The Laryngoscope
Lippincott Williams & Wilkins, Inc., Philadelphia
© 2002 The American Laryngological,
Rhinological and Otological Society, Inc.

The Reliability of the Assessment of Endoscopic Laryngeal Findings Associated With Laryngopharyngeal Reflux Disease

Ryan C. Branski, MA; Neil Bhattacharyya, MD; Jo Shapiro, MD

Objective: **To determine the reliability of the assessment of laryngoscopic findings potentially associated with laryngopharyngeal reflux disease (LPRD).** ***Study Design:*** **Prospective randomized blinded study.** ***Methods:*** **One hundred twenty video segments of rigid fiberoptic laryngeal examinations were prospectively analyzed by five otolaryngologists blinded to patient information and were scored according to several variables potentially associated with LPRD. Separate assessments of the degree of erythema and degree of edema were scored on a five-point scale for the anterior commissure, membranous vocal fold, and interarytenoid region. Similarly, interarytenoid pachydermia, likelihood of LPRD involvement, and severity of LPRD findings were assessed. For each of these scored physical findings, inter-rater and intrarater reliabilities were determined.** ***Results:*** **The inter-rater reliabilities of the laryngoscopic findings associated with LPRD were poor. Intraclass correlation coefficients were 0.161 and 0.461 for edema of the arytenoids and membranous vocal folds, respectively ($P < .001$). Intraclass correlation coefficients were 0.181 and 0.369 for erythema of the arytenoids and membranous vocal folds, respectively ($P < .001$). Raters demonstrated poor agreement as to the severity of LPRD findings (intraclass correlation coefficient, 0.265) and the likelihood of an LPRD component for dysphonia (intraclass correlation coefficient, 0.248). Similarly, intrarater reliability was extremely variable for the various physical findings, with Kendall correlation coefficients ranging from −0.121 to 0.837.** ***Conclusions:*** **Accurate clinical assessment of laryngeal involvement with LPRD is likely to be difficult because laryngeal physical findings cannot be reliably determined from clinician to clinician. Such variability makes the precise laryngoscopic diagnosis of LPRD highly subjective.** ***Key Words:*** **Gastroesophageal reflux disease, chronic laryngitis, reliability.**
Laryngoscope, **112:1019–1024, 2002**

INTRODUCTION

The deleterious effects of gastric material on the larynx have been well documented. Early research by Delahunty and Cherry[1] found that inflammation, ulceration, and the formation of granulation tissue on the vocal fold mucosa followed extended exposure to gastric material. More recent research has suggested that patients with chronic reflux present with increased interarytenoid or posterior glottic inflammation or erythema (or a combination of these), hypertrophy of the posterior commissure (cobblestoning), and granulation tissue formation in severe cases.[2] Koufman et al.[3] recently suggested that laryngeal edema (not erythema) is the most common laryngeal finding associated with reflux. However, there are relatively few studies correlating abnormal pH monitoring with laryngeal signs of reflux, and the overall sensitivity and specificity of various findings corresponding to LPRD have not yet been determined.

Double-barrel pH monitoring is currently considered by many authorities to be the gold standard for the diagnosis and quantification of supraesophageal reflux.[4] However, pH monitoring has not realized widespread use in the everyday clinical diagnosis and management of laryngopharyngeal reflux disease (LPRD). Several practical issues limit the use of pH monitoring outside the research realm.[5] Many patients resist undergoing pH monitoring because of the inconvenience and discomfort associated with probe placement. In addition, insurance coverage for the evaluation may be problematic. Sasaki and Toohill[6] reported a general consensus among otolaryngologists that when appropriate symptoms and laryngoscopic findings are present, patients may be placed on a trial of acid suppression therapy as both a diagnostic and a therapeutic maneuver. Therefore, visual examination of the larynx and pharynx are crucial to the initial diagnosis and subsequent treatment of patients with dysphonia or other complaints consistent with LPRD.

Among studies proposing the possible and probable laryngeal manifestations of reflux, there appears to be

From the Division of Otolaryngology (R.C.B., N.B., J.S.), Brigham and Women's Hospital, and the Department of Otology and Laryngology (N.B., J.S.), Harvard Medical School, Boston, Massachusetts, U.S.A.

Editor's Note: This Manuscript was accepted for publication February 13, 2002.

Send Correspondence to Neil Bhattacharyya, MD, Division of Otolaryngology, Brigham and Women's Hospital, 333 Longwood Avenue, Boston, MA 02115, U.S.A. E-mail: neiloy@massmed.org

Copyright © The American Laryngological, Rhinological and Otological Society, Inc. Unauthorized reproduction of this article is prohibited.

vast discrepancy in the visual criteria used by clinicians to diagnose LPRD.[2] Even if accurate, consistent laryngeal findings that correlate with LPRD were identified, variability among clinicians in terms of what constitutes these abnormal findings might exist. For example, one clinician's "erythema" may be another's "pallor." The current study was designed to determine the level of general agreement or disagreement among otolaryngologists with respect to the presence of and degree of severity of laryngoscopic findings thought to be clinically consistent with LPRD. Unless consistency in the identification and quantification of visual laryngeal findings can be demonstrated, an accurate diagnosis of LPRD based primarily on laryngoscopic findings is likely to be fraught with variability from clinician to clinician and therefore from study to study.

PATIENTS AND METHODS

Patients

The study population consisted of 100 consecutive patients presenting to the voice division of an urban academic otolaryngology practice with a primary complaint of dysphonia during a 6-month period. Each of these patients (68 female and 32 male patients) underwent transoral rigid endoscopy and stroboscopy with digital video recording as a component of their complete vocal function evaluation. Patients were asked whether they experienced any classic symptoms of gastroesophageal reflux disease (GERD), such has heartburn or excessive belching, as well as symptoms of extraesophageal reflux, such as halitosis, a burning sensation in the back of the throat, or excessive throat clearing. In addition, current use of antireflux medications including both prescription and over-the-counter medicines was noted. Of the 100 patients, 38 reported at least occasional reflux-related symptoms. Of those 38 patients, 18 were currently taking antireflux medications (most commonly, omeprazole or lansoprazole). Of the remaining 62 patients who denied any reflux symptoms, 13 were currently taking prescription antireflux medication. One patients who denied having any reflux symptoms had recently undergone a Nissen fundoplication. Of the 100 patients, 47 presented with lesions of the vocal folds, 24 were diagnosed with functional voice disorders, 11 presented with a neurological source of dysphonia, 8 presented with a diagnosis of acute laryngitis, 7 had lesions of either the false vocal fold(s) or subglottis, and 3 presented with vocal fold atrophy. We chose to include patients with identifiable lesions (e.g., vocal fold granulomas) because reflux may coexist with or exacerbate these conditions.

Endoscopic views were collected using the Digital Stroboscopy System (model 9200, Kay Elemetrics, Lincoln Park, NJ). A rigid 70° endoscope (model SFT-I, Nagashima, Tokyo, Japan) was coupled to a Stryker three-chip camera (model 782, Stryker, San Jose, CA). Five-second segments of complete or near-complete visualizations of the larynx from the anterior commissure to the posterior pharyngeal wall were selected and edited using the Kay Digital Stroboscopy Editor software package (Kay Elemetrics).

Raters

Five board-certified otolaryngologists with diverse clinical interests served as raters. All of the rating otolaryngologists see patients with dysphonia as part of their routine clinical practice. The raters were full-time academic faculty members of Harvard Medical School, Boston, and their time after residency ranged from 3 to 19 years (mean period, 9.2 y). One hundred video laryngoscopy segments were placed in random order and combined with 20 video segments, which were randomly selected to be viewed twice by each rater to determine intrarater reliability. All five of the raters viewed and rated the 120 video segments separately under identical conditions without information about the diagnosis, demographic information, or treatment history of individual patients. Their responses were kept confidential.

For each video segment, the raters were asked to rate the presence of laryngeal features potentially associated with reflux, as described previously in the literature, using a five-point scale (Fig. 1). We asked raters to include visual findings consistent with hyperemia or hypervascularity as part of the erythema score. In addition, they were asked to rate the overall severity of gastroesophageal reflux (GER) findings, the likelihood that GER was a component in the patients' complaints of dysphonia, and whether or not they would prescribe antireflux medication based solely on the video laryngoscopic examination. The raters were kept blind to medical history of individual patients, including the presence of any reflux-related symptoms. The raters were aware that each patient was being evaluated for voice-related complaints.

Data Analysis

The evaluation data were imported into SPSS version 10.0 software (Chicago, IL) for subsequent statistical analysis. Standard demographic and descriptive summary data were computed for each otolaryngologist-rater. Inter-rater reliability was determined using the intraclass correlation coefficient (ICC) as described by Shrout and Fleiss[7] for a two-way random effects model. This model assesses absolute agreement among several raters for scale data. The threshold for statistical significance was set at $P < .05$. Intrarater reliability was determined for each rater based on the interpretations of the 20 repeated patient recordings. Kendall bivariate correlation coefficients were computed for each rater for each of the measured variables to assess intrarater reliability.

RESULTS

Each rater evaluated each videostroboscopic recording for a total of 600 scored patient evaluations (5×120). Means and standard deviations for each of the rated variables for each otolaryngologist across all patients are presented in Table I. Overall, 185 video studies (30.8%) were

	None/ Unlikely				Severe/ Very Likely
1. Rate the severity of edema present:					
A) at the anterior commisure	0	1	2	3	4
B) along the musculomembranous vocal folds	0	1	2	3	4
C) on the arytenoids/interarytenoid space	0	1	2	3	4
2. Rate the severity of erythema present:					
A) at the anterior commisure	0	1	2	3	4
B) along the musculomembranous vocal folds	0	1	2	3	4
C) on the arytenoids/interarytenoid space	0	1	2	3	4
3. Rate the severity of interarytenoid pachydermia	0	1	2	3	4
4. Rate the severity of GER findings	0	1	2	3	4
5. How likely is it that GER is a component of this patient's dysphonia?	0	1	2	3	4
6. Based solely on this examination, would you place this patient on anti-reflux medications of any type?		YES		NO	

Fig. 1. Laryngeal appearance rating form.

Copyright © The American Laryngological, Rhinological and Otological Society, Inc. Unauthorized reproduction of this article is prohibited.

TABLE I.
Summary of Individual Rater Data for Each Measured Descriptive Variable.

	Rater 1		Rater 2		Rater 3		Rater 4		Rater 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Edema										
Anterior	0.2	0.6	1.2	0.9	0.7	0.7	0.7	0.7	0.4	0.9
Membranous folds	1.0	0.9	1.4	1.0	1.4	0.6	1.5	1.0	0.8	1.0
Arytenoids	1.1	1.0	0.2	0.5	1.9	0.7	1.9	0.9	0.7	1.0
Erythema										
Anterior	0.2	0.6	1.3	1.0	1.1	0.6	0.9	0.7	0.3	0.9
Membranous folds	0.6	0.8	1.3	1.0	1.4	0.7	1.0	0.8	0.4	0.9
Arytenoids	1.1	1.0	0.2	0.6	1.9	0.7	1.8	0.9	0.8	1.0
Pachydermia	1.3	1.0	0.5	0.8	1.6	0.7	1.1	0.8	0.6	0.9
Severity of GER	1.2	1.0	0.5	0.8	1.8	0.7	1.7	0.8	0.7	1.0
Likelihood GER component	1.4	1.2	0.5	0.8	1.7	0.7	1.6	0.8	0.6	1.0

*All scores may range from 1–5.
SD = standard deviation; GER = gastroesophageal reflux.

interpreted as "normal," with no evidence for GER findings. The distribution of normal versus abnormal findings for the aggregate 600 evaluations according to each anatomical subsite are displayed in Table II.

Inter-rater Reliability

The single-measure and average-measure intraclass correlation coefficients for measured variables are displayed in Table III, along with respective statistical significances. Both the single-measure and average-measure intraclass correlation coefficients were found to be statistically significant, but the overall magnitude of the coefficient (and therefore measured reliability) was poor. The single-measure intraclass correlation coefficient corresponded to the inter-rater reliability from otolaryngologist to otolaryngologist for any given rated variable. The average-measure intraclass correlation indicates the precision that could be obtained by using multiple raters, as in the present study, and then determining an average score for each variable. The coefficients obtained indicate that, when averaging the rating among several otolaryngologists, reliability increases significantly. With respect to both edema and erythema of the glottis, raters were most likely to agree on the degree of involvement of the musculomembranous vocal folds and least likely to agree on the degree of arytenoid abnormality. Relatively poor agreement was noted for the degree of pachydermia, the overall severity of GER findings, and the likelihood that GER was contributing to the dysphonia. According to established criteria, all coefficients for single-measure reliability were poor, except for edema of the musculomembranous folds, which exhibited only fair reliability.[8]

Intrarater Reliability

The intrarater reliabilities for each variable as measured by the Kendall bivariate correlation coefficient are presented in Table IV. Only one of the raters (Rater 2) was able to demonstrate consistent intrarater reliability for multiple scale variables, with fair to excellent reliability coefficients. However, even when the coefficients of reliability reached statistical significance, the overall magnitudes of the correlations for the various raters was weak. Again, edema of the musculomembranous vocal folds exhibited the greatest intrarater reliability across each of the raters. Comparisons were conducted for each rater to determine whether the rater was consistent in recommending GER treatment. Table V presents mean GER severity ratings for each rater based on patient groups for which the rater would recommend and would not recommend GER treatment. Each rater was able to consistently segregate patients into treatment groups based on severity of GER findings.

DISCUSSION

In the late 1960s, Delahunty and Cherry[1] described patterns of injury to the larynx caused by gastric material.

TABLE II.
Distribution of Normal and Abnormal Scores for Anatomic Subsites (n = 600).

	Edema			Erythema			
Score	AC	MMF	ARY	AC	MMF	ARY	Pachydermia
0 (Normal)	255	142	219	231	235	208	168
1–4 (Abnormal)	210	445	361	229	353	367	342
Unscored	135	13	20	140	12	25	90

AC = anterior commissure; MMF = musculo-membranous folds; ARY = arytenoids.

Copyright © The American Laryngological, Rhinological and Otological Society, Inc. Unauthorized reproduction of this article is prohibited.

TABLE III.
Intraclass Correlation Coefficients (ICC) for Reliability of Examination Findings Between Raters.

	Single measure		Average measure	
	r	*P*	r	*P*
Edema				
Anterior	0.363	<.001	0.740	<.001
Membranous folds	0.461	<.001	0.810	<.001
Arytenoids	0.161	<.001	0.490	<.001
Erythema				
Anterior	0.293	<.001	0.675	<.001
Membranous folds	0.369	<.001	0.745	<.001
Arytenoids	0.181	<.001	0.525	<.001
Pachydermia	0.351	<.001	0.730	<.001
Severity of GERD	0.265	<.001	0.644	<.001
Likelihood GERD component	0.248	<.001	0.623	<.001

*Interpretation of ICC: <0.40 = poor; 0.40–0.59 = fair; 0.60–0.74 = good; >0.74 = excellent.
GERD = gastroesophageal reflux disease.

Since that time, awareness of the potential impact of LPRD on the larynx has steadily increased. Chronic hoarseness, globus pharyngeus, and muscle tension dysphonia have long been purportedly associated with LPRD.[9] Recently, evidence for a relationship between LPRD and mass lesions of the larynx including glottic carcinoma has emerged.[10,11] Given the prevalence of laryngeal complaints among patients seeking otolaryngologic care, LPRD is often entertained as the potential underlying diagnosis for many of these symptoms. Although double-lumen pH monitoring is considered the gold standard for the diagnosis of both esophageal and laryngeal reflux disease, cost factors, patient inconvenience and, to some degree, its unavailability have prevented it from becoming widely used as a diagnostic modality in chronic laryngitis. More often, otolaryngologists base the diagnosis of LPRD on the patient's dysphonic manifestations and associated reflux symptomatology, as well as the office-based laryngeal examination. However, because classic symptoms of GERD are absent in 57% to 80% of patients with significant clinical manifestations of LPRD, otolaryngologists often rely heavily on their endoscopic assessment of the larynx in determining the presence of LPRD.[2]

Reflux-related injury to the larynx has also been termed acid laryngitis, acid posterior laryngitis, peptic

TABLE IV.
Intrarater Reliabilities for Laryngoscopic Physical Findings Associated With GER.

	Rater 1		Rater 2		Rater 3		Rater 4		Rater 5	
	Coefficient	Significance	Coefficient	Significance	Coefficient	Significance	Coefficient	Significance	Coefficient	Significance
Edema										
Anterior	0.594	***0.010***	0.702	***0.001***	0.000	1.000	0.445	0.054	0.509	0.051
Membranous folds	0.669	***0.001***	0.507	***0.016***	0.336	0.139	0.575	***0.005***	0.617	***0.002***
Arytenoids	0.413	***0.037***	0.813	***0.000***	0.350	0.112	0.661	***0.001***	0.594	***0.006***
Erythema										
Anterior	−0.121	0.608	0.834	***0.000***	0.047	0.870	0.289	0.196	0.681	***0.011***
Membranous folds	0.584	***0.005***	0.778	***0.000***	0.256	0.247	0.451	***0.029***	0.750	***0.001***
Arytenoids	0.304	0.125	0.837	***0.000***	0.331	0.193	0.586	***0.004***	0.553	***0.008***
Pachydermia	0.401	***0.049***	0.678	***0.002***	0.494	***0.038***	0.540	***0.012***	0.265	0.241
Severity of GERD	0.342	0.085	0.778	***0.000***	0.589	***0.011***	0.504	***0.015***	0.979	***0.000***
Likelihood GERD component	0.275	0.156	0.778	***0.000***	0.364	0.112	0.552	***0.008***	0.846	***0.000***
Recommend GER treatment	0.492	***0.032***	0.793	***0.001***	0.329	0.175	0.390	0.089	0.192	0.402

**P* values <.05 are marked as bold italic.
†Coefficients represent Kendall's bivariate correlation coefficient.
GER = gastroesophageal reflux; GERD = gastroesophageal reflux disease.

Copyright © The American Laryngological, Rhinological and Otological Society, Inc. Unauthorized reproduction of this article is prohibited.

TABLE V.
Mean Severity Scores for GER Findings According to Intent to Treat.

	Recommend Treatment Group	No GER Treatment Group	*P* value
Rater 1	1.83	0.40	<.001
Rater 2	1.71	0.16	<.001
Rater 3	2.19	1.12	<.001
Rater 4	2.07	0.94	<.001
Rater 5	2.00	0.22	<.001

GER = gastroesophageal reflux.

laryngitis, reflux laryngitis, and LPRD.[2] Various laryngeal findings have been associated with the diagnosis of LPRD. These include erythema or edema of the posterior one-third of the glottis, hyperemia of the posterior larynx, cobblestoning, and "heaping up" or thickening of the interarytenoid mucosa (pachydermia laryngis).[12] The wide array of nonspecific laryngeal findings and vague diagnostic terminology all suggest that the true pathophysiology of LPRD is poorly understood. If otolaryngologists are to be successful in making a diagnosis of LPRD based on clinical assessment without pH probe data, two criteria must be satisfied. First, sensitivities and specificities should be determined for various laryngoscopic findings in the diagnosis of LPRD. Second, otolaryngologists must be able to demonstrate reliability in identifying these physical findings to make accurate diagnoses. If the laryngeal findings cannot be reliably determined among different otolaryngologists, even with knowledge of the sensitivities and specificities for the various physical findings, accurate clinical diagnosis of LPRD will be difficult. Therefore, we sought to determine the reliability among otolaryngologists for the identification and quantification of various laryngeal physical findings potentially associated with LPRD.

Our data indicate that otolaryngologists vary significantly in their ratings of the various laryngoscopic physical findings that could be associated with LPRD. We found relatively poor inter-rater reliability for all of the visually assessed variables. This indicates that, even if the most sensitive or specific physical examination findings among these variables were known for the diagnosis of LPRD, different otolaryngologists might be unable to accurately diagnose LPRD based solely on such findings. We were not entirely surprised that such variability among the otolaryngologist-raters was encountered because all of these clinical variables are subjective in interpretation. Otolaryngologists are not alone in their difficulties with rating mucosal disease potentially attributable to GER. Studies in the gastroenterology literature have documented poor correlation coefficients ranging from 0.15 to 0.40 for the assessment and grading of reflux esophagitis among different endoscopists.[13] Furthermore, the poor reliability of the ratings for the "severity of GER" and "likelihood GER component" variables suggests additional inconsistency. Because these correlation coefficients were lower than those of the physical finding variables, otolaryngologists are also likely to disagree on the degree of impact that GER has on patients' symptoms given a set of physical findings.

We were somewhat surprised that the otolaryngologists were unable to demonstrate good intrarater reliability (Table IV). It is apparent from the statistical analysis that individual otolaryngologists may have difficulty in consistently identifying and rating physical findings in their endoscopic assessments of the larynx. However, posthoc analysis of intrarater consistency found that the otolaryngologist-raters were able to separate patients into what they thought were treatment-appropriate and treatment-inappropriate groups (Table V), which were consistent with their mean severity scores for LPRD findings. Furthermore, Table V does provide affirmative data to suggest that there can be consistency in the evaluation of LPRD. The intent-to-treat data indicate that, overall, the otolaryngologists each individually used a consistent method or methods to determine whether or not to recommend treatment for presumed LPRD. Therefore, it is possible that despite inter-rater and intrarater variability for the *individual* laryngeal findings associated with LPRD, some overall consistency may be expected in the diagnosis based on physical findings. This important fact implies that other factors in the examination may be important in suggesting LPRD. These factors may include other features such as subcordal edema, hypervascularity, or other features not yet determined.

The average-measure intraclass correlation coefficients were reasonably high. This indicates that when multiple raters are used to evaluate these physical examination variables, reasonable reliability can be expected from the average of their ratings. This indicates that in future studies attempting to correlate physical findings with pH monitoring results or other tests for LPRD, multiple raters should be used for each patient examination because the reliability of the average measure among those raters is acceptable. However, the lowest values for the average-measure intraclass correlation coefficients were noted for both edema and erythema of the arytenoids. The posterior larynx, which has traditionally been thought to be the primary site affected by LPRD, is the most difficult area to rate and assess consistently. It is likely that this stems from the fact that the musculomembranous folds are sharply demarcated anatomic entities and are generally white and lustrous. Therefore, edema and erythema, when present, are relatively easily discerned. The same does not hold true for the arytenoid region, which has an exceedingly variable natural architecture from patient to patient.

Notably, this study used an idealized form of laryngeal assessment for comparisons. In clinical practice, otolaryngologists may use indirect laryngoscopy or flexible fiberoptic laryngoscopy to visualize the larynx. Also, patients may be examined at different times of the day or after different dietary challenges the night before their evaluation. Thus, additional variability may be introduced into the examination process, which could lead to further unreliability in assessment of the laryngeal findings in LPRD. It is possible that otolaryngologists who "subspecialize" in the management of voice disorders would exhibit better inter-rater and intrarater reliabilities for

Copyright © The American Laryngological, Rhinological and Otological Society, Inc. Unauthorized reproduction of this article is prohibited.

these clinical variables. However, because the vast majority of patients with dysphonia are not initially treated by subspecialty laryngologists, establishing reliable laryngeal features among a wide array of otolaryngologists is important. Several additional laryngeal findings that could be attributed to LPRD were not assessed in the current study. These include sulcus vocalis and subcordal edema. Both of these findings can be difficult to assess and often require deeper placement of the scope. In our video segments the three-dimensional view is limited, so we did not think it appropriate to assess for these findings. Hypervascularity of the interarytenoid region has also been associated with LPRD, and we asked our raters to include hyperemia in consideration of "erythema," to include this potential finding.

One potential solution to the problem of the unreliability of the laryngeal examination findings is to "train" otolaryngologists with a set of standardized examinations, to foster consistency among the otolaryngologists for scoring these variables. This would be similar to creating "trained listeners" who evaluate patient voices for dysphonia rating scales and so forth. Once reliability can be established for various laryngeal findings, further work can be conducted to determine which findings are sensitive and specific for LPRD. This would allow otolaryngologists to make the best evidence-based treatment decisions for patients with dysphonia potentially attributable to LPRD. Awareness of the laryngeal manifestations of GER continues to grow, and the accurate diagnosis of LPRD may allow otolaryngologists to identify and treat the significant fraction of patients with LPRD-related dysphonia.

BIBLIOGRAPHY

1. Delahunty JE, Cherry J. Experimentally produced vocal cord granulomas. *Laryngoscope* 1968;78:1941–1947.
2. Ulualp SO, Toohill RJ. Laryngopharyngeal reflux: state of the art diagnosis and treatment. *Otolaryngol Clin North Am* 2000;33:785–802.
3. Koufman JA, Amin MR, Panetti M. Prevalence of reflux in 113 consecutive patients with laryngeal and voice disorders. *Otolaryngol Head Neck Surg* 2000;123:385–388.
4. Little JP, Matthews BL, Glock MS, et al. Extraesophageal pediatric reflux: 24-hour double-probe pH monitoring of 222 children. *Ann Otol Rhinol Laryngol Suppl* 1997;169:1–16.
5. Hanson DG, Conley D, Jiang J, Kahrilas P. Role of esophageal pH recording in management of chronic laryngitis: an overview. *Ann Otol Rhinol Laryngol Suppl* 2000;184:4–9.
6. Sasaki CT, Toohill RJ. Introduction: 24-hour ambulatory pH monitoring for patients with suspected extraesophageal complications of gastroesophageal reflux: indications and interpretations. *Ann Otol Rhinol Laryngol Suppl* 2000;184:2–3.
7. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull* 1979;86:420–428.
8. Cicchetti DV, Sparrow SS. Developing criteria for establishing interrater reliability of specific items in a given inventory. *Am J Ment Deficiency* 1981;86:127–137.
9. Koufman JA. The otolaryngologic manifestations of gastroesophageal reflux disease. *Laryngoscope* 1991;101(Suppl 53):1–78.
10. Harrill WC, Stasney CR, Donovan DT. Laryngopharyngeal reflux: a possible risk factor in laryngeal and hypopharyngeal carcinoma. *Otolaryngol Head Neck Surg* 1999;120: 598–601.
11. Koufman JA, Burke AJ. The etiology and pathogenesis of laryngeal carcinoma. *Otolaryngol Clin North Am* 1997;30: 1–19.
12. Gaynor EB. Laryngeal complications of GERD. *J Clin Gastroenterol* 2000;30(Suppl):31–34.
13. Lundell LR, Dent J, Bennett JR, et al. Endoscopic assessment of oesophagitis: clinical and functional correlates and further validation of the Los Angeles classification. *Gut* 1999; 45:172–180.

Copyright © The American Laryngological, Rhinological and Otological Society, Inc. Unauthorized reproduction of this article is prohibited.