Published in final edited form as:

Laryngoscope. 2014 October ; 124(10): 2345–2351. doi:10.1002/lary.24737.

Correlation between reflux and multichannel intraluminal impedance pH monitoring in untreated volunteers

Marie E. Jetté, M.S.,

Department of Surgery, Department of Communication Sciences and Disorders University of Wisconsin-Madison, Madison, WI, USA

Eric A. Gaumnitz, M.D.,

Department of Medicine University of Wisconsin-Madison, Madison, WI, USA

Martin A. Birchall, M.D.,

University College London Ear Institute The Royal National Throat, Nose and Ear Hospital, London, U.K.

Nathan V. Welham, Ph.D., and

Department of Surgery University of Wisconsin-Madison, Madison, WI, USA

Susan L. Thibeault, Ph.D.

Department of Surgery University of Wisconsin-Madison, Madison, WI, USA

5107 WIMR, 1111 Highland Avenue, Madison, WI 53705 Division of Otolaryngology-Head and Neck Surgery, Department of Surgery School of Medicine and Public Health University of Wisconsin-Madison

Abstract

Objective—While probable causative agents have been identified (e.g., refluxate components, tobacco smoke), the definitive mechanism for inflammation-related laryngeal mucosal damage remains elusive. Multichannel intraluminal impedance combined with pH monitoring (MII/pH) has emerged as a sensitive tool for diagnosis and characterization of gastroesophageal reflux disease (GERD) with laryngopharyngeal manifestations. To determine the relationship between laryngeal signs and MII/pH, we examined correlations between Reflux Finding Score (RFS) ratings of videostroboscopic laryngeal examinations and findings from MII/pH.

Study Design—Correlational study.

Methods—Healthy, untreated volunteers (n = 142) underwent reflux diagnosis using data acquired from MII/pH testing. Eight trained clinicians performed RFS ratings of corresponding laryngeal examinations. Averaged RFS ratings were compared to MII/pH data using Pearson correlation coefficients. The relationship between RFS and MII/pH findings and demographic/

Correspondence, reprint requests, and proofs should be sent to: Susan Thibeault 5107 Wisconsin Institutes for Medical Research 1111 Highland Avenue, Madison, WI 53705 Tel: (608) 263-0121 thibeaul@surgery.wisc.edu.

Conflict of Interest: None declared

Presented at the 2014 Combined Otolaryngological Spring Meetings, American Laryngological Society, Las Vegas, NV, May 15, 2014

clinical information (age, sex, smoking status, reflux) was assessed using general linear modeling. Rater reliability was evaluated.

Results—Posterior commissure hypertrophy was negatively correlated with minutes of nonacid refluxate ($R=-0.21$, $p=0.0115$). General linear modeling revealed that 28-40% of the variance in ratings of ventricular obliteration, erythema/hyperemia, vocal fold edema, diffuse laryngeal edema, posterior commissure hypertrophy, and granulation/granuloma could be explained by main and interaction effects of age, sex, smoking status, and reflux. Intra- and inter-rater reliability for RFS were poor-fair.

Conclusion—These results support the theory that the RFS is not specific for reflux in healthy, untreated volunteers, suggesting there may be alternate explanations for inflammatory clinical signs commonly ascribed to reflux in this population.

Keywords

impedance monitoring; pH monitoring; gastroesophageal reflux; laryngopharyngeal reflux; LPR diagnosis; Reflux Finding Score

INTRODUCTION

Chronic laryngitis, one of the most commonly diagnosed dysphonias among health care professionals¹, is characterized by a variety of inflammatory changes observed in patients with an array of symptoms. Gastroesophageal reflux disease (GERD) has been implicated as a probable etiologic factor for chronic laryngitis²⁻⁴, though treatment with proton pump inhibitors (PPI), the current standard of care for GERD, demonstrates a nonsignificant benefit over placebo⁵. In spite of lack of efficacy data supporting the use of PPI, 46.2% of patients with a diagnosis of chronic laryngitis receive medication⁶. While reflux with laryngeal manifestations (laryngopharyngeal reflux; LPR) may indeed be an activator of laryngeal inflammation, the extent to which the effects of LPR alone contribute to the clinical picture of chronic laryngitis is unknown.

The Reflux Finding Score (RFS) was developed by Belafsky et al⁷ to document physical LPR findings on a standardized scale, with scores ranging from 0 (no evidence of reflux) to 26 (severe evidence of reflux). To validate this scale, RFS scores from 40 patients with clinically diagnosed LPR documented by esophageal-pharyngeal pH monitoring were compared to scores from 40 age-matched, asymptomatic controls who had not undergone confirmatory pH monitoring and a statistically significant difference in scores was found⁷. Based on these results, the authors concluded with 95% certainty that a person with RFS greater than 7 has LPR. Other researchers have determined that findings and symptoms ascribed to LPR are not specific to LPR⁸. Milstein et al.⁹ found at least one sign of laryngeal tissue irritation in the majority of volunteers undergoing laryngoscopy with no history of ear-nose-throat complaints or diagnosis of reflux. Similarly, Hicks et al.¹⁰ demonstrated that 86% of normal, healthy, adult volunteers had findings commonly associated with reflux. Moreover, studies examining reliability of subjective laryngoscopic ratings of LPR have revealed mixed results ranging from poor to good¹¹.

Ambulatory pH monitoring has been lauded as gold standard for diagnosis of acid reflux, however its role in diagnosing LPR remains controversial. In a review of multiple studies, Vaezi et al.¹² revealed that only 54% of patients with laryngoscopic signs of reflux have abnormal esophageal acid exposure on pH probe. They suggest that such low accuracy demonstrates either overdiagnosis of reflux as the cause of laryngeal pathology or lack of sensitivity of pH monitoring in documenting LPR¹². Other diagnostic tools developed more recently include multichannel intraluminal impedance (MII), pharyngeal pH monitoring¹³, and hypopharyngeal MII (HMII)¹⁴. Impedance (including MII and HMII) monitoring measures both acid and nonacid reflux in liquid and gaseous forms by measuring electrical resistance between different points along the esophagus. Combined with pH monitoring, impedance may offer improved detection of reflux events associated with LPR, though its role in LPR diagnosis has not been established.

The primary goal of this study was to examine correlations between endoscopic findings using RFS and measures acquired from MII with pH (MII/pH) monitoring in healthy, untreated volunteers. Given that the pathophysiology of laryngeal inflammation has not yet been defined and concerns published in the literature regarding the specificity of the RFS, we hypothesized that there would be poor correspondence between these sets of variables.

MATERIALS AND METHODS

Participant selection

Participants aged 21-65 years were recruited with newspaper and email advertisements and signs in the clinic and around the University of Wisconsin-Madison. Participants underwent videolaryngostroboscopic examination and 24-hour MII/pH, with each procedure performed on separate dates. The protocol was approved by the Institutional Review Board of University of Wisconsin-Madison and informed consent was obtained from all participants.

Participants were excluded from the study if they had a history of radiation therapy to the head and neck within the past five years, lung or gastroesophageal surgery, chronic sinusitis or rhinitis in the last year, an acute traumatic event near the larynx in the last year, tracheostomy or other significant laryngeal or tracheal surgery, and substance or alcohol abuse in the past year. Consumption of more than 10 (women) and 17 (men) units of alcohol per week (means of United Kingdom and United States recommended weekly limits) excluded participants¹⁵. Further exclusion criteria included malignancy (except superficial basal cell carcinoma) within the past five years, presence of an infectious cause of laryngitis in the past three months, need for continuous therapy with diazepam, phenytoin, mephenytoin, warfarin, anticholinergics, antineoplastics, prostaglandin analogs, H₂-receptor antagonists, steroids (inhaled, oral or intravenous), promotility drugs and sucralfate, use of any PPI or H₂ blockers in the past year, theophylline or any other investigational compound or participation in an investigational drug study in the previous 60 days. Women were excluded if pregnant or lactating. Nonsmokers had not smoked during the previous year. Smokers were defined by consumption of a minimum of 5 cigarettes/5g of tobacco per day for the duration of one or more years, thereby distinguishing them from light smokers^{16,17}.

Laryngoscopy

Participants underwent videolaryngostroboscopic examination using rigid or flexible endoscope (Pentax Medical, Lincoln Park, N.J.). Topical anesthetic was avoided unless the participant exhibited extreme gag reflex and was unable to tolerate examination. The larynx was visualized during sustained phonation on /i/ and quiet breathing. Digital recordings of laryngoscopic examinations were edited, randomized by clip number (List Randomizer, random.org), and organized into two video montages (iMovie, Apple) representing two randomizations. Sixteen video clips were chosen randomly (List Randomizer, random.org) and included at the end of each video montage to assess intra-rater reliability.

Reflux Finding Score

Eight raters provided ratings for this analysis using an adapted RFS (Table 1). Raters included clinicians with 55 combined years of experience in voice disorders. A 45-minute training presentation was developed demonstrating published photographic examples of each RFS item^{7,18,19} as well as their descriptions. Following training, raters were presented with still images from 5 examinations and performed group consensus ratings. Notes from the presentation and consensus ratings were saved and raters were able to access these while completing RFS. Six of 8 raters completed the training session with consensus. Two raters that did not attend reviewed the presentation and consensus notes before completing ratings. No demographic or pH/impedance data were provided to raters. Raters were also blinded to the purpose of investigation and participant classification.

Combined multichannel intraluminal impedance and 24 hour pH probe

After a four-hour fast, participants underwent conventional esophageal manometry (circumferential probe, Medtronic, Shoreview, MN) to locate lower and upper esophageal sphincters (LES and UES, respectively). The MII/pH catheter had two antimony electrodes placed such that proximal sensor was positioned 1 cm below and distal sensor 15 cm below UES. Impedance was measured through 7 sensors placed along a 2.3 mm polyurethane catheter. This catheter was placed transnasally immediately following manometry. Configuration of the catheter allowed recording of changes in intraluminal impedance at 3, 5, 7, 9, 15, and 17 cm above LES. Data from impedance channels and pH electrodes were transmitted at 50 Hz and stored together on a portable data recorder (Sleuth; Sandhill Scientific Inc., Highlands Ranch, Colorado) for later synchronization. Participants were monitored for 18-24 hours and encouraged to eat regular meals and participate in routine activities. Change in position (upright and supine) and symptomatic events including heartburn or regurgitation were documented by using buttons on the data recorder. Data were uploaded and analyzed using commercially available software (Bioview Analysis; Sandhill Scientific Inc., Highlands Ranch, Colorado).

Analysis of pH data—Acid reflux episodes were defined as drops in pH to less than 4 for at least 5 seconds. Total acid exposure time (%) was calculated as total time of acid reflux episodes divided by monitoring time. Johnson/DeMeester score²⁰ was obtained using 6 parameters: (1) total percentage time pH <4.0; (2) percentage time pH <4.0 in upright position; (3) percentage time pH <4.0 in recumbent position; (4) total number acid reflux

episodes; (5) total number acid reflux episodes longer than 5 minutes; and (6) duration of longest acid reflux episode.

Analysis of MII data—Recorded meal periods were excluded from analysis. On impedance, gas reflux was defined as rapid ($>3000\Omega/s$) retrograde moving increase in impedance in at least two impedance sites. Liquid reflux was defined as retrograde moving 40% fall in impedance in two distal impedance sites. Proximal reflux was considered when refluxate reached the 15 cm impedance sensor. Total bolus exposure time (%) was defined as the combination of durations of gas and liquid reflux events divided by total time monitored.

Interpretation of combined dual channel MII/pH Data—Participants were assigned to cohorts – GERD, LPR, normal – based on MII/pH data. GERD was defined by acid exposure percent time of the distal pH probe >4.0 , DeMeester score >14.7 , and/or bolus exposure percent time of more than 1.4%²¹. LPR was defined by >31 proximal reflux events^{22,23}. Normal was defined by the following criteria: acid exposure percent time of the distal pH probe <4.0 ; DeMeester score <14.7 ; and <31 proximal reflux events²².

Statistical analysis

To determine inter-rater reliability, intraclass correlation coefficients (ICC) were calculated. Pearson correlation coefficients were used to evaluate intra-rater reliability. Average within rater agreement across all 8 raters was computed for each RFS item. RFS ratings for each videostroboscopic examination were averaged across all ratings from 8 individual raters. Pearson correlation coefficients were used to determine correlations between average RFS ratings and findings on MII/pH and correlations between age and average RFS ratings. General linear models, including repeated measures analysis of variance (ANOVA) and analysis of covariance (ANCOVA), were fitted to assess main effects of age, cohort, sex, and smoking status, as well as the two-, three-, and four-way interaction effects of age*sex, age *cohort, age*smoking status, cohort*sex, cohort*smoking status, sex*smoking status, age*sex*smoking status, age*cohort*smoking status, age*sex*cohort, cohort*sex*smoking status, and age*cohort*sex*smoker for all RFS ratings. T-tests were used to examine differences in variables that could not be accounted for by linear modeling. All analyses were performed with SAS software (SAS Institute Inc., Cary, N.C.) with type I error set at 0.05.

RESULTS

Clinical and demographic characteristics

Of 155 original video clips included in the montages provided to raters, 13 were excluded from rating and analysis due to insufficient views from anterior commissure to posterior pharyngeal wall. Data from 142 participants including videolaryngostroboscopic recordings, MII/pH variables (Table 2), and averaged RFS ratings (Table 2) were therefore included in final analysis. Analysis of MII/pH data revealed 38 participants with GERD (27%), 44 with LPR (31%), and 60 normal (42%). Of 142 participants, 116 (82%) had total RFS >7 and 55 (39%) had total RFS >11 . Age, sex, smoking, reflux cohort, and total RFS characteristics of

these participants are summarized in Table 3. Videostroboscopic examination and MII/pH testing were completed with an average of 61 days between each procedure.

RFS rater reliability and agreement

ICC for intra-rater reliability ranged from 0.05 to 0.45 (Table 4). Results demonstrate poor to fair reliability for all RFS rating items. Inter-rater reliability was assessed on 256 observations from 8 raters. ICC ranged from 0.21 to 0.48 (Table 4), indicating poor to fair inter-rater reliability for all RFS rating items. Average intra-rater agreement examines overall levels of rater self-consistency, for each rater and RFS rating. Results are based on repeated ratings of 16 video clips and indicate that individual raters were 54.8-71.7% reliable across all ratings and that they produced the same value for any individual variable 48.75-78.75% of the time (Table 5).

Correlations between RFS and MII/pH

Average RFS ratings for each videostroboscopic examination were compared to individual MII/pH variables resulting in 144 analyzed correlations across 142 participants. There was a single significant correlation between posterior commissure hypertrophy and minutes of nonacid refluxate ($R=-0.21$, $p=0.0115$). No other correlations were significant (data not shown).

Effect of clinical and demographic characteristics on RFS

Average RFS ratings for each variable were analyzed relative to clinical and demographic data including cohort, sex, and smoking status. Age was analyzed as a main effect and also included in a separate interaction effects model (Table 6). Interaction effects of cohort, sex, smoking status, and age influenced averaged RFS ratings. General linear modeling including all variables and their interactions (Table 6, Model 2) explained 25-40% of the variance observed in many RFS ratings. While both models tested could not account for variance in ratings of subglottic edema and thick endolaryngeal mucus, further analysis revealed main effects of sex on both of these variables ($p=0.025$, $p=0.049$, respectively).

DISCUSSION

The major finding of this study was a single statistically significant correlation between RFS and MII/pH variables in a group of healthy, non-treatment-seeking, untreated volunteers. We found a negative correlation between posterior commissure hypertrophy and duration (minutes) of nonacid reflux ($R=-0.21$, $p=0.0115$), suggesting that posterior commissure hypertrophy is decreased with greater duration of nonacid reflux. This result is supported by biological evidence demonstrating less proinflammatory cytokine gene expression with greater acid exposure in biopsies taken from the posterior commissure²⁴. Though this correlation coefficient is statistically significant, it is meaningless unless properly interpreted for clinical relevance. Calculating coefficient of determination (R^2) yields 0.044, meaning that 4.4% of variation in ratings of posterior commissure hypertrophy can be explained or accounted for by variation in duration of nonacid reflux. This interpretation of the data suggests there are other factors (e.g., demographic characteristics) aside from reflux findings

measured by MII/pH that may explain variability in RFS ratings. It is also possible that there is inherent lack of RFS validity for specific reflux diagnosis.

The primary outcome measures of our study were 8 RFS ratings in addition to total RFS averaged across 8 trained clinician raters and 16 MII/pH variables. Though averaged RFS ratings were used for analysis, it is worth noting that inter- and intra-rater reliability for RFS was poor-fair. In a review of the literature examining reliability for laryngopharyngeal findings in LPR, Powell and Cocks¹¹ presented a summary from 9 publications demonstrating variable reliability ranging from poor-good. They suggested variability might be related to methods of assessment or statistical tests used. Potential explanations for poor intra-rater reliability observed in our study relate to the inherent limits of human raters' visual-perceptual systems and the RFS scale itself. Rosen²⁵ has suggested several limitations and possible errors associated with visual-perceptual ratings of videostroboscopy, including rater fatigue and lack of variability of videos. Additionally, whereas some variables (e.g., subglottic edema) can be scored as 0 (absent) or 2 (present), other variables (e.g., vocal fold edema) are scored on a 5 point scale (0, 1, 2, 3, 4). When data are pooled for statistical calculation of intra-rater reliability, the difference between ratings of 0 and 2 is given greater weight than the difference between ratings on a five-point scale. Examining agreement in conjunction with reliability gives an indication of statistical penalties resulting from limits of the scale. For example, upon repeat rating of thick endolaryngeal mucus, clinicians on average agreed with their initial rating 72.4% of the time, whereas intra-rater reliability was calculated at $R=0.12$ ($p=0.0001$) indicating poor reliability. Agreement implies that two raters assign identical meanings to each score for each variable, whereas reliability indicates that raters rate variables in parallel fashion, without implying that score values have the same meaning. If the range of scores is restricted (e.g., raters consistently avoid extremes of a scale or scores vary little with respect to variable rated), reliability coefficients may be low, even if raters agree. In this study, it is possible raters avoided severe extremes of the RFS given they were rating images from non-treatment-seeking volunteers as opposed to a pathologic population.

To bolster the clinical relevance of our findings, we used combined MII/pH variables semi-diagnostically to categorize our study population into cohorts including LPR, GERD, and normal based on normative data²¹⁻²³. Our study is the first to report on the incidence of GERD and LPR based on MII/pH in untreated, non-treatment-seeking healthy volunteers. Within our participant group, more than half (58%) was categorized as either LPR or GERD, whereas 42% demonstrated normal findings on MII/pH. Similarly, categorization of participants using published thresholds for total RFS of 7⁷ and 11²⁶ yielded 82% and 39% (respectively) categorized as LPR, supporting Hicks et al.'s finding that 86% of normal, healthy, adult volunteers had signs associated with reflux¹⁰. In a study investigating the diagnostic usefulness of MII/pH in 98 patients with suspected LPR off PPI therapy for at least 2 weeks, Lee et al. found that 54% demonstrated pathologic GERD²⁷, a finding consistent with our data in spite of the difference in study populations. It should be noted that in our study LPR was determined based on impedance and pH findings in the proximal esophagus, not in the hypopharynx, which may have resulted in overestimation of incidence of LPR. Supporting this possibility, an investigation of 34 asymptomatic, untreated research participants using hypopharyngeal MII/pH revealed a single LPR event recorded from one

participant (3%), whereas in symptomatic, untreated patients, 24/184 (13%) had at least one LPR event documented¹⁴. In clinical practice, gastroenterologists use MII/pH to diagnose reflux in patients with persistent symptoms despite acid-suppressive therapy. Diagnosis includes examining symptom association²⁸ (i.e., determining whether episodes recorded by MII/pH are associated with a corresponding symptom) and comparing MII/pH variables in patients on therapy to normative values²⁹. As we were attempting to use MII/pH as the sole objective measure of reflux in a non-treatment-seeking population, symptom association and treatment response were not evaluated within the present research design.

The clinical/demographic interaction and main effects observed within our dataset provide insight into factors that explain some variance in RFS ratings. General linear modeling including main and interaction effects of age, reflux cohort, sex, and smoking status could explain 25-40% of the variance observed in all RFS variables except subglottic edema and thick endolaryngeal mucus, suggesting that RFS ratings are influenced by clinical and demographic factors to a greater extent than MII/pH measures. Inflammatory signs measured with RFS are in part related to the combinations of sex, smoking status, and age of the larynx being rated, as opposed to reflux alone. Subglottic edema, also referred to as pseudosulcus and infraglottic edema, has long been thought to be predictive of,³⁰ and specific for,¹⁹ LPR; however, our results demonstrate that males receive greater ratings than females on this variable regardless of reflux cohort, smoking status, and age. It seems possible that this finding so commonly ascribed to inflammation from reflux may be a result of anatomic differences between males and females. Males also received greater ratings than females for thick endolaryngeal mucus, suggesting that this finding provides more information about the sex of the person being examined than it does about reflux.

While attempts were made to eliminate bias, we recognize limitations in our study design that may have prejudiced our results. Of primary consideration is that we examined data from non-treatment-seeking volunteers, a population not representative of a typical clinical population. It would be ideal to repeat the study in treatment-seeking patients for whom laryngeal inflammation impacts vocal function, thereby addressing the role of reflux specific to diagnosis of chronic laryngitis. We also recognize that we persisted in analyzing averaged RFS ratings in spite of poor reliability, though we attempted to avoid this issue by providing raters with training. Finally, we acknowledge that reflux status may have changed in the time between videostroboscopic examination and MII/pH testing. This could be avoided in future studies by completing videostroboscopic examination immediately prior to MII/pH.

CONCLUSIONS

Our data demonstrate an overall lack of correlation between RFS and MII/pH, supporting the hypothesis that RFS is not specific for reflux in non-treatment-seeking, untreated volunteers. Our findings also illustrate that in spite of training, raters demonstrated poor-fair inter- and intra-rater reliability on RFS, consistent with results from other studies. Finally, we suggest that clinical and demographic characteristics, including sex, smoking status, and age, contribute to differences in RFS ratings.

Acknowledgments

This work was supported by grants T32 DC-009401 and R01 DC-009600 from the NIH/NIDCD. The authors would like to thank Dr. Glen Levenson for statistical support.

References

1. Cohen SM, Kim J, Roy N, Asche C, Courey M. Prevalence and causes of dysphonia in a large treatment-seeking population. *Laryngoscope*. 2012; 122:343–348. [PubMed: 22271658]
2. Koufman JA. The otolaryngologic manifestations of gastroesophageal reflux disease (GERD): a clinical investigation of 225 patients using ambulatory 24-hour pH monitoring and an experimental investigation of the role of acid and pepsin in the development of laryngeal injury. *Laryngoscope*. 1991; 101:1–78. [PubMed: 1895864]
3. Koufman J, Sataloff RT, Toohill R. Laryngopharyngeal reflux: consensus conference report. *J Voice*. 1996; 10:215–216. [PubMed: 8865091]
4. Hanson DG, Jiang JJ. Diagnosis and management of chronic laryngitis associated with reflux. *The American journal of medicine*. 2000; 108(Suppl 4a):112S–119S. [PubMed: 10718463]
5. Qadeer MA, Phillips CO, Lopez AR, et al. Proton pump inhibitor therapy for suspected GERD-related chronic laryngitis: a meta-analysis of randomized controlled trials. *The American journal of gastroenterology*. 2006; 101:2646–2654. [PubMed: 17037995]
6. Cohen SM, Kim J, Roy N, Courey M. Assessing factors related to the pharmacologic management of laryngeal diseases and disorders. *Laryngoscope*. 2013; 123:1763–1769. [PubMed: 23720214]
7. Belafsky PC, Postma GN, Koufman JA. The validity and reliability of the reflux finding score (RFS). *Laryngoscope*. 2001; 111:1313–1317. [PubMed: 11568561]
8. Barry DW, Vaezi MF. Laryngopharyngeal reflux: More questions than answers. *Cleveland Clinic journal of medicine*. 2010; 77:327–334. [PubMed: 20439565]
9. Milstein CF, Charbel S, Hicks DM, Abelson TI, Richter JE, Vaezi MF. Prevalence of laryngeal irritation signs associated with reflux in asymptomatic volunteers: impact of endoscopic technique (rigid vs. flexible laryngoscope). *Laryngoscope*. 2005; 115:2256–2261. [PubMed: 16369176]
10. Hicks DM, Ours TM, Abelson TI, Vaezi MF, Richter JE. The prevalence of hypopharynx findings associated with gastroesophageal reflux in normal volunteers. *J Voice*. 2002; 16:564–579. [PubMed: 12512644]
11. Powell J, Cocks HC. Mucosal changes in laryngopharyngeal reflux--prevalence, sensitivity, specificity and assessment. *Laryngoscope*. 2013; 123:985–991. [PubMed: 23208751]
12. Vaezi MF, Hicks DM, Abelson TI, Richter JE. Laryngeal signs and symptoms and gastroesophageal reflux disease (GERD): a critical assessment of cause and effect association. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association*. 2003; 1:333–344. [PubMed: 15017651]
13. Becker V, Graf S, Schlag C, et al. First agreement analysis and day-to-day comparison of pharyngeal pH monitoring with pH/impedance monitoring in patients with suspected laryngopharyngeal reflux. *Journal of gastrointestinal surgery : official journal of the Society for Surgery of the Alimentary Tract*. 2012; 16:1096–1101. [PubMed: 22450948]
14. Hoppo T, Sanz AF, Nason KS, et al. How much pharyngeal exposure is “normal”? Normative data for laryngopharyngeal reflux events using hypopharyngeal multichannel intraluminal impedance (HMII). *Journal of gastrointestinal surgery : official journal of the Society for Surgery of the Alimentary Tract*. 2012; 16:16–24. discussion 24–15. [PubMed: 22033702]
15. Furtwaengler NA, de Visser RO. Lack of international consensus in low-risk drinking guidelines. *Drug and alcohol review*. 2013; 32:11–18. [PubMed: 22672631]
16. Kenford SL, Wetter DW, Welsch SK, Smith SS, Fiore MC, Baker TB. Progression of college-age cigarette smokers: what influences outcome. *Addictive behaviors*. 2005; 30:285–294. [PubMed: 15621399]
17. Husten CG. How should we define light or intermittent smoking? Does it matter? *Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco*. 2009; 11:111–121. [PubMed: 19246425]

18. Belafsky PC, Postma GN, Amin MR, Koufman JA. Symptoms and findings of laryngopharyngeal reflux. *Ear, nose, & throat journal*. 2002; 81:10–13.
19. Belafsky PC, Postma GN, Koufman JA. The association between laryngeal pseudosulcus and laryngopharyngeal reflux. *Otolaryngol Head Neck Surg*. 2002; 126:649–652. [PubMed: 12087332]
20. Johnson LF, DeMeester TR. Development of the 24-hour intraesophageal pH monitoring composite scoring system. *Journal of clinical gastroenterology*. 1986; 8(Suppl 1):52–58. [PubMed: 3734377]
21. Lee JH, Park SY, Cho SB, et al. Reflux episode reaching the proximal esophagus are associated with chronic cough. *Gut and liver*. 2012; 6:197–202. [PubMed: 22570748]
22. Shay S, Tutuian R, Sifrim D, et al. Twenty-four hour ambulatory simultaneous impedance and pH monitoring: a multicenter report of normal values from 60 healthy volunteers. *The American journal of gastroenterology*. 2004; 99:1037–1043. [PubMed: 15180722]
23. Cho YK. How to interpret esophageal impedance pH monitoring. *Journal of neurogastroenterology and motility*. 2010; 16:327–330. [PubMed: 20680174]
24. Thibeault SL, Smith ME, Peterson K, Ylitalo-Moller R. Gene expression changes of inflammatory mediators in posterior laryngitis due to laryngopharyngeal reflux and evolution with PPI treatment: a preliminary study. *Laryngoscope*. 2007; 117:2050–2056. [PubMed: 17721405]
25. Rosen CA. Stroboscopy as a research instrument: development of a perceptual evaluation tool. *Laryngoscope*. 2005; 115:423–428. [PubMed: 15744150]
26. Postma GN, Halum SL. Laryngeal and pharyngeal complications of gastroesophageal reflux disease. *GI Motility Online*. 2006
27. Lee BE, Kim GH, Ryu DY, et al. Combined Dual Channel Impedance/pH-metry in Patients With Suspected Laryngopharyngeal Reflux. *Journal of neurogastroenterology and motility*. 2010; 16:157–165. [PubMed: 20535346]
28. Weusten BL, Roelofs JM, Akkermans LM, Van Berge-Henegouwen GP, Smout AJ. The symptom-association probability: an improved method for symptom analysis of 24-hour esophageal pH data. *Gastroenterology*. 1994; 107:1741–1745. [PubMed: 7958686]
29. Tutuian R, Castell DO. Review article: complete gastro-oesophageal reflux monitoring - combined pH and impedance. *Aliment Pharmacol Ther*. 2006; 24(Suppl 2):27–37. [PubMed: 16939430]
30. Hickson C, Simpson CB, Falcon R. Laryngeal pseudosulcus as a predictor of laryngopharyngeal reflux. *Laryngoscope*. 2001; 111:1742–1745. [PubMed: 11801937]

Table 1

Reflux Finding Score rating rubric adapted from Belafsky, Postma, & Koufman

REFLUX FINDING SCORE		COMMENTS
Subglottic Edema (pseudosulcus; aka “infraglottic edema”)	2 = present 0 = absent	
Ventricular Obliteration (false vocal fold edge is indistinct; “complete” refers to the true and false folds appearing to touch)	2 = partial 4 = complete	
Erythema/Hyperemia (redness)	2 = arytenoids only 4 = diffuse	
Vocal Fold Edema (mild is slight swelling, moderate is more perceptible, severe is sessile)	1= mild 2= moderate 3= severe 4= polypoid	
Diffuse Laryngeal Edema (size of airway relative to size of larynx)	1= mild 2= moderate 3= severe 4= obstructing	
Posterior Commissure Hypertrophy (pachydermia; mild is mustache-like appearance, moderate is straight line across back of larynx, severe is bulging into airway, and obstructing is airway obliterated)	1= mild 2= moderate 3= severe 4= obstructing	
Granuloma/Granulation	2 = present 0 = absent	
Thick Endolaryngeal Mucus	2 = present 0 = absent	
Total =		

Table 2

Summary of Reflux Finding Score (RFS) and multichannel intraluminal impedance pH monitoring (MII/pH) variables.

RFS Variables	Mean	S.D.	Min	Max
Subglottic Edema	0.78	0.74	0	2
Ventricular Obliteration	1.77	0.83	0	3.75
Erythema/Hyperemia	2.91	0.87	0.5	4
Vocal Fold Edema	1.25	0.75	0	3.75
Diffuse Laryngeal Edema	1.05	0.57	0	3
Posterior Commissure Hypertrophy	1.63	0.66	0	3
Granulation/Granuloma	0.38	0.45	0	2
Thick Endolaryngeal Mucus	1.18	0.72	0	2.25
Total	10.38	3.63	1.75	19.875
MI/pH Variables				
<i>Measured by pH monitoring</i>				
% total time pH <4	3.34	7.95	0	80.5
% upright time pH <4	3.69	7.70	0	62.8
% supine time pH <4	2.56	9.41	0	94.9
Number of reflux episodes	19.50	14.20	0	76
Number of reflux episodes > 5 minutes	1.09	2.64	0	18
Longest reflux episode (minutes)	13.84	43.69	0	444.8
Johnson/DeMeester Score	14.78	31.85	0.8	256.0
<i>Measured by multichannel intraluminal impedance</i>				
Acid refluxate (minutes)	13.67	15.63	0	102.9
Nonacid refluxate (minutes)	7.08	9.78	0	102.3
Total % time reflux (minutes)	1.79	1.66	0.1	12
Number of reflux events	44.3	21.01	7	105
Number of acid reflux events	24.42	17.28	0	91
Number of nonacid reflux events	19.78	11.56	0	62
Number of reflux events that reached the proximal esophagus	23.75	13.97	3	72
Number of acid reflux events that reached the proximal esophagus	14.69	11.32	0	52
Number of nonacid reflux events that reached the proximal esophagus	9.06	6.65	0	38

Table 3

Participant characteristics

Characteristic	No. (%)	Mean Age (years)
Sex		
Male	64 (45)	40.1
Female	78 (55)	43.5
Cigarette smoking		
Nonsmoker	107 (75)	40.6
Smoker	35 (25)	42.9
Reflux cohort		
GERD	38 (27)	43.4
LPR	44 (31)	37.5
Normal	60 (42)	42.5
Total Reflux Finding Score		
<7	26 (18)	42.0
>7	116 (82)	41.3
<11	87 (61)	41.4
>11	55 (39)	41.2

Table 4

Intra-rater and inter-rater reliability

<i>RFS Variable</i>	Intra-rater Reliability		Inter-rater Reliability	
	R	p-value	R	p-value
Subglottic Edema	0.05	0.06	0.48	<0.0001
Ventricular Obliteration	0.45	<0.0001	0.24	<0.0001
Erythema/Hyperemia	0.10	0.001	0.34	<0.0001
Vocal Fold Edema	0.29	<0.0001	0.39	<0.0001
Diffuse Laryngeal Edema	0.17	<0.0001	0.29	<0.0001
Posterior Commissure Hypertrophy	0.021	0.38	0.34	<0.0001
Granulation/Granuloma	0.20	<0.0001	0.21	<0.0001
Thick Endolaryngeal Mucus	0.12	0.0001	0.43	<0.0001
Total	0.21	0.0001	0.48	<0.0001

Interpretation of correlation coefficients: <0.40 = poor; 0.40–0.59 = fair; 0.60–0.74 = good; >0.74 = excellent.

Table 5

Percent intra-rater agreement for each individual rater and RFS rating, as well as averages across raters and RFS ratings.

<i>RFS Variable</i>	Intra-rater % Agreement								Average % Agreement
	R1	R2	R3	R4	R5	R6 *	R7	R8 *	All Raters
Subglottic Edema	70.0	81.8	81.3	72.7	50.0	60.0	90.9	53.8	70.1
Ventricular Obliteration	81.8	53.9	100.0	71.4	75.0	88.9	72.7	56.3	75.0
Erythema/Hyperemia	73.3	57.2	62.5	66.7	61.5	87.5	61.5	66.7	67.1
Vocal Fold Edema	76.9	71.4	68.8	58.3	63.6	25.0	30.8	46.7	55.2
Diffuse Laryngeal Edema	75.0	76.9	18.8	50.0	54.5	50.0	27.3	37.5	48.7
Posterior Commissure Hypertrophy	42.9	58.3	56.3	58.3	46.2	62.5	66.7	57.2	56.0
Granulation/Granuloma	75.0	100.0	81.3	90.0	75.0	77.8	66.7	64.3	78.7
Thick Endolaryngeal Mucus	78.6	57.1	68.8	68.7	92.9	66.7	92.3	56.3	72.6
Average % Agreement	71.7	69.6	67.2	67.0	64.8	64.8	63.6	54.8	

* Indicates rater did not attend training.

Table 6

Summary of results of generalized linear modeling demonstrating the total variance (R^2) accounted for by: 1) the main effect of age; 2) the main and interaction effects of cohort, sex, smoking status; and 3) the main and interaction effects of cohort, sex, smoking status, and age for each RFS variable.

RFS Variable	Age		Model 1		Model 2 (with Age)	
	R^2	P	R^2	P	R^2	P
Subglottic Edema	0.001	0.69	0.11	0.19	0.16	0.51
Ventricular Obliteration	0.08	0.0006	0.29	<0.0001	0.40	<0.0001
Erythema/Hyperemia	0.01	0.19	0.35	<0.0001	0.39	<0.0001
Vocal Fold Edema	0.04	0.03	0.33	<0.0001	0.39	<0.0001
Diffuse Laryngeal Edema	0.03	0.04	0.28	<0.0001	0.37	<0.0001
Posterior Commissure Hypertrophy	0.03	0.04	0.17	0.01	0.25	0.03
Granulation/Granuloma	0.03	0.03	0.05	0.80	0.28	0.01
Thick Endolaryngeal Mucus	0.03	0.70	0.10	0.20	0.16	0.47
Total	0.02	0.09	0.35	<0.0001	0.39	<0.0001