Validation of a Quality-of-Life Instrument for Laryngopharyngeal Reflux

Ricardo L. Carrau, MD; Aliaa Khidr, MD, PhD; Karen F. Gold, PhD; Joseph A. Crawley, MS; Eric M. Hillson, PhD; Jamie A. Koufman, MD; Chris L. Pashos, PhD

Objectives: To establish the reliability, validity, and responsiveness of a new, disease-specific assessment tool, the LPR-HRQL, which assesses patient-reported outcomes (PRO) with regard to health-related quality of life (HRQL) of patients with laryngopharyngeal reflux (LPR).

Design: A prospective, open-label, repeated-measures study.

Setting: Six centers in 4 states in the eastern United States.

Patients: Patients with LPR.

Interventions: Open-label treatment with 20 mg of omeprazole twice daily. Clinical and PRO HRQL data were collected. Several PRO instruments were administered to patients at each of several time points; these instruments included the Medical Outcomes Study Short-Form 36-Item Health Survey (SF-36), a general HRQL tool; the Voice Handicap Index (VHI), a symptom-specific tool for assessing voice problems; and the QOLRAD instrument (Quality of Life in Reflux and Dyspepsia), used to assess the impact of gastroesophageal reflux disease.

Results: Factor analyses of the LPR-HRQL scales confirmed single dimensions for each. All LPR-HRQL items contributed to internal consistency of scales and had substantial variability permitting useful information. Substantial evidence of convergent and divergent validity with SF-36, VHI, and QOLRAD items was observed. Test-retest validity was adequate for the time interval tested. Changes in domain scores of the LPR-HRQL at 4 and 6 months documented its responsiveness.

Conclusions: The LPR-HRQL displays reliability, validity, and responsiveness, has face validity, and is simple and not burdensome to administer, score, and analyze. Accordingly, it may be used to assist physicians and patients in understanding the HRQL burden of LPR and the impact of therapy.


The clinical impact of laryngopharyngeal reflux (LPR) has recently been confirmed in a position statement of the American Academy of Otolaryngology–Head and Neck Surgery by Koufman et al,1 which outlines symptoms, clinical manifestations, diagnosis, and treatment. Laryngopharyngeal reflux is a gastrointestinal and otolaryngologic condition related to, but distinct from, gastroesophageal reflux disease (GERD). It is estimated that 4% to 10% of patients presenting to an otolaryngology practice have symptoms and/or findings related to LPR.2 Laryngopharyngeal reflux is increasingly recognized as a probable contributing factor to nonallergic asthma and many ear, nose, and throat complaints. Studies suggest that acid reflux is present in 50% to 80% of patients with asthma, 10% to 20% of patients with chronic cough, up to 80% of patients with difficult-to-manage hoarseness, and 25% to 50% of patients with globus sensation.3-6

The manifestations of LPR vary considerably.7 In contrast to GERD, in which heartburn is the primary symptom, there is no predominant sign or symptom for LPR. Typical symptoms of LPR are chronic but intermittent8 and may include dysphagia, throat clearing, hoarseness, chronic cough, globus sensation, and laryngospasm. Approximately 20% to 43% of patients with LPR experience heartburn,9 and 18% have esophagitis.10

Health-related quality of life (HRQL) has been studied in depth for many diseases, including gastrointestinal disorders.11-14 A variety of generic and disease-specific instruments have been used to measure the impact on HRQL of GERD,15-22 and other instruments have been used to evaluate the impact of voice disorders.23-26 However, the HRQL impact of LPR has only recently begun to be evaluated and is not well under-
Although 1 study has compared the general HRQL impact of LPR with that of GERD and with a general US population, use of a validated disease-specific HRQL instrument would provide physicians, other health care providers, and patients more information on the disease-related HRQL impact of LPR. An LPR-specific instrument would address the HRQL domains affected by the disorder (i.e., it would be more specific) and would be more sensitive to changes in HRQL, notably those due to treatment. Ultimately, such an instrument would then have greater utility in clinical practice as well as in research into new therapies.

Therefore, the objective of the present study was to evaluate the reliability, validity, and responsiveness of a new LPR-specific HRQL instrument, the LPR-HRQL, which was designed to be self-administered by the patient. While several related symptom scales exist, the LPR-HRQL is the first HRQL instrument to be developed specifically for LPR.

**METHODS**

**THE INSTRUMENT**

The LPR-HRQL was developed based on a literature review, patient input obtained in focus-group settings, and input from an expert panel of physicians specializing in otolaryngology, gastroenterology, and pulmonology. These 3 sources provided support for an instrument that would measure symptom distress and the important effects of LPR on social and occupational functioning, vitality, well-being, and perceived health. Accordingly, a brief 43-item questionnaire that uses Likert response scales was constructed that may be self-administered and fulfills these broad measurement needs. In its field-test version, the instrument had 43 items—specifically, questions about LPR and how it affects the patient. The recall period was the last 4 weeks.

A standard 7-point Likert scale was used to assess how much or how often each item described the feelings of the patient. Twelve questions assessed talking, singing, and voice (Voice/Hoarse domain) with a 13th asking how these voice issues affected overall quality of life. Six questions assessed coughing (Cough domain) with a follow-up asking how problems with coughing affected overall quality of life. Another 6 questions addressed clearing the throat (Clear Throat domain) with a follow-up asking how problems with clearing the throat affected overall quality of life. The next 5 questions assessed globus and general throat symptoms (Swallow domain) with a follow-up about general related problems affecting overall quality of life. The final part of the instrument, consisting of questions 34 through 43, assessed the combined impact of acid reflux-related symptoms (Overall Impact of Acid Reflux domain).

**PATIENTS**

To validate the LPR-HRQL, we enrolled patients presenting with symptoms at 6 sites in 4 states in the southern, mid-Atlantic, and northeastern regions of the United States. Sites included 4 academic medical centers, 1 regional medical center, and a multispecialty group practice, so that the study population included various sociodemographic groups receiving care in a variety of health care settings. Local institutional review boards at each site reviewed and approved the implementation of the study protocol.

Patients were all newly diagnosed as having LPR (diagnosis <1 month before enrollment) or were patients with relapse and not under current treatment; this was confirmed by a physician in each case by direct patient examination, as is typical practice. Each patient gave written informed consent prior to any study procedures. Enrolled patients were between 19 and 80 years old and had at least 1 of the following symptoms commonly associated with LPR within the previous month: hoarseness, chronic cough (defined as cough lasting >1 month), globus, laryngospasm, chronic throat clearing, or difficulty swallowing. Patients were not enrolled in the study if they manifested only GERD symptoms and not LPR.

A laryngoscopic examination documenting clinical signs consistent with a diagnosis of LPR was conducted within 1 month prior to patient enrollment for all patients. Furthermore, the presence and severity of LPR clinical signs for each patient were assessed using the Reflux Finding Score developed by Belafsky et al. The Reflux Finding Score facilitated evaluation of the presence and severity of the following clinical signs: subglottic edema, ventricular obliteration, aryttenoid erythema, vocal fold edema, diffuse laryngeal edema, posterior commissure hypertrophy, granuloma and/or granulation, and pachydermia laryngis.

Patients with cancer, major psychiatric illness, and/or unstable chronic illnesses (such as diabetes) were excluded to eliminate comorbidities that might confound or inhibit an assessment of the effect of LPR on HRQL. In addition, patients requiring medications with known drug-drug interactions with omeprazole were also excluded, further reducing the number of participants with comorbid conditions.

**STUDY DESIGN**

Data were collected from patients with LPR at a baseline visit and 3 times subsequently during regularly scheduled study visits. Baseline and 2-month postbaseline data were used to establish the validity and reliability of the LPR-HRQL. Data from the 4-month and 6-month study visits were used to establish the instrument’s responsiveness.

**INSTRUMENTATION**

A battery of instruments was administered to patients at baseline and at 2, 4, and 6 months after baseline. Aside from the LPR-HRQL, these instruments included the Medical Outcomes Study Short-Form 36-Item Health Survey (SF-36), the Voice Handicap Index (VHI), the QOLRAD instrument (Quality of Life in Reflux and Dyspepsia), and Overall Treatment Effect (OTE) questionnaires. The SF-36, a self-administered, generic HRQL questionnaire containing 36 items, measures health in 8 multi-item dimensions, covering functional status, well-being, and overall evaluation of health. The reliability and validity of the SF-36 is well documented. The 30-item VHI measures the impact of voice problems on a person’s life. The QOLRAD measures the HRQL associated with GERD and dyspepsia. The 4-item OTE questionnaire was designed to address treatment effectiveness and the impact of treatment on HRQL of patients with LPR. The instrument’s responsiveness was assessed by measuring changes in HRQL scores with respect to the time of the follow-up assessment. The study protocol ensured the validity and reliability of the LPR-HRQL instrument.
was based on a recommendation by a panel of otolaryngologists with a high volume of referred patients with LPR as the minimum required to achieve relief of symptoms. For the purposes of homogeneity of effects in testing the sensitivity of the instrument to change, patients were instructed that no other prescription therapy for acid-related symptoms was permitted during the study. Previous use of antisecretory or promotility agents was permitted, as long as these treatments were discontinued at least 10 days prior to study enrollment and the baseline evaluation.

STATISTICAL ANALYSIS

Validity

Face validity was established by the evaluation of the substantive content by a panel of physician experts. In addition, information was collected from patients during the pilot test to assess the patient perception of the face validity of the instrument. Statistical, construct, and concurrent validity were assessed. The primary tools for assessing statistical validity were descriptive statistics and distribution graphs. In the assessment of construct validity, factor analysis was conducted of all domains to determine if they reflected individual, coherent factors. Zero order correlations among the items and subscale scores were examined for discriminant and convergent validities, which provided an assessment of concurrent validity. Concurrent validity was further examined by analyzing the relationships between the LPR-HRQL and the VHI, SF-36, and QOLRAD.

Reliability

Reliability was established through item analysis. To assess internal consistency reliability, we computed a Cronbach’s α for each domain of the LPR-HRQL. To ensure that the questionnaire was stable across time, we performed a test-retest reliability analysis by examining Pearson correlations and comparing 4-month and 6-month evaluations. These time points were selected to coincide with the time subsequent to treatment initiation when the therapy should have achieved full efficacy (ie, month 4 and later) and the patient could be considered “treatment stable.”

Responsiveness

The responsiveness index was calculated as the change from baseline to end point to represent the improvement in function under a known effective treatment regimen. A paired t test compared baseline and 6-month scores. To facilitate a meaningful clinical interpretation of domain score changes, we calculated the minimum clinically meaningful change in each domain to correspond to a significant (1-point) decrease in the physician-reported symptom severity score.

Burden

The ability of respondents to complete the instrument was evaluated, and factors of time needed for completion and language complexity were considered.

RESULTS

In the assessment of the patient perception of the face validity of the instrument, no items were flagged by subjects as inappropriate or unclear.

Of the 117 subjects enrolled, 78 (66.7%) were women, 101 (86.3%) were white, and the mean (SD) age was 48.4 years. Overall, 28.2% of the sample were younger than 40 years; 31.6% were aged between 40 and 50 years; and 40.2% were older than 50 years. This age group distribution did not differ significantly between men and women. At baseline, 85.5% reported chronic throat clearing, 82.1% globus, 80.3% hoarseness, 53.9% difficulty swallowing, 44.3% chronic cough, and 33.3% laryngospasm.

Table 1 lists the item-level descriptive statistics grouped by domain for the LPR-HRQL. All items showed good variability, and the ranges spanned all possible val-
Table 2. Scores on LPR-HRQL Domains, VHI, and QOLRAD

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Mean (SD)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPR-HRQL domain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voice/Hoarse</td>
<td>19.71 (16.09)</td>
<td>0-65</td>
</tr>
<tr>
<td>Cough</td>
<td>7.61 (8.76)</td>
<td>0-35</td>
</tr>
<tr>
<td>Clear Throat</td>
<td>9.63 (7.58)</td>
<td>0-36</td>
</tr>
<tr>
<td>Swallow</td>
<td>7.64 (5.84)</td>
<td>0-30</td>
</tr>
<tr>
<td>Overall Impact of Acid Reflux</td>
<td>32.14 (25.93)</td>
<td>9-90</td>
</tr>
<tr>
<td>VHI</td>
<td>29.86 (25.93)</td>
<td>0-109</td>
</tr>
<tr>
<td>QOLRAD</td>
<td>1.83 (1.40)</td>
<td>0-5.4</td>
</tr>
</tbody>
</table>

Abbreviations: LPR-HRQL, Laryngopharyngeal Reflux Health-Related Quality of Life questionnaire; QOLRAD, Quality of Life in Reflux and Dyspepsia instrument; VHI, Voice Handicap Index.

Table 3. Correlation of LPR-HRQL Domains With Validated Instruments*

<table>
<thead>
<tr>
<th>LPR-HRQL Domain</th>
<th>VHI</th>
<th>QOLRAD</th>
<th>Acid Reflux Finding Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice/Hoarse</td>
<td>0.88 (.01)</td>
<td>0.24 (.01)</td>
<td>0.16 (.08)</td>
</tr>
<tr>
<td>Cough</td>
<td>0.45 (.01)</td>
<td>0.37 (.01)</td>
<td>−0.06 (.52)</td>
</tr>
<tr>
<td>Clear Throat</td>
<td>0.56 (.01)</td>
<td>0.41 (.01)</td>
<td>0.24 (.01)</td>
</tr>
<tr>
<td>Swallow</td>
<td>0.41 (.01)</td>
<td>0.59 (.01)</td>
<td>0.23 (.01)</td>
</tr>
<tr>
<td>Overall Impact of Acid Reflux</td>
<td>0.61 (.01)</td>
<td>0.66 (.01)</td>
<td>0.18 (.06)</td>
</tr>
</tbody>
</table>

Abbreviations: LPR-HRQL, Laryngopharyngeal Reflux Health-Related Quality of Life questionnaire; QOLRAD, Quality of Life in Reflux and Dyspepsia instrument; VHI, Voice Handicap Index.

*Data are reported as Pearson correlation coefficient (P value).

Histograms were examined, and no items showed evidence of multimodal distributions. There were no substantial floor or ceiling effects. There were no missing item patterns to be analyzed. Respondents completed the LPR-HRQL in approximately 30 minutes or less.

All but 1 domain showed a single-factor structure with items loading substantially on that factor. The Voice/Hoarse domain was the only domain that generated more than a 1-factor structure. Examination of the results revealed that the second factor, with an eigenvalue marginally greater than 1, was attributable to the item that required reverse scoring to be integrated into the scale correctly.

Table 2 lists the descriptive statistics for the 5 domains of the LPR-HRQL instrument as well as for 2 validated instruments used to assess convergent and divergent validity. Examination of the histograms of the scales indicates normal distributions with no substantial ceiling or floor effects. The variability, as compared with the potential range for each of the scores of the LPR-HRQL, compares favorably with that of the VHI and the QOLRAD.

Table 3 outlines the evidence of substantial convergent validity based on the VHI and QOLRAD but not the Reflux Finding Score. The LPR-HRQL Voice/Hoarse score is most highly correlated with the VHI, and the magnitude of the correlation is substantial. Both are indications of convergent validity for the LPR-HRQL Voice/Hoarse domain component. For the remaining 4 domains of the LPR-HRQL, while they are significantly correlated with the VHI (which is attributable to a general association among symptoms), the magnitude of the correlation is moderate. This indicates that the Cough, Clear Throat, Swallow, and Overall Impact of Acid Reflux scores represent variability in subject symptoms independent of the issues represented by the Voice/Hoarse questions.

Likewise, the relatively moderate or higher correlations with the QOLRAD (which assesses GERD and dyspepsia) are in the Swallow and Overall Impact of Acid Reflux scores. The LPR-HRQL scores reflecting the symptoms that we would expect to distinguish GERD and dyspepsia from LPR (Voice/Hoarse, Cough and Clear Throat) have moderate to low correlation with the QOLRAD, which provides divergent validity evidence for the LPR-HRQL. The Reflux Finding Score, documented by the physician at baseline, has somewhat limited convergent validity. It is an amalgam of symptom questions relating to all the domains covered in the LPR-HRQL, so the correlation reflects the relationship of a part (LPR-HRQL domain) to the whole (overall reflux finding score).

Table 4 reports the associations of the domains of the LPR-HRQL with the domains of the SF-36. These associations provide evidence of concurrent validation and also reveal how scores on the new disease-specific HRQL instrument relate to general domains of HRQL in an established instrument. The signs of the correlations are appropriately negative. The LPR-HRQL domains are most closely associated with the Vitality and Social Function domains of the SF-36. Voice/Hoarse issues are substantially correlated with Social Function (r = −.51). Cough, Clear Throat, and Swallow are similarly associated with both Vitality and Social Function (mean r = −.41). In addition, Swallow is substantially associated with Pain, Role Emotional, and Role Physical. There are smaller, though significant, levels of association across all the LPR-HRQL domains and the Mental Health domain. Significant correlations across the board indicate that the LPR-HRQL captures the impact of LPR on the overall HRQL as measured by a validated, general-profile instrument. Finally, the Overall Impact of Acid Reflux score has the highest correlations with the SF-36 domains compared with the other LPR-HRQL domains.

The overall reliabilities of the LPR-HRQL domains tend to be either superior or highly acceptable (α = .84-.93), though the Swallow score is somewhat lower (α = .69). Removal of any given single item from the score did not improve the LPR-HRQL’s reliability.

The data provide evidence of substantial test-retest reliability. The test-retest reliability correlation coefficients were all significant and substantial, ranging from 0.90 to 0.64, indicating very high to moderate levels of test-retest reliability. Specifically, correlation coefficients for domains were 0.77 for Voice, 0.64 for Cough, 0.86 for Clear Throat, 0.83 for Swallow, and 0.89 for Overall Impact of Acid Reflux. In comparison, the test-retest reliability coefficients for the VHI and the QOLRAD are 0.70 and 0.64, respectively.

The final set of results establishes the responsiveness of the instrument to changes over time and helps establish the minimal clinically meaningful difference for the LPR-HRQL scores. Table 5 lists the dependent t test results for each of the 5 domains of the LPR-HRQL for changes.
from baseline to 4 and 6 months. For all domains of the LPR-HRQL for both the 4- and 6-month measurements, HRQL problems significantly decreased. The minimum clinically meaningful change for each of the Voice/Hoarse, Cough, Clear Throat, and Swallow domains was 5 points, whereas the comparable minimum change for the Overall Impact of Acid Reflux score was 10 points. The Voice/Hoarse, Swallow, and Overall Impact of Acid Reflux scores of the LPR-HRQL exhibited average change in scores between baseline and 6 months that were at or greater than this clinically meaningful minimum difference. Given the standard deviations associated with each domain score, the instrument should detect this 5-point difference with a power of 80% or better in a study comparing 2 arms each of between 20 and 30 subjects.

The LPR-HRQL has been shown to be a reliable and valid instrument that poses no unusual burden on the subject, has face validity, and is simple to administer, score, and analyze. Furthermore, it has been documented to show significant responsiveness to changes in subjects’ HRQL associated with LPR.

Specific findings related to the LPR-HRQL are noteworthy. That there were no missing item patterns to be analyzed suggests that subjects understood the questions and were comfortable answering all of them. That all subjects completed the entire instrument also indicates that the instrument is not overly burdensome.

Significant correlations across the board between the LPR-HRQL domains and those of the SF-36 indicate that the LPR-HRQL captures the impact of LPR on the overall HRQL as measured by a validated, general-profile instrument. That the Overall Impact of Acid Reflux score has the highest correlations with the SF-36 domains compared with the other LPR-HRQL domains makes sense in that an overall disease-specific HRQL score should have a greater correlation with overall health domains than with more limited aspects such as Voice/Hoarse or Cough (the “whole to whole” correlations should be greater than the “part-to-whole” correlations).

The overall reliabilities of the LPR-HRQL domains tend to be highly acceptable, though the Swallow score was in the modest range. Examination of the Swallow scale revealed that the items had a far greater span of symptoms than the other scales, which would explain its reduced reliability. The decision was made to retain the domain to maintain the span of the instrument’s coverage. Given the standard for reliability in related disease instruments, it was decided that maintaining all domains would produce a more valuable instrument than one that was shortened and thereby reduced in its span of coverage. The test-retest reliability correlation coefficients were all significant and favorable compared with those of the VHI and QOLRAD, 2 frequently used and validated instruments, which suggests that the test-retest reliability for the LPR-HRQL meet the same standard of existing instruments.

We anticipated that an instrument useful in the study of LPR would show a significant improvement in scores between pretreatment and posttreatment status. This occurred for all domains of the LPR-HRQL at both the 4- and 6-month periods, which suggests that all LPR-HRQL domains are able to capture the response to an intervention as early as 4 months into treatment. The Voice/Hoarse, Swallow, and Overall Impact of Acid Reflux scores of the LPR-HRQL exhibited average changes between baseline
and 6 months that were at or greater than the clinically meaningful minimum difference, which suggests that the minimum clinically meaningful difference of 5 points for each domain score may be conservative.

The introduction of this instrument to the armamentarium of disease-specific HRQL instruments will enable physicians and health care practitioners to assess the HRQL of their patients with LPR and to evaluate their response to therapy. This instrument can also assist researchers who are conducting clinical studies to assess the ability of new and/or existing therapies to reduce the HRQL burden of reflux laryngitis.

Submitted for Publication: June 1, 2004; final revision received November 15, 2004; accepted January 14, 2005. Author Affiliations: Department of Otolaryngology, University of Pittsburgh Medical Center, Pittsburgh, Pa (Dr Carrau); Otolaryngology Department, University of Virginia School of Medicine, Charlottesville (Dr Khidir); Health Economic Research and Quality of Life Evaluation Services (HERQuLES), Abt Associates Inc, Cambridge, Mass (Drs Gold and Pashos); AstraZeneca LP, Wilmington, Del (Mr Crawley and Dr Hillson); and Center for Voice Disorders of Wake Forest University, Winston-Salem, NC (Dr Koufman).

Correspondence: Chris L. Pashos, PhD, HERQuLES, Abt Associates Inc, 181 Spring St, Lexington, MA 02421 (chris_pashos@abtassoc.com).

Funding/Support: Funding for this research was provided by AstraZeneca LP.

Acknowledgment: We thank William Lenderking, PhD, Richard Berzon, DrPh, and Abt Associates Inc for their contributions to the design, data collection, and instrument development; Sandra Sweeney, AB, for initial project contributions to the design, data collection, and instrument development; Judy Davis, MPH, MSW, for project management; and Lark Madoo, BA, for data processing and analysis. We also thank David Moore, MD, James Reibel, MD, Paul Castellanos, MD, Beverly Prince, MD, Kirk Tolhurst, MD, and their clinic staffs as well as the clinic staffs of Drs Carrau, Khidir, and Koufman for their assistance in enrollment of the subjects into the study.

REFERENCES


(Reprinted) ARCH OTOLARYNGOL HEAD NECK SURG/VOL. 131, APR 2005 WWW.ARCHOTO.COM

©2005 American Medical Association. All rights reserved.