

Prospective trial of voice outcomes after thyroidectomy: Evaluation of patient-reported and clinician-determined voice assessments in identifying postthyroidectomy dysphonia

Alexander Stojadinovic, MD,^a Leonard R. Henry, MD,^b Robin S. Howard, MA,^c
Joyce Gurevich-Uvena, MA,^d Matthew J. Makashay, PhD,^d George L. Coppit, MD,^e
Craig D. Shriver, MD,^a and Nancy P. Solomon, PhD,^d Washington, DC, and Bethesda, Md

Background. Reliable voice grading systems to identify postoperative voice dysfunction by surgeons are needed.

Purpose. To examine the utility of patient-reported and clinician-determined voice assessment in identifying postthyroidectomy voice dysfunction.

Patients and Methods. Fifty patients enrolled in a prospective observational trial evaluating voice function perioperatively by patient-reported symptoms (Voice Case History [VCHx]) and perceived voice handicap (Voice Handicap Index [VHI]), clinician-determined judgment of voice quality (Consensus Auditory-Perceptual Evaluation—Voice [CAPE-V]), and laryngeal examination via video laryngoscopy (VLS). Voice dysfunction at first postoperative visit in symptomatic patients was defined by objective laryngeal abnormalities on VLS. Postoperative changes from baseline in voice parameters were compared between patients with and without voice dysfunction using the Wilcoxon rank sum test. Receiver operating characteristics were evaluated to determine area under the curve (AUC) for tested parameters.

Results. Eight (16%) had early transient and 1 (2%) had permanent postoperative voice dysfunction. VCHx symptoms had negative (NPV) and positive (PPV) predictive values of 96%–100% and 39%–53%, respectively for voice dysfunction. The rating of overall severity from the CAPE-V was highly predictive (AUC = 0.96), and a change in severity from preoperative baseline $\geq 20\%$ at 1–2 weeks had a PPV of 86% and NPV of 95% for postoperative dysphonia. Patient-reported total VHI score was most predictive (AUC = 0.97) and a change in VHI from preoperative baseline ≥ 25 early postoperatively had a PPV of 88% and NPV of 97% for postoperative dysphonia.

Conclusion. Patient self-assessment of voice handicap using the VHI reliably identifies voice dysfunction after thyroidectomy. Patients with a change in VHI ≥ 25 from preoperative baseline warrant early referral to speech pathology and laryngology. (Surgery 2008;143:732-42.)

From the Department of Surgery, Division of Surgical Oncology,^a Walter Reed Army Medical Center, and the United States Military Cancer Institute, Washington, DC; the Department of Surgery,^b National Naval Medical Center, Bethesda, Maryland; the Department of Clinical Investigation, Division of Biostatistics,^c Walter Reed Army Medical Center, Washington, DC; the Department of Surgery, Army Audiology and Speech Center,^d Walter Reed Army Medical Center, Washington, DC; and the Department of Surgery, Division of Otolaryngology, Head and Neck Surgery,^e Walter Reed Army Medical Center, Washington, DC

The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or reflecting the views of the Department of the Army, Department of the Navy, or the Department of Defense.

Accepted for publication December 7, 2007.

Reprint requests: Alexander Stojadinovic, MD, Department of Surgery, Walter Reed Army Medical Center, 6900 Georgia

Avenue, Room 5C27A, NW, Washington, DC 20307. E-mail: alexander.stojadinovic@na.amedd.army.mil.

0039-6060/\$ - see front matter

© 2008 Mosby, Inc. All rights reserved.

doi:10.1016/j.surg.2007.12.004

THE INTRICATE AND DELICATE ANATOMIC STRUCTURES FOR VOICE PRODUCTION lie in proximity to the thyroid gland and are vulnerable during operations on the thyroid. Voice disturbances after thyroidectomy have been attributed traditionally to direct injury to the efferent laryngeal nerves sustained during the course of operation. Laryngeal nerve injury is a disturbing, functionally limiting sequela of thyroidectomy; it can engender considerable impairment in the patient's ability to perform daily, social and work-related tasks. Laryngeal nerve-related morbidity is particularly disabling for those reliant on voice to fulfill their professional duties. Singers and actors rely on the endurance and quality of voice. Other professional voice users, such as public speakers, teachers, and coaches, also depend on vocal staying power and effectiveness.

Functional voice outcome after thyroidectomy is highly dependent on surgical technique. Much like that of complex thoracic, pancreatic, and hepatobiliary operations, the volume–outcome relationship holds true for thyroidectomy.¹⁻³ Laryngeal nerve preservation during thyroidectomy is imperative for sustaining voice function; however, not all postoperative voice disturbances are related to nerve injury. The results of our prior clinical trial, which characterized postthyroidectomy functional voice changes, indicated that voice alterations are common after operation despite preservation of the laryngeal nerves.⁴ About one third of patients experience early and 14% have late voice symptoms after thyroidectomy. Other researchers have also confirmed that a high proportion of patients suffer postoperative deterioration in voice despite absence of objective signs of nerve injury.⁵⁻⁷ This has led some to question the contributory role of other mechanical factors (strap muscle division or laryngotracheal fixation, surgical trauma to the cricothyroid muscle or cricoarytenoid joint, endotracheal intubation, etc.) to postoperative dysphonia.⁸⁻¹⁰

Multidimensional objective acoustic, aerodynamic, and stroboscopic functional voice assessment in the modern speech pathology laboratory characterizes the anatomy and pathophysiology of postthyroidectomy voice dysfunction. However, specialized voice testing does not satisfactorily assess the degree of voice disability, and fails to account for why patients with similar vocal deficits experience markedly different voice handicaps. Conversely, identification of patients with significant postoperative voice dysfunction through straightforward, reliable grading systems utilized by nonexperts is needed. If such clinical tools had high predictive value, then they would contribute to improved resource utilization and identification of the small percentage of postthyroidectomy patients

who could benefit from early speech pathology referral and testing. Those few patients found to have functionally limiting laryngeal nerve dysfunction would be candidates for temporary vocal fold medialization procedures (injection medialization laryngoplasty), given the prolonged period of time required for idiopathic reinnervation, for which referral to laryngology would be appropriate.

Voice-related quality-of-life instruments have been applied with increasing frequency to allow assessment of the physical, functional, and affective aspects of voice disability. The Voice Handicap Index (VHI) is a brief, simple, and reliable outcome instrument increasingly used to monitor response to voice therapy, which has been validated for revealing the biopsychosocial impact of voice disorders.¹¹ The VHI relies on a patient's self-assessment of perceived severity of voice handicap.¹²⁻¹⁴ It consists of 30 statements that are rated from 0 to 4, for a maximum score of 120. The VHI has been used in both the clinical and research realms for the purpose of evaluating adult and pediatric patient responses to therapy for medical conditions and surgical interventions.¹⁵⁻²⁰ Despite this widespread application in clinical situations pertaining to the head and neck region in general, the VHI has been employed infrequently as a research tool or with clinical intent in patients undergoing thyroidectomy in particular. Only 2 previous reports available in the literature have addressed use of the VHI for patients undergoing thyroid operations; however, neither addressed the predictive role of the VHI in the perioperative setting for postthyroidectomy dysphonia.^{6,21} Nonetheless, the VHI is gaining widespread acceptance as a useful and simple tool, evidenced by its translation into at least 5 languages (Chinese, Dutch, French, Hebrew, and Spanish), and the tool's increasing popularity in the voice literature in general.

The ability of the VHI to predict dysphonia after thyroidectomy is undefined. Similarly, the utility of patient-reported symptoms (Voice Case History) and clinician judgments of voice quality (Consensus Auditory-Perceptual Evaluation-Voice [CAPE-V]) remains to be determined, although these assessments are utilized routinely in speech pathology clinics for determining postthyroidectomy voice dysfunction.^{22,23} The optimal clinical tool to identify significant alterations in postthyroidectomy voice would be easy to use, noninvasive, simple to interpret, inexpensive, and would not require physician interpretation. In the search of such a tool, this prospective clinical study examines the utility of Voice Case History, CAPE-V, and VHI in determining objective voice dysfunction after thyroid surgery.

MATERIAL AND METHODS

Study design. A prospective, single-arm, observational cohort trial addressing voice outcomes after thyroidectomy was begun under Walter Reed Army Medical Center Institutional Review Board approval from October 2004 and is currently ongoing. All subjects enrolled into the study voluntarily agreed to participate and gave written informed consent.

The principal aims of the study were to characterize changes in voice after thyroidectomy, specifically, to determine the natural history and level of functional impairment of voice after thyroid surgery using patient- and clinician-based assessment as well as objective laryngeal image recordings. This report includes data from the first 46 study subjects enrolled in the trial with preoperative and ≥ 1 postoperative visits.

By participating in this study, all patients undergoing thyroidectomy received a comprehensive voice evaluation, incorporating the following measures included in this report:

1. Patient-reported symptoms (Voice Case History);
2. Auditory perceptual measures of voice (i.e., judgments of voice quality by experienced clinicians using the CAPE-V);
3. Visual perceptual measures of anatomic laryngeal images (judgments of videostroboscopic laryngeal recordings by experienced clinicians); and,
4. Self-report of the physical, functional, and emotional impact of voice dysfunction using the VHI.

Postoperative voice dysfunction at first (early, 1-2 week) clinic visit after operation was defined as objective laryngeal abnormalities (e.g., arytenoid asymmetry, reduced vocal cord lengthening, impaired vocal cord mobility/adduction, arytenoid subluxation) evident on video laryngostroboscopy (VLS) in symptomatic patients reporting change in voice. Other assessments not reported herein involved acoustic analysis of voice and aerodynamic testing. This voice assessment was intended to assist investigators in identifying and tracking measures indicative of significant deficits in voice production that could potentially result in more effective and efficient efforts to improve the postthyroidectomy voice, or to prevent the adoption of maladaptive compensatory behaviors through early referral for speech therapy.

Study population. Participants (age, 18-77 years) included 19 male and 31 female military health care beneficiaries who were scheduled to undergo thyroid resection. The primary indication for operation in the 20% of patients with coexisting

physiologic thyroid disease (eg, lymphocytic thyroiditis) was a dominant nodule(s). Subjects had no preoperative laryngeal pathology or prior neck or thyroid surgery. After operation, 46 returned for the early postoperative assessment, 35 after 3 months, and 30 after 6 months.

Functional voice outcome assessment. At each of 4 study time points (before operation, and 1-2 weeks and 3 and 6 months after thyroidectomy), patients underwent real-time laryngeal imaging to assess laryngeal function as part of standard-of-care assessment. At the same scheduled study visits, the Voice Case History and VHI were administered. An experienced voice researcher and speech-language pathologist judged the sustained vowels and short sentences included in the CAPE-V uttered by the participants in the clinical setting. CAPE-V ratings are made according to 10-cm undifferentiated lines (visual analog scale) for each of the following: overall severity (grade), roughness, breathiness, strain, pitch, and loudness.

VLS was conducted in the current study primarily to correlate changes in voice with objective signs of vocal dysfunction. Laryngeal imaging during phonation was performed with fiberoptic flexible laryngoscopy utilizing halogen and stroboscopic light sources. VLS assessed the appearance and mobility of the vocal folds utilizing a standardized vocal task protocol. During the laryngeal examination, tasks included phonation of sustained vowels at low, comfortable, and high pitch and loudness levels, ascending and descending pitch glides, syllable repetitions, and short sentence productions. An experienced examiner made ratings of laryngeal structure and function according to a standard grading protocol. Assessment involved a rating of normal, mild, moderate, or severe for supraglottic activity, right and left vocal fold (vocal process of the arytenoid) mobility and symmetry, vocal fold edge, lengthening and mucosal wave, phase symmetry, and glottic closure pattern. In questionable cases, 3 experienced raters reviewed the images and came to a consensus opinion regarding final observations and diagnosis.

The Voice Case History involved a brief interview during which subjects answered specific questions about their primary voice symptoms and social and occupational vocal demands. The subjects stated whether their voices were normal and typical for them. They were further queried about their perceived presence of hoarseness, vocal fatigue, difficulty projecting, difficulty singing, and difficulty phonating at high pitches (see [Appendix](#) for Voice Case History).

The VHI is a brief, validated outcome instrument that relies on a patient's self-assessment of perceived

severity of voice handicap.¹¹ The patient's self-appraisal consists of 10 items in each of 3 domains: functional, physical, and emotional. The subject responds according to his or her perception of the impact of the voice disorder on daily activities.

Auditory-perceptual data was obtained from acoustic data recordings taken during each study subject visit. Voice samples selected for analysis were a sustained vowel "ee" and "ah" (3–5 seconds, on a comfortable breath), and 6 sentences read aloud from the CAPE-V. A speech-language pathologist/voice researcher experienced in judging abnormal voice rated each sustained-vowel sample for the following parameters on a visual analog scale (CAPE-V; 10 cm undifferentiated line) from normal (left extreme) to most abnormal (right extreme): overall severity (grade) of dysphonia, roughness, breathiness, strain, pitch, and loudness. Additional features were noted as necessary (eg, fry, pitch instability, diplophonia).²²

Statistical analysis. Differences in baseline characteristics were evaluated using Fisher's exact test for categorical variables and the Wilcoxon rank-sum test for continuous or ordinal data. Changes from baseline to follow-up time points postoperatively were compared between groups using the Wilcoxon rank-sum test. Spearman's rho correlation coefficient was calculated to examine nonparametric measures of association between variables. Diagnostic accuracy of voice parameters were presented using receiver operating characteristic curves and described using the area under the curve (AUC), as well as sensitivity, specificity, positive predictive values (PPV) and negative predictive values (NPV) together with 95% confidence intervals for these proportions. Continuous data were presented as either medians with ranges or mean values with standard deviations ($\pm$ SD) in the text and tables. Logistic regression was used to compute the relationship between voice parameters (CAPE-V severity and VHI) and the probability of postoperative dysphonia. Significance was determined at the $P \leq .05$ level. Statistical analyses were carried out using JMP (5.1 SAS Institute, Inc., Cary, NC) and SPSS statistical software (version 13.0, SPSS, Inc., Chicago, Ill).

RESULTS

Patients. Fifty patients were enrolled in this study, predominantly female, with mean age of 48 years. As seen in the demographic data provided in Table I, most patients were neither smokers nor voice professionals, and half underwent less than total thyroidectomy. Nearly all operations were conducted under general anesthesia without difficulty intubating the patient, and without the requirement for any strap

muscle division. Nearly one third of operations were undertaken for differentiated early stage thyroid carcinoma, the remainder for benign disease. Mean operating time was 152 minutes (SD = 78 minutes), and median resected thyroid volume was 65 cm³ (range, 6–2,873 cm³).

Postoperative voice dysfunction (dysphonia). Eight subjects (16%) had early transient postoperative voice dysfunction after operation, which was defined as objective laryngeal abnormalities evident on VLS in patients reporting change in voice (Table II). One (2%) had permanent voice dysfunction >6 months after operation owing to transection of the recurrent laryngeal nerve (RLN) during resection of a large posterior multinodular goiter, despite immediate nerve repair and voice rehabilitation. Speech therapy was recommended to the 8 patients with postoperative dysphonia; however, only 5 of the 8 patients pursued voice rehabilitation. The mean number of abnormal VLS parameters at time of first postoperative visit (1–2 weeks) for those with and without voice dysfunction was 6.6 ± 0.5 and 0.9 ± 0.2 , respectively ($P < .001$).

Voice case history: Patient reported. Reported symptoms [voice not typical ($P < .0005$), hoarseness ($P < .0005$), pitch loss ($P = .002$), voice fatigue ($P = .004$)] at 1–2 weeks after operation each correlated significantly with voice dysfunction. However, despite a high NPV (96–100%), Voice Case History had a low PPV for postoperative voice dysfunction (39–53%; Table III).

Consensus auditory-perceptual evaluation of voice: Clinician determined. Auditory perceptual measures of voice according to the CAPE-V require the judgment of voice quality by an experienced voice clinician. Four CAPE-V parameters significantly correlated with postoperative voice dysfunction, overall severity (grade) of dysphonia, habitual pitch, habitual loudness, and roughness ($P < .001$). Mean change from preoperative baseline in these CAPE-V measures at first postoperative (1–2 weeks) visit for those with and without voice dysfunction was: severity, 48.5 ± 30.6 versus -0.8 ± 13.4 ($P < .0005$); pitch, 20.9 ± 19.9 versus 0.7 ± 7.2 ($P < .0005$); loudness, 33.9 ± 23.3 versus -0.2 ± 8.3 ($P < .0005$); roughness, 33.1 ± 31.7 versus 0.8 ± 16.2 ($P = .001$). All 4 parameters returned to baseline, on average, by 3 months after operation (Fig 1). The receiver operating characteristics were most predictive for the rating of perceived overall voice severity (grade; AUC, 0.96; Table IV). A change in severity from preoperative baseline ≥ 20 , 1–2 weeks after operation had a PPV and NPV for postoperative dysphonia of 86% and 95%, respectively (Table III).

Table I. Patient and thyroid-related characteristics of 50 consecutive patients enrolled in voice outcomes study

Characteristics	No. patients	%
Age (y), median (range)	49 (18–77)	
Gender		
Male	19	38
Female	31	62
Ethnicity		
Caucasian	32	64
African American	13	26
Latino	4	8
Asian	1	2
Voice professional		
Yes	13	26
No	37	74
Tobacco use		
Never	30	60
Former	15	30
Current	5	10
Anesthesia		
General, endotracheal	47	94
Regional, superficial cervical block	3	6
Extent of thyroidectomy		
Lobe and isthmus	25	50
Total	25	50
Neuromonitoring of RLN		
Yes	9	18
No	41	82
Strap muscle division		
Yes	6	12
No	44	88
Postoperative voice dysfunction		
Yes	8	16
No	42	84

RLN, recurrent laryngeal nerve.

Voice Handicap Index: Patient reported. The patient's self-assessment of perceived severity of voice handicap was analyzed on the basis of responses to 30 questions divided equally across 3 VHI domains (Functional, Physical and Emotional; each question having a maximum score of 4). Mean change from baseline in total VHI score (range, 0–120) at first postoperative (1–2 weeks) visit for those with and without voice dysfunction was 59.6 ± 33.4 versus 3.8 ± 15.2 , respectively ($P < .0005$; Fig 2). Significant changes were identified at the first postoperative visit (1–2 week) for the Functional (20.8 ± 12.0 vs. 1.6 ± 6.9 ; $P < .0005$), Physical (25.4 ± 10.3 vs. 2.7 ± 7.8 ; $P < .0005$), and Emotional (13.5 ± 13.8 vs. 1.2 ± 6.7 ; $P = .002$) domains of the VHI between patients with and without dysphonia. Significant changes in VHI scores were demonstrated at 3-month follow-up visits in patients with voice dysfunction, which returned

Table II. Patient and thyroid-related characteristics of patients enrolled in voice outcomes study stratified by postoperative voice dysfunction

Characteristic	Voice dysfunction		P value
	Yes	No	
Total patients, n (%)	8 (16)	42 (84)	
Age (y), mean $\pm$ SD	50.0 ± 17.8	47.8 ± 13.3	.62
Gender, n (%)			.13
Male	1 (12)	18 (43)	
Female	7 (88)	24 (57)	
Voice professional, n (%)			.99
Yes	2 (25)	11 (26)	
No	6 (75)	31 (74)	
Anesthesia, n (%)			.99
General, endotracheal	8 (100)	39 (93)	
Regional, superficial cervical block	0 (0)	3 (7)	
Extent of thyroidectomy, n (%)			.25
Lobe and isthmus	2 (25)	23 (55)	
Total thyroidectomy	6 (75)	19 (45)	
RLN neuromonitoring, n (%)			.99
Yes	1 (12)	8 (19)	
No	7 (88)	34 (81)	
Strap muscle division, n (%)			.99
Yes	1 (12)	5 (12)	
No	7 (88)	37 (88)	
Thyroid volume resected (cm ³), Mean $\pm$ SD	129 ± 108	212 ± 468	.89

RLN, recurrent laryngeal nerve; SD, standard deviation.

to baseline by postoperative month 6 (Fig 3). The receiver operating characteristics were most predictive for total VHI score (AUC, 0.97; Table IV; Fig 4). A change in VHI score from preoperative baseline ≥ 25 , 1–2 weeks after operation had a PPV and NPV for postoperative dysphonia of 88% and 97%, respectively (Table III).

Correlations of CAPE-V (clinical assessment) and VHI (patient-reported) scores. Change from baseline in CAPE-V parameters of overall severity (grade), pitch, and loudness at 1–2 weeks after operation correlated significantly with total change in VHI, as well as the individual index domains, Functional, and Physical ($P < .05$ for all comparisons); although, the Spearman's ρ correlation coefficient was <0.60 ($r = 0.31$ – 0.58) for all non-parametric correlations between auditory perceptual measures of voice according to CAPE-V and the patient's self-assessment of perceived severity of voice handicap using the VHI. Only changes

Table III. Predictive value of Voice Case History (VCHx), Voice Handicap Index (VHI), and Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) Severity Score for postoperative voice dysfunction

Category	Sensitivity	Specificity	PPV	NPV
	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)
VCHx symptom				
Voice not typical	100 (71–100)	68 (52–81)	40 (21–61)	100 (88–100)
Hoarse	100 (71–100)	82 (66–91)	53 (30–75)	100 (90–100)
Pitch loss	88 (51–99)	76 (59–87)	44 (23–67)	97 (81–99)
Voice fatigue	88 (51–99)	70 (54–83)	39 (20–61)	96 (80–99)
Change in VHI ≥ 25 , 1–2 weeks postop	88 (51–99)	97 (84–99)	88 (51–99)	97 (84–99)
Change in CAPE-V Severity ≥ 20 , 1–2 weeks postop	75 (40–94)	97 (84–99)	86 (46–99)	95 (81–99)

PPV, predictive values; NPV, negative predictive values.

in Pitch and Loudness were significantly correlated with the Emotional domain ($r = 0.37$ and $r = 0.48$, respectively; $P < .014$). However, the combined change in CAPE-V severity from preoperative baseline ≥ 20 and change in total VHI score from preoperative baseline ≥ 25 , 1–2 weeks after operation, had a PPV and NPV for postoperative dysphonia of 86% and 95%, respectively (Fig 5).

DISCUSSION

The present study was conducted to determine the utility of simple bedside questionnaires and auditory perceptual analysis in identifying patients with significant dysphonia after thyroid resection. Postoperative voice dysfunction was defined as objective laryngeal abnormalities evident on VLS in symptomatic patients reporting change in voice. The findings of transient (16%) and permanent (2%) alteration in voice after thyroidectomy reflect the experience of other endocrine surgical units.

Although the Voice Case History had high NPV (96–100%), its low PPV limits its clinical utility. The rating of overall severity from the CAPE-V was highly predictive, and a change in severity from preoperative baseline ≥ 20 early postthyroidectomy had a PPV and NPV of 86% and 95%, respectively, for postoperative dysphonia. However, nonexperts should not utilize the CAPE-V; this tool is intended for use by an experienced speech-language pathologist (“a trained ear”) to make judgments about the various aspects of voice quality. The patient’s self-assessment of perceived severity of voice handicap in terms of the physical, functional, and emotional impact of voice dysfunction recorded on a standardized questionnaire, the VHI, was highly predictive of voice outcome in this study. A change in VHI score from preoperative baseline ≥ 25 early postoperatively had a positive and NPV of 88% and 97%, respectively, for postoperative

voice dysfunction. Combined CAPE-V and VHI provided no incremental predictive value.

These findings are important; the prevalence of thyroid disease, thyroid resections, and potential for morbidity are high. In the United States, approximately 15 million adults have unrecognized thyroid disease and roughly 11% of persons suffer significant endocrine, inflammatory, or neoplastic disorders of the thyroid gland.²⁴ Surgical resection of the thyroid gland is performed approximately 80,000 times per year in the United States alone.²⁵ The RLN is highly relevant in thyroid surgery because it innervates all of the intrinsic muscles of the larynx except the cricothyroid muscle, which is innervated by the external branch of the superior laryngeal nerve. Operative trauma to the RLN—the most feared complication of thyroidectomy—occurs in 2–13% of patients, and results in vocal fold paresis or paralysis, both of which can cause the voice to sound hoarse, breathy, and quiet.^{26,27} Inspiratory stridor may also occur in cases of bilateral vocal fold paralysis. Swallowing impairment may also ensue, leaving patients vulnerable to aspiration. Injury to the external branch of the superior laryngeal nerve occurs during thyroidectomy in up to 30% of patients and may result in profound morbidity for professional voice users owing to voice fatigue and impaired ability to project the voice and achieve a high pitch.^{28,29} Professional voice users, particularly singers and public speakers, who have an isolated superior laryngeal nerve deficit, may in fact have seemingly disproportionate vocal dysfunction when arytenoid abduction/adduction appears normal. Furthermore, functional voice disorders have been documented in the absence of major laryngeal nerve injury with as many as 30% of patients suffering early and 14% suffering lasting functional voice changes after thyroid resection.⁴ The high incidence of thyroidectomy coupled with the significant rates of postoperative voice morbidity

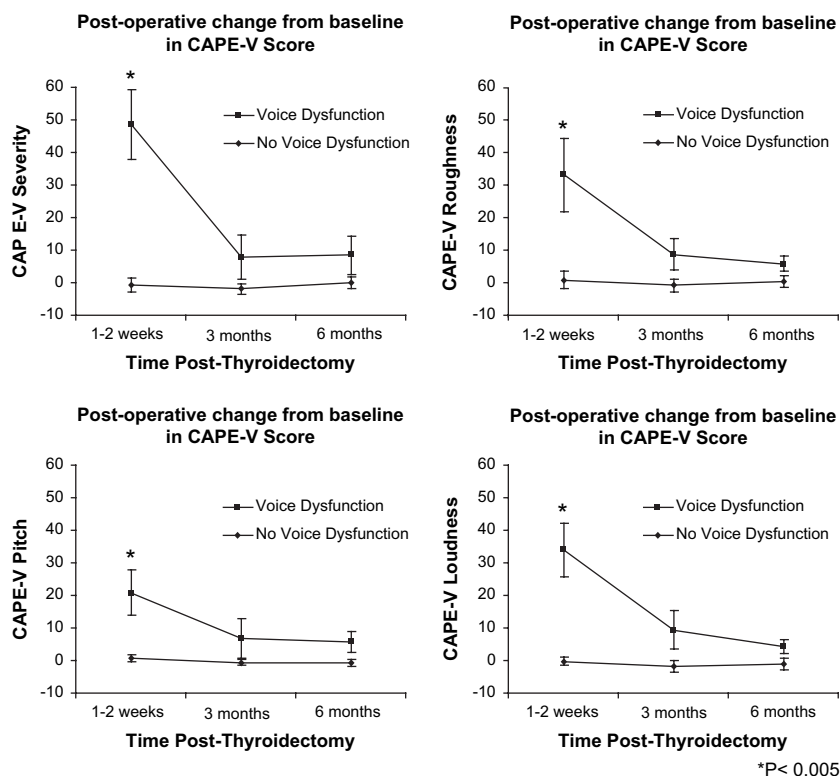


Fig 1. Mean changes ($\pm$ standard errors) in CAPE-V overall severity (grade) of dysphonia, habitual pitch, habitual loudness and roughness (mean) in patients with and without dysphonia, 1–2 weeks, and 3 and 6 months after thyroidectomy. * $P < .005$; voice dysfunction versus no voice dysfunction.

Table IV. Area under the receiver operating characteristics curve (AUC) for Voice Handicap Index (VHI) and Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). Change in scores at first (1–2 week) postoperative visit.

Test result variable	AUC
CAPE-V Severity	0.96
CAPE-V Pitch	0.88
CAPE-V Loudness	0.91
CAPE-V Roughness	0.85
Total VHI Score	0.97
VHI-Functional	0.94
VHI-Physical	0.94
VHI-Emotional	0.82

reported in the literature suggests that dysphonia after thyroidectomy (despite widespread awareness of recurrent nerve injury) remains a significant and likely underappreciated sequela of the operation.

A standard practice pattern is lacking among thyroid surgeons for identifying patients suffering early postoperative dysphonia, and no decision tool has been applied generally for predicting which patients are at risk for persistent dysphonia after thyroidectomy. Surgical practice setting determines

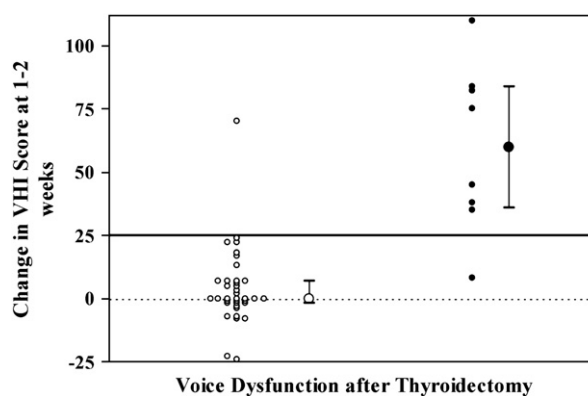


Fig 2. Changes in total VHI score (median and interquartile range) from preoperative baseline in patients with and without dysphonia, 1–2 weeks after thyroidectomy.

availability of sophisticated laryngoscopic, acoustic, and aerodynamic voice testing technology, and many surgeons will have variable access to this instrumentation. The level of personal comfort with office-based direct or indirect laryngoscopy and ability to interpret subtle vocal fold findings indicating vocal dysfunction is likewise inconsistent among thyroid surgeons.

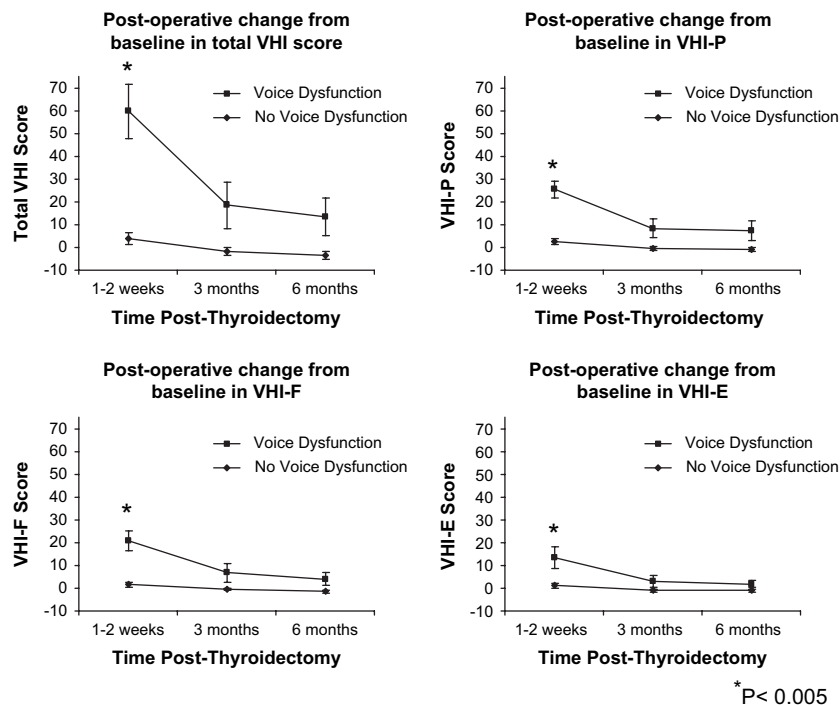


Fig 3. Mean changes ($\pm$ standard errors) in total VHI score (mean) and Functional, Physical and Emotional VHI domain scores from preoperative baseline in patients with and without dysphonia, 1–2 weeks and 3 and 6 months after thyroidectomy. * $P < .005$: voice dysfunction versus no voice dysfunction.

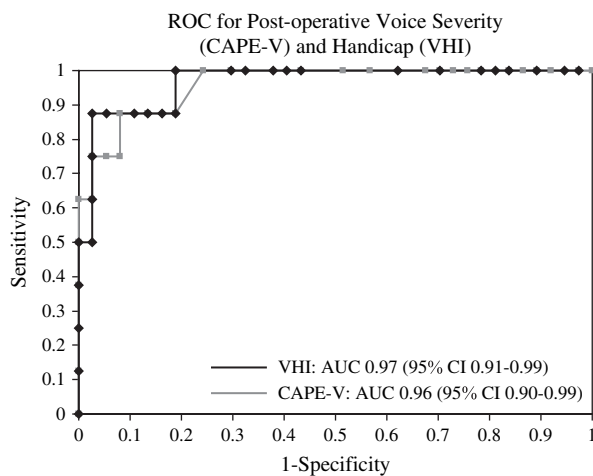


Fig 4. Receiver operating characteristics (ROC) for early (1–2 weeks) postoperative CAPE-V overall severity (grade) and total VHI score changes from preoperative baseline.

Theoretically, the optimal surveillance strategy to detect functional voice changes after thyroidectomy would be noninvasive, inexpensive, quick, and easy to use, both in terms of completion by patients and interpretation by surgeons. To be clinically relevant, such a strategy must possess excellent PPV and NPV. It would function efficiently and equally well in the small community hospital as in a tertiary

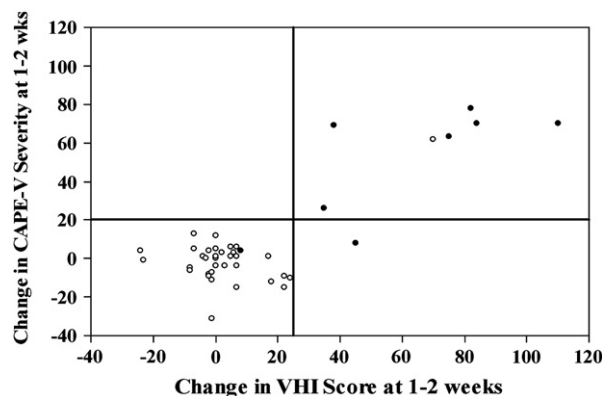


Fig 5. Plot of change in CAPE-V overall severity (grade) versus change in total VHI score from preoperative baseline, 1–2 weeks after thyroidectomy. Threshold change in these parameters was established using logistic regression at 20 and 25 for CAPE-V severity and VHI score, respectively.

referral center. To that end, we report here our findings in search of such a postthyroidectomy surveillance strategy.

Results from laryngoscopic examination conducted in a clinical voice research center were correlated with simple clinician- and patient-derived assessments of voice disturbance after thyroidectomy. Based on the results of this study, the VHI

appears to fulfill the intended postoperative surveillance strategy. The VHI proved to be a quick test easily completed in the clinic setting. Moreover, a change from baseline of ≥ 25 points on the VHI score after thyroidectomy was highly predictive of vocal dysfunction based on visual examination of the larynx during phonation, with a PPV of 88% and NPV of 97%.

The excellent PPV and NPV of the change in VHI found in the present study may allow for this questionnaire to be incorporated in the postoperative decision making process for patients undergoing thyroidectomy. Those with no voice complaints reported by either Voice Case History or with minimal (< 25 point) change in VHI score postoperatively are highly unlikely to suffer dysphonia or functional voice limitation at any time in the postoperative period. Long-term follow-up for vocal concerns for these patients appears unnecessary. Based on our findings, those with changes in the VHI of ≥ 25 points after thyroidectomy are predicted to suffer dysphonia in nearly 90% of cases, which can be functionally limiting for months after operation. Although our data suggest that return to normal voice is anticipated in most patients between 3 and 6 months, early referral to a voice specialist may still be justified for more sophisticated assessment and interventions to prevent the development of maladaptive compensatory vocal behaviors that may persist even after the resolution of underlying neuropathology. Those patients that screen positive by VHI (with changes in the VHI of ≥ 25 points after thyroidectomy) and are confirmed to have phonatory decrements via objective voice measures can be referred for laryngeal electromyography, which has been shown to be a useful adjunct to the diagnosis, management, and prediction of poor functional outcome in patients with dysphonia and impaired vocal fold movement.^{30,31}

Optimal treatment options for unilateral vocal fold paresis (partial interruption of nerve impulse resulting in weak or abnormal motion of laryngeal muscle(s), namely, partial loss of or impaired vocal fold movement) includes temporary vocal fold medialization (injection medialization laryngoplasty), usually by injecting resorbable material (such as Gelfoam paste, collagen, fat, calcium hydroxylapatite, nonanimal stabilized hyaluronic acid, or micronized particulate form of decellularized human donor tissue, Alloderm) into the lateral aspect of the paretic vocal fold,³²⁻³⁴ in conjunction with behavioral voice therapy.³⁵⁻³⁸ Permanent dysphonia resulting from unilateral vocal fold paralysis (owing to total interruption of nerve impulse

resulting in no vocal fold movement) is typically treated surgically with a permanent injection laryngoplasty (eg, vocal fold injection of nonresorbable Teflon, expanded polytetrafluoroethylene, or polyacrylamide hydrogel), medialization thyroplasty, arytenoid adduction or arytenoidopexy, RLN electrostimulation, or reinnervation procedures, usually in conjunction with behavioral voice therapy to facilitate patients learning how to optimize residual laryngeal function and avoid maladaptive compensatory behaviors.³⁹⁻⁴⁴ Although we have found in this study that those with a change of ≤ 25 points in the VHI score from baseline are unlikely to suffer long-term dysphonia, we believe reevaluation at a 3-month interval in symptomatic patients to be reasonable given our modest data set (see [Appendix](#)).

The auditory-perceptual evaluation of voice by an experienced speech-language pathologist before and after thyroid resection using the CAPE-V identifies patients with early and significant postthyroidectomy dysphonia according to perceived overall voice severity (grade), habitual pitch, loudness, and roughness. However, the CAPE-V is not patient derived, and relies heavily on voice expertise, making it impractical in surgery practice where the proficiency to interpret it is not readily available. Given the high predictive value observed with the perioperative change in VHI in this study, particularly considering the practicality issues associated with the CAPE-V mentioned above, we favor the utilization of the VHI by most thyroid surgeons before and after thyroidectomy for identifying patients with severe postoperative dysphonia.

It is important to acknowledge that although sophisticated acoustic, aerodynamic, and videostroboscopic data were recorded in this prospective cohort study, we have not presented or analyzed these data here. It is our intention to have the patient-reported and clinician-derived data presented in this study assist thyroid surgeons in all practice settings, and to encourage thyroid surgeons to adopt the VHI given its simplicity, practicality, utility, and potential value in clinical decision making. The overall clinical impact could be significant, given the potential for streamlining and improving postoperative care, and containing costs through selective referral for specialized voice testing and care for those patients most likely to benefit. Moreover, the VHI provides surgeons a way to better inform their patients about not only the anticipated postoperative voice outcomes but also natural history of changes in voice that follow thyroidectomy.

Important limitations of this study deserve mention. This analysis is based on a relatively small study population and even smaller number

of patients with postoperative dysphonia. Although we would consider these preliminary data supportive of the proposed clinical utilization of the VHI, prospective validation with a larger sample size is warranted before the widespread acceptance of our findings and implementation of our recommendations. Moreover, the clinicians who diagnosed dysphonia on the basis of objective VLS criteria in this study, although experts in the field, were not blinded to the subjects' identity at time of assessment of vocal fold abnormalities. For that reason, one cannot exclude recall bias as a potential factor in their assessments.

Continued study of functional voice changes after thyroid operation is warranted and its characterization important for timely rehabilitation, particularly in professional voice users. We think that early referral to a voice specialist is warranted to help patients with postoperative vocal dysfunction to learn how to best use their optimal voice and avoid developing maladaptive compensatory vocal behaviors while the voice recovers. This study provides insights into the utility of patients' self-assessment of voice handicap. Patient-reported severity of voice impairment using the standardized VHI reliably identifies voice dysfunction after thyroidectomy. Using the VHI in this setting may prove useful in selecting patients who may benefit from referral to a voice specialist. Patients with a change in VHI score ≥ 25 from preoperative baseline should be considered for early referral to speech pathology. This study represents an important step forward in postoperative decision making and treatment planning by establishing the basis for early recognition of vocal disturbances and selective referral for consideration of voice rehabilitation. This prospective trial is ongoing to further validate the VHI in patients undergoing thyroid operations.

REFERENCES

1. Fong Y, Gonen M, Rubin D, Radzyner M, Brennan MF. Long-term survival is superior after resection for cancer in high-volume centers. *Ann Surg* 2005;242:540-4.
2. Birkmeyer JD, Dimick JB, Staiger DO. Operative mortality and procedure volume as predictors of subsequent hospital performance. *Ann Surg* 2006;243:411-7.
3. Sosa JA, Bowman HM, Tielsch JM, Powe NR, Gordon TA, Udelsman R. The importance of surgeon experience for clinical and economic outcomes from thyroidectomy. *Ann Surg* 1998;228:320-30.
4. Stojadinovic A, Shaha AR, Orlikoff RF, et al. Prospective functional voice assessment in patients undergoing thyroid surgery. *Ann Surg* 2002;236:823-32.
5. McIvor NP, Flint DJ, Gillibrand J, Morton RP. Thyroid surgery and voice-related outcomes. *Aust N Z J Surg* 2000;70:179-83.
6. de Pedro Netto I, Fae A, Vartanian JG, et al. Voice and vocal self-assessment after thyroidectomy. *Head Neck* 2006;28:1106-14.
7. Lombardi CP, Raffaelli M, D'Alatri L, et al. Voice and swallowing changes after thyroidectomy in patients without inferior laryngeal nerve injuries. *Surgery* 2006;140:1026-32.
8. Hong KH, Kim YK. Phonatory characteristics of patients undergoing thyroidectomy without laryngeal nerve injury. *Otolaryngol Head Neck Surg* 1997;117:399-404.
9. Loucks TMJ, Poletto CJ, Saxon KG, Ludlow CL. Laryngeal muscle responses to mechanical displacement of the thyroid cartilage in humans. *J Appl Physiol* 2005;99:922-30.
10. Kark AE, Kissin MW, Auerbach R, Meikle M. Voice changes after thyroidectomy: role of the external laryngeal nerve. *Br Med J (Clin Res Ed)* 1984;289:1412-5.
11. Maertens K, de Jong FI. The voice handicap index as a tool for assessment of the biopsychosocial impact of voice problems. *B-ENT* 2007;3:61-6.
12. Jacobson BH, Johnson A, Grywalsky C, et al. The Voice Handicap Index (VHI): Development and Validation. *J Voice* 1998;12:540-50.
13. Webb AL, Carding PN, Deary IJ, et al. Optimizing outcome assessment of voice interventions, I: reliability and validity of three self-reported scales. *J Laryngol Otolaryngol* 2007;121:763-7.
14. Portone CR, Hapner ER, McGregor L, Otto K, Johns MM. Correlation of the Voice Handicap Index (VHI) and the Voice-Related Quality of Life Measure (V-RQOL). *J Voice* 2007;21:723-7.
15. Schindler A, Capaccio P, Maruzzi P, Ginocchio D, Bottero A, Otraviani F. Preliminary considerations on the application of the voice handicap index to paediatric dysphonia. *Acta Otorhinolaryngol Ital* 2007;27:22-6.
16. Piazza C, Bolzoni Villaret A, Redaelli De Zinis LO, Cattaneo A, Cocco D, Peretti G. Phonosurgery after endoscopic cordectomies: Delayed medialization techniques for major glottic incompetence after total and extended resections. *Eur Arch Otorhinolaryngol* 2007;264:1185-90.
17. Millqvist E, Bende M, Brynnel M, Johansson I, Kappel S, Ohlsson AC. Voice change in seasonal allergic rhinitis. *J Voice* 2007 May 16 [Epub ahead of print].
18. Schindler A, Cuccarini V, Bottero A, Dobe C, Capaccio P, Ottaviani F. Long-term vocal functional results after glottectomy: A multi-dimensional analysis. *Eur Arch Otorhinolaryngol* 2007;264:1039-44.
19. Spielman J, Ramig LO, Mahler L, Halpern A, Gavin WJ. Effects of an extended version of the Lee Silverman voice treatment on voice and speech in Parkinson's disease. *Am J Speech Lang Pathol* 2007;16:95-107.
20. Siupsinskiene N, Adamonis K, Toohill RJ. Quality of life in laryngopharyngeal reflux patients. *Laryngoscope* 2007;117:480-4.
21. Petro ML, Schweinfurth JM, Petro AB. Transcricothyroid, intraoperative monitoring of the vagus nerve. *Arch Otolaryngol Head Neck Surg* 2006;132:624-8.
22. 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.
23. Hogikyan ND, Sethuraman G. Validation of an instrument to measure voice-related quality of life (V-RQOL). *J Voice* 1999;13:557-69.
24. Canaris GJ, Manowitz NR, Mayor G, et al. The Colorado Thyroid Disease Prevalence Study. *Arch Intern Med* 2000;160:526-34.

25. Harness JK, Organ CHJ, Thompson NW. Operative experience of US general surgery residents in thyroid and parathyroid disease. *Surgery* 1995;118:1063-70.
26. Filho JG, Kowalski LP. Surgical complications after thyroid surgery performed in a cancer hospital. *Otolaryngol Head Neck Surg* 2005;132:490-4.
27. Roy AD, Gardiner RH, Niblock WM. Thyroidectomy and the recurrent laryngeal nerve. *Lancet* 1956;270:988-90.
28. Cernea C, Ferraz A, Fulani J, et al. Identification of the external branch of the superior laryngeal nerve during thyroidectomy. *Am J Surg* 1992;164:634-8.
29. Hurtado-Lopez LM, Pacheco-Alvarez MI, Montes-Castillo Mde L, Zaldivar-Ramirez FR. Importance of the intraoperative identification of the external branch of the superior laryngeal nerve during thyroidectomy: electromyographic evaluation. *Thyroid* 2005;15:449-54.
30. Heman-Ackah YD, Barr A. The value of laryngeal electromyography in the evaluation of laryngeal motion abnormalities. *J Voice* 2006;20:452-60.
31. Sittel C, Stennert E, Thumfart WF, Dapunt U, Eckel HE. Prognostic value of laryngeal electromyography in vocal fold paralysis. *Arch Otolaryngol Head Neck Surg* 2001;127:155-60.
32. Hartl DM, Travagli JP, Lebouilleux S, Baudin E, Brasnu DF, Schlumberger M. Clinical review: Current concepts in the management of unilateral recurrent laryngeal nerve paralysis after thyroid surgery. *J Clin Endocrinol Metab* 2005;90:3084-8.
33. Karpenko AN, Dworkin JP, Meleca RJ, Stachler RJ. Cymetra injection for unilateral vocal fold paralysis. *Ann Otol Rhinol Laryngol* 2003;112:927-34.
34. Saarinen A, Sovijärvi AR, Karhumäki L, Rihkanen H. Medicalization of paralyzed vocal fold does not increase respiratory resistance measured by impulse oscillometry. *Eur Arch Otorhinolaryngol* 2007;264:1323-7.
35. D'Alatri L, Galla S, Rigante M, Antonelli O, Buldrini S, Marchese MR. Role of early voice therapy in patients affected by unilateral vocal fold paralysis. *J Laryngol Otol* 2007;24:1-6.
36. Miller S. Voice therapy for vocal fold paralysis. *Otolaryngol Clin North Am* 2004;37:105-19.
37. Schindler A, Bottero A, Capaccio P, Ginocchio D, Adorni F, Ottaviani F. Vocal improvement after voice therapy in unilateral vocal fold paralysis. *J Voice* 2008;22:113-8.
38. Schindler A, Favero E, Nudo S, Albera R, Schindler O, Cavalot AL. Long-term voice and swallowing modifications after supracricoid laryngectomy: Objective, subjective, and self-assessment data. *Am J Otolaryngol* 2006;27:378-83.
39. Havas TE, Priestley KJ. Laser tenotomy and vocal process resection for bilateral midline vocal fold fixation. *Aust N Z J Surg* 2003;73:326-30.
40. Hertegård S, Hallén L, Laurent C, et al. Cross-linked hyaluronan versus collagen for injection treatment of glottal insufficiency: 2-year follow-up. *Acta Otolaryngol* 2004;124:1208-14.
41. Lee SW, Son YI, Kim CH, Lee JY, Kim SC, Koh YW. Voice outcomes of polyacrylamide hydrogel injection laryngoplasty. *Laryngoscope* 2007;117:1871-5.
42. Billante CR, Clary J, Sullivan C, Netterville JL. Voice outcome following thyroplasty in patients with longstanding vocal fold immobility. *Auris Nasus Larynx* 2002;29:341-5.
43. Shin JE, Nam SY, Yoo SJ, Kim SY. Analysis of voice and quantitative measurement of glottal gap after thyroplasty type I in the treatment of unilateral vocal paralysis. *J Voice* 2002;16:136-42.
44. Ptak M, Strack D. Voice exercise therapy versus electrostimulation therapy in patients with unilateral vocal fold paralysis. *HNO* 2005;53:1092-7.

APPENDIX

<i>Voice Case History Questions</i>	<i>Circle One Per Question</i>	
Is your voice typical today (compared with how it sounded before your operation)?	Yes	No
If "No," are you having problems today (<i>not</i> related to surgery) that are affecting your voice (e.g., allergies, upper respiratory infection, excessive talking, indigestion, reflux)?	Yes	No
How does your voice sound to you?	Normal	Abnormal
If abnormal, do you have any of these symptoms?		
Hoarse voice	Yes	No
Voice fatigues easily or effortful to talk	Yes	No
Difficulty being heard or projecting	Yes	No
Pain/hurts when using voice	Yes	No
Difficulty singing	Yes	No
Loss of upper pitch range	Yes	No