias S, de Alarcon A, Gerlach T (2015). Experimental investigation on minimum frame rate requirements of high-speed videoendoscopy for clinical voice assessment. *Biomedical Signal Processing and Control*, 17, 21–28. [PubMed: 28989342]

Figure 1.
Percent change in clinical ratings of vibratory features and diagnosis: videostroboscopy to high-speed videoendoscopy.

Table 1.

Demographics and clinical characteristics of the 151 study patients.

Variable	Value	Percent, %
Sample characteristics:		
Age - median	51	
Age range	22 to 84	
Sample size	151	100.0
Female	99	65.6
Male	52	34.4
Primary Diagnosis:		
Lesions ^a (total)	38	25.2
Bilateral	18	11.9
Unilateral with a reactive lesion	13	8.6
Unilateral	7	4.6
Edema	28	18.5
Dysphonia/voice disturbance	27	17.9
Laryngeal spasm	14	9.3
Paresis (bilateral)	11	7.3
Paresis (unilateral)	11	7.3
Paralysis (unilateral)	7	4.6
Reflux	5	3.3
Vocal fold weakness/bowing	4	2.7
Chronic laryngitis	2	1.3

^a Lesions = nodules, cysts, polyps, leukoplakia, granuloma

Table 2:

Features changed: videostroboscopy to HSV.

Features	Themes
Mucosal Wave	- increase in mucosal wave
Amplitude of Vibration	- increase in amplitude
Glottal Closure	- refinement in ability to describe closure pattern
Phase closure	- harsher initiation of closure
Glottal Edge	- refinement of mucosal swelling
Phase Symmetry	- either increase or decrease
Change in Diagnosis	- bowing (increased weakness) - increased movement of vocal folds ruling out paralysis - vocal fold appearance
Compression	- none noted
Level of Approximation	- none noted
Other	- visible onset and offset - HSV shows vibratory characteristics best - mucus better visualized on HSV - reflux – inconsistent – sometimes better seen on strobe and other times better seen on HSV

Table 3.

Description of the cases for which the additional assessment via high-speed videoendoscopy resulted in a 7% diagnostic change: 11 cases of total 151 patients included in the study.

Sex	Initial diagnosis	Final diagnosis	HSV finding
F	Mucus	Unilateral lesion	Increased mucosal wave Increased amplitude Glottic closure
F	Unilateral lesion	Bilateral lesions	Edge of vocal fold
F	Polyp	Cyst	Increased mucosal wave Increased amplitude Glottic closure
F	Bilateral Sulci	No sulci	Mucosal wave Straight vocal fold edges
F	Weakness	Adductor spasms	Increase amplitude Glottic closure
F	Edema	Lesion, bowing	Increase mucosal wave Glottic closure
M	Lesion (no movement)	Lesion (movement)	Increased mucosal wave Increased amplitude
F	Lesion	Mucus	Increased mucosal wave Increased amplitude Glottic closure
F	Mucus	Bilateral lesions	Increased mucosal wave Increased amplitude Onset/offset
M	Nodule	Cyst	Increase mucosal wave Increased amplitude Large anterior gap
F	Hoarse	Tremor/mixed SD	Tremor