The Reliability of the Reflux Finding Score Among General Otolaryngologists

Brent A. Chang, S. Danielle MacNeil, Murray D. Morrison, and Patricia K. Lee, *Vancouver, British Columbia, Canada*

Summary: Background. The reflux finding score (RFS) is a validated clinical severity scale for findings of laryngopharyngeal reflux (LPR) on fiberoptic laryngoscopy. To our knowledge, there have been no studies to determine whether severity of patient symptoms influence the RFS; in addition, the reliability of the RFS has not been tested for general otolaryngologists.

Objectives. The objectives of this study were (1) to determine whether the RFS for LPR is influenced by symptoms of reflux and (2) to determine the inter-rater reliability for general otolaryngologists in diagnosing LPR using the RFS.

Methods. Ten general otolaryngologists were selected to participate. Participants were asked to complete an Internet survey consisting of flexible endoscopic videos of larynges with varying physical findings of reflux and grade the severity of reflux using the RFS. The videos were randomly shown with and without accompanying patient symptoms.

Results. Our data suggest that patient symptoms influence the RFS. Inter-rater reliability for general otolaryngologists using the RFS is fair.

Conclusions. Among general otolaryngologists in our study, the reliability and objectivity of the RFS in diagnosing reflux cannot be demonstrated.

Key Words: Laryngopharyngeal reflux–Reflux finding score–Rater reliability.

INTRODUCTION

Laryngopharyngeal reflux (LPR) is the retrograde movement of gastric contents into the larynx, pharynx, and upper digestive tract.[1] Patients with LPR are predominantly upright (daytime) refluxers, and the primary defect in LPR is thought to possible be related to upper esophageal sphincter dysfunction.[2] LPR is a relatively common entity with one study reporting up to 10% of patients presenting to an otolaryngologist's office having the disease.[3]

Signs and symptoms of LPR include hoarseness, vocal fatigue, excessive throat clearing, globus pharyngeus, chronic cough, postnasal drip, and dysphagia.[4] The physical findings of LPR on fiberoptic endoscopy that are most commonly reported are edema and erythema of the larynx.[5,6] At least two studies have found that laryngeal abnormalities can be identified in the healthy asymptomatic population.[7,8] In addition, Milstein et al[8] found that the choice of endoscopic instrument influenced laryngoscopic findings with flexible fiberoptic laryngoscopy being more likely to falsely suggest the presence of erythema and edema than rigid laryngoscopy. Branski et al[9] found poor inter-rater and intrarater reliability for gastroesophageal reflux disease-related laryngeal signs among five otolaryngologists who analyzed and scored 120 video segments from rigid fiberoptic examinations.

Although individual findings of LPR alone are poor predictors of symptomatic LPR, Belafsky et al[1] were able to validate an RFS that is predictive of LPR. The reflux finding score (RFS) is a validated eight-item clinical severity scale based on findings from fiberoptic laryngoscopy. The scale ranges from 0 (no abnormal findings) to a maximum of 26 (worst possible score) (Appendix 1).[1] The eight items included in the scale are infraglottic edema, ventricular obliteration, erythema/hyperemia, vocal fold edema, diffuse laryngeal edema, posterior commissure hypertrophy, granuloma/granulation tissue, and thick endolaryngeal mucus. The most frequent finding of persons with LPR was posterior laryngeal hypertrophy, which was documented in 85% of all patients before initiation of treatment.[1] A significant difference was found between the scores of 40 individuals with hypopharyngeal reflux documented by pH study and 40 normal controls.[1] They found intraobserver reliability of 95% and interobserver reliability of 90% for two laryngology trained observers.[1] Ventricular obliteration is also a relatively frequent finding in patients with LPR (80%).[1] A patient with an RFS greater than 7 can be diagnosed with LPR based on a 95% confidence interval.[1]

Laryngologists have fellowship training beyond that required to practice general otolaryngology. Belafsky demonstrated fellowship-trained laryngologists have been shown to reliably diagnose LPR using the RFS; however, a great deal of LPR is diagnosed and treated by general otolaryngologists. Although Branski et al assessed reliability among "board-certified otolaryngologists with diverse clinical interests," to date, there has been a paucity of studies specifically in general otolaryngologists (without subspecialty laryngology training) and their reliability in diagnosing LPR using the RFS.[9] In addition, the RFS has not been independently validated.

Although the RFS is used to score physical findings of reflux, the reflux symptom index (RSI) is a nine-item patient questionnaire that can be used to quantify symptom severity in patients with LPR complaints (Appendix 2).[4] The index was validated by comparing patients with LPR documented by 24-hour pH study to asymptomatic controls and by documenting improvement of the index scores in patients after

Accepted for publication October 15, 2014.
This material has never been published and is not currently under evaluation in any other peer-reviewed publication.
No grants or financial support was used for this study.
This article was presented at 62nd Annual Meeting of the Canadian Society of Otolaryngology, Jasper, Alberta, June 1–4, 2008.
From the Department of Surgery, Division of Otolaryngology, University of British Columbia, Vancouver, British Columbia, Canada.
Address correspondence and reprint requests to Brent A. Chang, Department of Surgery, Division of Otolaryngology, University of British Columbia, Gordon and Leslie Diamond Health Care Center, 4th Floor ENT Clinic (Otolaryngology), 2775 Laurel St. Vancouver, British Columbia V5Z 1M9, Canada. E-mail: brent.a.chang@gmail.com
Journal of Voice, Vol. 29, No. 5, pp. 572-577
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2014.10.009

treatment with proton pump inhibitor therapy. An index score greater than 13 is considered indicative of LPR.[4] It is possible that patient symptoms may influence the RFS. We intended to determine whether the severity of endoscopic findings of LPR as rated by general otolaryngologists is influenced by patient symptoms of LPR.

METHODS

Video acquisition

One of the authors (M.D.M.) routinely generates laryngoscopic video recordings of patients presenting to his office, mostly using a rigid laryngoscope. Flexible endoscopic video recordings were selected because most general otolaryngologists use this technology. We obtained all the flexible laryngoscopic videos of patients with findings suggestive of LPR from January 2005 to January 2007, and 30 patient videos were obtained in total. Dr. Morrison, a fellowship-trained laryngologist, assigned an RFS for each video. The videos were then stratified into five equal groups by RFS to ensure sampling of varying severities. Two videos were randomly selected from each group. A 30-second video clip of each larynx was made and uploaded onto a Web site. The patient's charts were accessed to record the presenting LPR symptoms of each patient and to generate an RSI based on the patient's symptoms. University of British Columbia Research Ethics Board approval was obtained for the use of patient video clips, access to patient charts, and for the questionnaire distribution.

Questionnaire development

An Internet-based questionnaire was developed containing an information page, and photographic examples of pseudosulcus, ventricular obliteration, vocal fold edema, and posterior commissure hypertrophy taken from the original publication describing the RFS.[1] The remaining pages contained the scale for rating each video. Finally, the questionnaire contained demographic questions about the otolaryngologists. Participants were asked to provide their contact information at the end of the questionnaire if they wished to receive a summary of the results.

Participant recruitment

Practicing otolaryngologists who were members of the British Columbia Society of Otolaryngology were invited to participate. Exclusion criteria were subspecialty training in laryngology and otolaryngologists with a practice comprised of over 50% laryngology. Otolaryngologists were randomly contacted to participate until a total of 10 were recruited. For those otolaryngologists that wished to participate, an e-mail was sent to them containing an information letter acting as the consent form, a participant number to ensure anonymity, and access information to the survey Web site.

Questionnaire administration

The first 10 videos of larynges were presented in a random order with accompanying RSI score, followed by the same videos presented in a different random order without an accompanying RSI. Each larynx was rated by the otolaryngologist using the RFS. It was possible to watch each video clip more than once. Once the otolaryngologists had completed the RFS and advanced to the next page, it was not possible to go back and change any previous ratings.

Data analysis

Statistical analysis was performed with *SPSS 10* for Windows (SPSS Inc., Chicago, IL, USA). Overall descriptive statistics, including the mean, range, and standard deviation (SD) for continuous variables and frequencies were calculated. The primary end point was whether the impact of patient symptoms would bias the RFS. The random effect model was applied to evaluate the "symptom impact" on two sets of patients. Set 1 was the group of patients who had RSI less than 13, and set 2 was the group of patients who had RSI greater than 13. The random effect model estimated the mean of the response. Our secondary end point was to determine the inter-rater reliability of otolaryngologists. Inter-rater reliability was calculated based on Leiss's intraclass correlation coefficient (ICC) in which the raters (otolaryngologists) were treated as a random sample from a large population. Multirater kappa was used to assess inter-rater reliability for the dichotomous items in the RFS.

TABLE 1.
Mean, Standard Deviations, and Range of Total RFS for Each Patient When Rater Presented With Symptoms (RSI) and Without Symptoms (Without RSI)

	RFS With RSI			RFS Without RSI			
Patient	Mean	SD	Range	Mean	SD	Range	Provided RSI
1	16.8	3.88	9–23	18.1	3.78	13–24	22
2	4.2	2.67	0–9	2	1.33	0–4	10
3	10.2	3.79	1–16	9.5	4.01	3–16	36
4	4.7	3.13	0–9	4.5	3.34	0–8	6
5	5	3.7	0–10	4	2.75	0–9	3
6	3.2	3.94	0–13	4.1	5.49	0–19	30
7	15.1	4.63	7–21	14	3.74	8–21	19
8	9.3	4.74	2–15	9.8	4.32	3–16	7
9	12.9	4.09	5–19	11.5	6.74	0–21	4
10	9.5	4.22	4–16	8.6	3.78	2–13	14