The Laryngoscope

Systematic Review

Instruments Evaluating the Clinical Findings of Laryngopharyngeal Reflux: A Systematic Review

Jerome R. Lechien, MD, PhD, MSc; Antonio Schindler, MD, PhD; Lisa G. De Marrez, MD; Abdul Latif Hamdan, MD, MPH, FACS; Petros D. Karkos, MD, PhD; Bernard Harmegnies, PhD; Maria Rosaria Barillari, MD, PhD; Camille Finck, MD, PhD†; Sven Saussez, MD, PhD†

Objectives: To identify the instruments for evaluating the clinical findings (ICFs) of laryngopharyngeal reflux (LPR) designed for use with regard to diagnosis and treatment effectiveness.

Methods: The PubMed, Scopus, and Cochrane databases were used to search for subject headings following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) recommendations. Three investigators retrieved relevant studies published between 1990 and 2018 describing the evolution of laryngopharyngeal findings throughout LPR treatment. Issues of clinical relevance, that is, LPR diagnosis, treatments, and signs assessed for diagnosis or as therapeutic outcomes, were assessed. The investigators also evaluated the psychometric properties (conceptual model, content validity, consistency, reliability, concordance, convergent validity, known-groups validity, responsiveness to change, and interpretability) of the ICF. The risk of bias was assessed with the tool of the Clarity Group and Evidence Partners.

Results: The search identified 1,227 publications with a total of 4,735 LPR patients; of these studies, 53 met the inclusion criteria. Of these 53 studies, we identified 10 unvalidated and six validated ICFs. None of the validated ICFs included all the psychometric properties. The main identified deficiencies related to ICF psychometric validation included variable construct validity, disparate and uncertain reliabilities, and a lack of interpretability. The lack of consideration of certain LPR laryngeal and extralaryngeal signs is the main weakness of ICFs, biasing content, and construct validities.

Conclusion: The low specificity of LPR signs, the lack of consideration of many findings, and the absence of a gold standard for diagnosis constitute barriers to the further validation of these ICFs. Additional studies are needed to develop complete and reliable ICFs.

Key Words: Laryngopharyngeal, reflux, findings, treatment.

***Laryngoscope*, 00:1–17, 2018**

INTRODUCTION

Laryngopharyngeal reflux disease (LPRD) is the back flow of gastric or gastroduodenal contents into the laryngopharynx where it comes in contact with tissues of the upper aerodigestive tract.[1] According to some U.S. reports, LPRD affects approximately 10% of outpatients of the ear, nose, and throat (ENT) department consultation and up to 50% of patients in the voice center.[2,3] The most common complaints related to LPRD include hoarseness, sore throat, odynophagia, cough, throat clearing, globus sensation, and excessive phlegm.[4,5] Usual gastroesophageal reflux disease (GERD) findings such as heartburn, regurgitations, and esophagitis do not necessarily concern the majority of patients, leading some otolaryngologists to distinguish LPR and GERD as two different clinical entities.[4,6] Moreover, LPRD is characterized by a myriad of clinical signs such as laryngeal edema and erythema, ventricular obliteration, laryngeal keratosis, posterior commissure hypertrophy, pharyngeal wall edema and erythema, tongue tonsil hypertrophy, and erythema of the anterior pillar.[5,7–10] The exact incidence of these signs remains unknown given the mixed results in the literature.

Today, there is no gold standard for the diagnosis of LPRD because pH impedance monitoring has many

From the Laryngopharyngeal Reflux Study Group of Young Otolaryngologists of the International Federations of Oto-rhino-laryngological Societies (YO-IFOS) (J.R.L., A.S., L.G.DM., P.K., B.H., M.R.B., C.F., S.S.); the Department of Anatomy and Experimental Oncology, Mons School of Medicine, UMONS Research Institute for Health Sciences and Technology (J.R.L., L.G.DM., S.S.); the Laboratory of Phonetics, Faculty of Psychology, Research Institute for Language Sciences and Technology (J.R.L., B.H., C.F.), University of Mons (UMons), Mons; the Department of Otorhinolaryngology and Head and Neck Surgery, CHU Saint-Pierre, School of Medicine, Université Libre de Bruxelles (J.R.L., S.S.), Brussels; the Department of Otorhinolaryngology and Head and Neck Surgery, CHU de Liège, Faculty of Medicine, University of Liège (C.F.), Liège, Belgium; the Department of Biomedical and Clinical Sciences, Phoniatric Unit, L. Sacco Hospital, University of Milan (A.S.), Milan; the Division of Phoniatrics and Audiology, Department of Mental and Physical Health and Preventive Medicine, University of Naples SUN (M.R.B.), Naples, Italy; the Department of Otorhinolaryngology and Head and Neck Surgery, American University of Beirut-Medical Center (A.L.H.), Beirut, Lebanon; and the Department of Otorhinolaryngology and Head and Neck Surgery, Thessaloniki Medical School (P.K.), Thessaloniki, Greece

Additional supporting information may be found in the online version of this article.

Editor's Note: This Manuscript was accepted for publication on August 3, 2018.

†These authors contributed equally to this work.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Jerome R. Lechien, MD, PhD, MS, Laboratory of Anatomy and Cell Biology, Faculty of Medicine, University of Mons (UMONS), Avenue du Champ de mars, 6, B7000 Mons, Belgium. E-mail: jerome.lechien@umons.ac.be

DOI: 10.1002/lary.27537

weaknesses such as the cost of the procedure or high false-positive and false-negative rates. Regarding the weaknesses of pH impedance metry, many physicians consider the evolution of signs and symptoms throughout empirical treatment as cost-effective diagnosis methods for LPRD.[6,11] LPRD clinical findings must be assessed with an adequate knowledge of the pathology and with methodological rigor. In fact, the evaluation of both the signs and symptoms related to LPRD using poor instruments with defective methods can have substantial implications for treatment effectiveness, thereby leading to inaccurate and equivocal conclusions. Recent studies have summarized and analyzed patient-reported outcome measurements;[12] to date, however, no study has provided a systematic review of the clinical instruments (assessing signs) used with regard to LPRD.

The first objective of this systematic review was to identify the instruments assessing the clinical findings (ICFs) of LPRD used in both the diagnosis and treatment effectiveness. The second objective was to assess the frequency of ICFs. Finally, the third objective was to evaluate the psychometric properties of these ICFs.

MATERIALS AND METHODS

The criteria for considering studies for this systematic review were based on the Patient/Problem, Intervention, Comparison, Outcome (PICO) framework.

Types of Studies

Clinical and observational studies published as full-scale original articles in peer-reviewed journals. The studies should be written in English or French.

Participants

Adults with suspected or confirmed LPRD. Patients with positive pH metry/impedance and patients who positively responded to an empirical treatment were considered LPR patients. Patients included in the study on the basis of symptoms ± signs without additional examination were considered suspected LPR subjects.

Intervention

The patient may have been treated with medication (i.e., proton-pump inhibitors [PPIs]), alginate, antihistamine, gastroprokinetic), or diet and behavioral changes, or placebo for at least 4 weeks.

Comparison and Outcomes

Authors may have followed natural history of symptoms with no active treatment or not conducted any comparisons.

Search Strategy

The PubMED, Cochrane Library, and Scopus databases were searched to identify studies published between January 1990 and April 2018 that used ENT signs with or without validated ICFs for both the diagnosis and follow up of LPRD. The keywords applied were "reflux," "laryngopharyngeal," "laryngitis," "sign(s)," "measurement," and "gastroesophageal." These words were combined in distinct ways to generate broad research results (Fig. 1). References were also obtained from the citations of retrieved publications or the systematic review. The methodology of this review strictly adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.[13]

Data Selection, Extraction, and Analysis

Three authors (J.R.L., M.R.B., and L.G.DM.) independently reviewed the abstracts of all the identified studies and selected those that met the inclusion criteria for full texts. The authors of the publications were contacted for additional information about their studies when necessary. The investigators compared the age, gender, author, and center data whenever they were available to avoid multiple inclusions of patients. When the same patients were described in more than one publication, authors used only the data reported in the larger and more recent publication. The authors did not exclude any publication based on quality. The following types of studies were excluded: studies focusing on only patient-reported outcomes or health-related quality of life instruments.

Procedure of Outcomes Analysis

The investigators independently identified all studies in which LPRD signs were used as therapeutic outcomes. From this database, the investigators extracted potential ICFs that were assessed with regard to name; acronym; year and reference of the initial publication (development); language and country of origin; objective (diagnosis, therapeutic tool, or both) and target population; characteristics of the included patients for instrument development; setting of development (tertiary care or community); and the ICF characteristics including rated signs, type of scale (i.e., visual analog scale), number of items, calculation, and potential subscales. The investigators analyzed the following properties to assess the psychometric developmental properties of each ICF: 1) conceptual model, 2) content validity, 3) internal consistency, 4) test–retest reliability, 5) concordance, 6) convergent validity, 7) known-groups validity, 8) responsiveness to change, and 9) interpretability and scoring. The analysis of these properties was conducted with regard to definitions described in Table I. The properties' definitions were based on previous analyses of clinical outcome measurement instruments.[12,14,15] Discrepancies in the analysis of psychometric properties were resolved by discussion with senior otolaryngologists (C.F. and S. S.). To ensure wide inclusion, we considered ICFs as validated only when the authors validated at least one developmental measurement property. Thus, at the end of the analysis, we documented validated and unvalidated ICFs. The validation of ICFs in studies was also evaluated for the risk of epidemiological bias using the Tool to Assess the Risk of Bias in Cohort Studies developed by the Clarity Group and Evidence Partners.[16]

RESULTS

Search Results

A total of 72 relevant studies were identified from the PubMED, Scopus, and Cochrane databases. Of these references, 53 met our inclusion criteria (specifically, studies using ENT signs for LPRD diagnosis and as therapeutic outcomes). A total of 4,735 suspected or confirmed LPR patients were included within these articles.[4,5,9–11,17–63] A detailed description of included studies and characteristics (i.e., the number of patients, assessed signs, and treatment)

TABLE I.
Definition of the Measurement Properties of Signs of Instruments Analyzed in the Study.

Domain	Definition
Conceptual model	
Construct definition	It provides a rationale for and description of the concepts and target population that a measure is intended to assess.
Target population	
Expected subscales	
Content validity	It refers to evidence that an instrument is appropriate for its intended use. Items and conceptual domains must be relevant to the targeted population.
Content expert involved	The instrument's development of signs must include direct input from experts. There should be a clear description of the process by which included signs were derived.
Description of item development	The items described in the instrument must reflect the most common signs encountered in the disease.
Reliability	The degree to which scores are free from random (measurement) error.
Internal consistency reliability	Extent to which items within each domain are interrelated.*
Test–retest reliability	Stability of scores over time when no change is expected in the concept of interest.*
Concordance	The degree of agreement among raters.
Construct validity	It refers to whether an instrument measures intended theoretic constructs or traits and directly affects the appropriateness of the measurement-based inferences.
Responsiveness to change	The extent to which an instrument detects meaningful changes over time that have occurred after baseline.*
Convergent validity	The degree to which the sign score correlates with other instruments measuring the same construct or with related clinical indicators.†
Known-groups validity	The extent to which the instrument can discriminate between groups that are known to differ on the variables being measured.*
Interpretability and scoring	The degree to which the meaning of the scores can be easily understood.
Plan for scoring measure	A description of how to score the measure should be provided (sum, algorithm).
Plan for missing data	A prespecified plan for managing missing responses can mitigate the risk of bias resulting from the necessity to exclude cases with missing data.
Scaling described	The process of distributing the full range of respondents' possible scores with respect to the measured attribute.

*Consistent: > 0.70 for group-level comparisons and 0.90–0.95 for individual comparisons.
†< 0.30 = low correlation; 0.30–0.60 = moderate correlation; > 0.60 = strong correlation (Pearson or Spearman analysis).
*Large change: > 0.80; moderate change: 0.50–0.79; small change: 0.2–0.49.

is displayed in Table II. Of the 53 papers, six validated (i.e., Reflux Finding Score [RFS], Chronic Posterior Laryngitis Index [CPLI], Laryngeal Reflux Grade [LRG], laryngopharyngeal reflux disease index [LRDI] and Laryngeal Grading Scale [LGS]; and Table III) and 10 unvalidated ICFs were extracted. Unvalidated ICFs are available in an additional table (Supporting Table SI). The flowchart of this study is shown in Figure 1.

LPRD Findings and Instruments

All standardized or unstandardized instruments were developed at secondary or tertiary academic centers between 1997 and 2014. Specifically, most of the ICFs were elaborated within otolaryngology or head and neck surgery departments with or without collaboration with gastroenterology departments for monocenter (N = 15) or multicenter (N = 1) prospective studies. The instruments were developed in the United States (N = 8), Austria (N = 3), India (N = 1), Japan (N = 1), Lithuania (N = 1), Spain (N = 1), and Australia (N = 1). All studies were published in the English language. The targeted populations were patients with confirmed or suspected LPR, with mean ages of 50.12 and 40.70 years in the validated and unvalidated ICF groups, respectively. The sex ratio of both groups was 6 females/4 males As Tables III and Supporting Table SI). show, ICFs were mainly developed as a therapeutic outcome (N = 13), occasionally for both diagnosis and therapeutic outcomes (N = 3).

The majority of the ICFs used a visual analog scale (VAS) to rate the severity or presence of LPRD signs with numerous items ranging from 3 to 12 (Table III). When the score was not used to assess the prevalence of signs, the calculation method was primarily based on the sum of each item score (N = 10). Three studies did not calculate a total score and only reported individual item scores.

An important level of heterogeneity characterizes the different ICFs with regard to the evaluated findings. In fact, certain laryngeal signs were frequently evaluated (i.e., laryngeal/arytenoids erythema, edema, posterior commissure hypertrophy, vocal fold edema, erythema, granuloma, and granulation) by comparison with pharyngeal and other signs. This heterogeneity was more prevalent in the current literature when focusing on all studies that assessed LPRD signs throughout treatment (Table IV). The most frequently assessed findings as therapeutic outcomes were those described with regard to RFS.

TABLE II.
Characteristics of Included Publications

References	Design	Characteristics	Inclusion Criteria	Signs	Sign Outcomes	Results	ET	Treatment
El-Serag, 2001 (17)	Monocentric	Gr1: suspected LPR (N = 10)	LPR symptoms and signs	LE, EH, GG, UC	Comp. Signs Score	Gr1 = Gr2	12 w	Gr1: Lansoprazole (30 mg 2/d)
	Placebo–RCT	Gr2: suspected LPR (N = 10)						Gr2: Placebo
								Diet: –
Noordzij, 2001 (18)	Monocentric	Gr1: LPR (N = 15)	LPR symptoms	VE, EH, GG, LE,	Comp Signs Score 3	Gr1 = Gr2	8 w	Gr1: Omeprazole (40 mg, 2/d)
	Placebo–RCT	Gr2: LPR (N = 15)	Dual-probe pH metry	TM				Gr2: Placebo (2/d)
								Diet: –
Eherer, 2001 (18)	Monocentric	Gr1: LPR (N = 7)	Laryngeal symptoms	GG, EH, PY, VE,	Comp. Signs Score 7	Gr1 = Gr2	12 w	Gr1: Pantoprazole (40 mg, 2/d)
	Placebo–RCT	Gr2: LPR (N = 7)	Dual-probe pH metry	VR, PH, TM, SE,				Gr2: Placebo
				SP, SR, PP				Diet: –
Steward, 2004 (20)	Monocentric	Gr1: LPR (N = 21)	LPR symptoms and signs	EH, VE, LE, PH,	Comp. Signs Score	Gr1 = Gr2	8 w	Gr1: Rabeprazole (20 mg 2/d)
	Placebo–RCT	Gr2: LPR (N = 21)	Dual-probe pH metry	VR, GG, ND, UC,	Gr1 and Gr2	t1 > t0		Gr2: Placebo
				SE				Diet: +
Vaezi, 2006 (21)	Multicentric	Gr1: suspected LPR (N = 95)	LPR symptoms	EH, GG, LE, PW	CPLI	Gr1 = Gr2	16 w	Gr1: Esomeprazole (40 mg, 2/d)
	Placebo–RCT	Gr2: suspected LPR (N = 50)	CPLI≥5	PH, VR, VE,				Gr2: Placebo
								Diet: –
Wo, 2006 (22)	Monocentric	Gr1: LPR (N = 19)	LPR symptoms (3/w)	SE, VV, EH, VE,	RFS	Gr1 and 2: t1 > t0	12 w	Gr1: Pantoprazole (40 mg/d)
	Placebo–RCT	Gr2: LPR (N = 20)	Triple-probe pH metry	LE, PH, GG, TM		Gr1 = Gr2		Gr2: Placebo
								Diet: –
Reichel, 2008 (23)	Monocentric	Gr1: suspected LPR (N = 30)	RSI > 13 and RFS > 7	SE, VV, EH, VE,	RFS 6 w (Gr1 and 2)	t1 > t0; t2 > t0	6, 12 w	Gr1: Esomeprazole (20 mg, 2/d)
	Placebo–RCT	Gr2: suspected LPR (N = 28)		LE, PH, GG, TM	Intergroup (6 w)	Gr1 = Gr2		Gr 2: Placebo
					RFS 12 w (Gr1 and 2)	t1 > t0; t2 > t0		Diet: –
					Intergroup (12 w)	Gr1 > Gr2		
Vashani, 2010 (24)	Monocentric	Gr1: suspected LPR (N = 16)	LPR symptoms and signs	VE, LE, EH, PW,	Vocal folds erythema and edema	t1 > t0, Gr2 > Gr1	6 w	Gr1: voice therapy (2/w)+
	Placebo–RCT	Gr2: suspected LPR (N = 16)	Hoarseness	PO	Pharyngeal erythema and edema	t1 > t0, Gr2 > Gr1		Omeprazole (20 mg, 2/d)
			Esophagitis		Hypopharyngeal erythema	t1 > t0, Gr2 > Gr1		Gr 2: Placebo (2/d)
					Hypopharyngeal edema	t1 > t0, Gr2 > Gr1		Diet: –

(Continues)

TABLE II.
Continued

References	Design	Characteristics	Inclusion Criteria	Signs	Sign Outcomes	Results	ET	Treatment
Vashani, 2010 (24)	Monocentric	Gr1: suspected LPR (N = 16)	LPR symptoms and signs	VE, LE, EH, PW,	Vocal folds erythema and edema	t1 > t0, Gr2 > Gr1	6 w	Gr1: voice therapy (2/w)+
	Placebo–RCT	Gr2: suspected LPR (N = 16)	Hoarseness	PO	Pharyngeal erythema and edema	t1 > t0, Gr2 > Gr1		Omeprazole (20 mg, 2/d)
			Esophagitis		Hypopharyngeal erythema	t1 > t0, Gr2 > Gr1		Gr 2: Placebo (2/d)
					Hypopharyngeal edema	t1 > t0, Gr2 > Gr1		Diet: –
McGlashan, 2009 (25)	Monocentric	Gr1: suspected LPR (N = 24)	RSI > 10 and RSF > 5	SE, VV, EH, VE,	RFS (8 w)	t1 > t0; Gr1 = Gr2	8, 24 w	Gr1: Gaviscon (4/d)
	Placebo–RCT	Gr2: suspected LPR (N = 25)		LE, PH, GG, TM	RFS (24 w)	t2 > t0; Gr1 = Gr2		Gr2: Placebo
								Diet: +
Fass, 2010 (26)	Monocentric	Gr1: suspected LPR (N = 24)	LPR symptoms and signs	SE, VV, EH, VE,	RFS	Gr1 = Gr2	12 w	Gr1: Esomeprazole (20 mg, 2/d)
	Placebo–RCT	Gr1: suspected LPR (N = 17)		LE, PH, GG, TM				Gr2: Placebo
								Diet: +
Lam, 2010 (27)	Monocentric	Gr1: suspected LPR (N = 42)	LPR symptoms	SE, VV, EH, VE,	RFS (6, 12 and 18 w)	Gr1 = Gr2	6, 12,	Gr1: Rabeprazole (20 mg, 2/d)
	Placebo–RCT	Gr2: suspected LPR (N = 40)	RFS > 7	LE, PH, GG, TM			18 w	Gr2: Placebo
								Diet: +
Ezzat, 2011 (28)	Monocentric	Gr1: suspected LPR (N = 42)	LPR symptoms and signs	SE, VV, EH, VE,	RFS	Gr1 > Gr2	8 w	Gr1: Pantoprazole (40 mg/d) and
	Placebo–RCT	Gr2: suspected LPR (N = 45)		LE, PH, GG, TM				Itopride (50 mg, 3/d)
								Gr2: Pantoprazole (40 mg/d) and
								Placebo
								Diet: +
Tseng,	Monocentric	Gr1: LPR (N = 39)	LPR symptoms and signs	SE, VV, EH, VE,	RFS	t1 > t0; Gr1 = Gr2	8 w	Gr1: Alginate
2018 (29)	Placebo–RCT	Gr2: LPR (N = 40)	RSI > 10 and RFS > 5	LE, PH, GG, TM				Gr2: Placebo
			pH metry/impedance					Diet: +
Siupsinkiene, 2003 (30)	MPC	Gr1: suspected LPR (N = 113)	LPR symptoms and signs	LE, GG, PH	Comp. Signs Score (2w)	Gr1: S; Gr2 and 3: NS	2 and 5 w	Gr1: Omeprazole (20 mg, 1–2/d)
		Gr2: healthy (N = 113)			Comp. Signs Score (5w)	Gr1 and 3: S; Gr2: NS		Gr2: Diet: +
								Gr3: Nothing

(Continues)

TABLE II.
Continued

References	Design	Characteristics	Inclusion Criteria	Signs	Sign Outcomes	Results	ET	Treatment
Park, 2005 (31)	MPC	Gr1: suspected LPR (N = 30)	LPR symptoms and signs	PW, PY, EH, LE,	Comp. PL Score	Prevalence of signs	16 w	Gr1: Lansoprazole (30 mg, 2/d)
		Gr2: suspected (N = 30)		PH, VR, VV, VE,		No statistical evolu		Gr2: Omeprazole (20 mg, 2/d)+
		Gr3: suspected (N = 25)		SP, SR, GG		tion provided		Ranitidine (300 mg/d)
								Gr3: Esomeprazole (40 mg, 1/d)
								Diet: +
Swoger, 2006 (32)	MPC	Gr1: Uncured LPR (N = 10)	LPR symptoms and signs	PW, PY, EH, LE,	Comp. PL Score	Gr1 = Gr2	12, 52 w	Gr1: Fundoplication
		Gr2: Uncured LPR (N = 15)	pH impedance/monitoring	PH, VR, VV, VE,				Gr2: Omeprazole (80 mg/d)
			Esophagoduodenoscopy	SP, SR, GG				or Lansoprazole (120 mg/d)
			Resistance to PPIs treatment					Diet: +
Chung, 2012 (33)	MPC	Gr1: suspected LPR (N = 22)	RSI > 13	SE, VV, EH, VE,	Gr1: RFS (4 w)	t0 = t1, t1 > t0	4, 8 w	Gr1: Lansoprazole (30 mg 1/d)
		Gr2: suspected LPR (N = 20)		LE, PH, GG, TM	Gr1: RFS (8 w)	t1 > t0		Gr2: Lansoprazole + SGB
					Gr2: RFS (4 and 8 w)	t1 > t0		Diet: NA
					Gr Comparison	Gr2 > Gr1 (4 and 8 w)		
Oridate, 2012 (34)	MPC	Gr1: suspected LPR (N = 60)	LPR symptoms	LE, PH, SE, GG,	Modified RFS (t0, t1)	Gr1 = Gr2	4 w	Gr 1: Rabeprazole (10 mg/d)
		Gr2: suspected LPR (N = 13)	Since at least 1 m	TM, PI, EH				Gr 2: No treatment
								Diet. NA
Chun, 2013 (35)	MPC	Gr1: suspected LPR (N = 32)	RSI > 13 and RFS > 7	SE, VV, EH, VE,	Gr1 and 2 (RFS)	t1,2 > t0	6, 12 w	Gr1: Lansoprazole (30 mg/d)
		Gr2: suspected LPR (N = 29)		LE, PH, GG, TM	Gr differences (6 and 8 w, RFS)	Gr1 = Gr2		Gr2: Lansoprazole and
								Itopride (50 mg, 3/d)
								Diet: NA
Chappity, 2014 (36)	Monocentric	Gr1: suspected LPR (N = 117)	RSI > 13	DT, EH, LE, NC,	Comp. Signs Score 8	Gr1 = Gr2	12 w	Gr1: Omeprazole (20 mg, 2/d)
	RCT	Gr2: suspected LPR (N = 117)		PY, PW, VR, PH				Gr2: diet
								Diet: +
	MPC		Gr1: RSI > 13 and RFS > 7		RFS	t1 > t0; Gr1 = Gr2	4 w	

(Continues)

TABLE II.
Continued

References	Design	Characteristics	Inclusion Criteria	Signs	Sign Outcomes	Results	ET	Treatment
Wan, 2014 (37)		Gr1: suspected LPR (N = 35)		SE, VV, EH, VE,				Esomeprazole (20 mg, 2/d)
		Gr2: LPR (N = 23)	Gr2: Dual-probe pH/impedance	LE, PH, GG, TM				Diet: +
		Gr3: CT (N = 58)	Metry					
Ozturan, 2016 (38)	MPC	Gr1: suspected LPR (N = 65)	LPR symptoms	SE, VV, EH, VE,	RFS	t1, t2 > t0	4, 8 w	Esomeprazole (20 mg, 2/d)
		Gr2: Control (N = 35)		LE, PH, GG, TM,	t2	Gr1 = Gr2		Diet: +
				AN, UV, PO,				
Hanson, 1995 (39)	MPUC	Suspected LPR (N = 141)	LPR symptoms and signs	EH, LE, PH, PW,	Hypopharyngeal and	t1 > t0*	4 w	Omeprazole (20 mg, 1/d)
				GG, TM, KT	Laryngeal signs			Diet: +
Shaw, 1997 (40)	MPUC	Suspected LPR (N = 96)	LPR symptoms and signs	EH, LE, UC, GG	Comp. Signs Score 1	t1 > t0	12 w	Omeprazole (20 mg/d)
					Except granulation			Diet: +
Habermann, 1999 (41)	MPUC	Suspected LPR (N = 29)	LPR symptoms and signs	SP, SR, PH, VR,	Comp. Signs Score 2	t1 > t0	6 w	Pantoprazole (40 mg/d)
			Voice disorder	VE, GG, PP, UC,				Diet: –
				LE, EH, LO, TI				
Belfasky, 2001 (42)	MPUC	LPR (N = 39)	LPR symptoms and signs	SE, VV, EH, VE,	RFS	t1, t2, t3 > t0	8, 16, 24 w	Omeprazole (20 mg, 2/d) or
			Dual-probe pH metry	LE, PH, GG, TM				Rabeprazole (20 mg, 2/d) or
								Lansoprazole (30 mg, 2/d)
								Diet: +
Rodriguez, 2002 (43)	MPUC	Suspected LPR (N = 21)	LPR symptoms and signs	TM, LH, EH, UC,	Comp. Signs Score	t1 > t0#	12 w	Omeprazole (20 mg, 2/d)
				GG	#except granulation score			Diet: –
Habermann, 2002 (44)	MPUC	Suspected LPR (N = 24)	LPR symptoms and signs	VV, EH, VE, LE,	Comp. Signs Score 4	t1 > t0	6 w	Pantoprazole (40 mg/d)
				PH, GG, UC, VR,				Diet: NA
				SP, SR, PP				
DelGaudio, 2003 (45)	MPUC	Gr1: LPR responder (N = 19)	LPR symptoms and signs	VE, VR, LE, GG,	Comp. Signs Score 5		4, 8 w	Esomeprazole (40 mg 1/d)
		Gr2: nonresponder (N = 11)	Responder (8 w therapy)	PH, TT, TM, EH	Responder	t1 > t0; t2 > t0		Diet: +
			Nonresponder (pH metry+)		Nonresponder	t1 > t0; t2 > t0		

(Continues)

TABLE II.
Continued

References	Design	Characteristics	Inclusion Criteria	Signs	Sign Outcomes	Results	ET	Treatment
Bilgen, 2003 (46)	MPUC	Gr1: suspected LPR (N = 36)	LPR symptoms and signs	SE, VV, EH, VE,	RFS	t3 > t1 > t0	12, 16, 24 w	Lansoprazole (30 mg, 2/d)
		Gr2: CT (N = 23)		LE, PH, GG, TM				Diet: +
Garrigues, 2003 (47)	MPUC	Suspected LPR (N = 91)	LPR symptoms and signs	GG, UC, LO, TM,	Comp. Signs Score 6	t2 > t1 > t0*	12 and 24 w	Omeprazole (20 mg, 2/d)
				EH, VR, SR				Diet: NA
Beaver, 2003 (9)	MPUC	Suspected LPR (N = 49)	LPR symptoms and signs	PH, SP, SE, VR,	LPR disease index (photos)	t1 > t0	6 w	Lansoprazole (30 mg, 2/d)
				SR, SU, ND, PP,				Omeprazole (20 mg, 2/d)
				LL, GG, WW				Pantoprazole (40 mg, 2/d)
								Rabeprazole (20 mg, 2/d)
								Diet: NA
Williams, 2004 (48)	MPUC	Suspected LPR (N = 20)	LPR symptoms and signs	LE, EH, VE, VR	Laryngoscopic Grading Score	t1 > t0	12 w	Omeprazole (20 mg, 3/d)
			Since at least 3 m	SE, SU, UC				Diet: +
Sereg-Bahar, 2005 (49)	MPUC	Suspected LPR (N = 43)	LPR symptoms and signs	SE, VV, EH, VE,	RFS	t1 > t0	8 w	Esomeprazole (40 mg/d)
				LE, PH, GG, TM				Diet: +
Qaader, 2005 (10)	MPUC	Gr1: LPR (N = 72)	LPR symptoms and signs	PY, PW, GG, EH,	Comp. Signs Score v1	t1 > t0; t2 > t1	16 w, 54 w	Gr1: Omeprazole (40 mg, 2/d)
		Gr2: LPR (N = 10)	pH metry	PH, KT, LE, VE,				or Lansoprazole (60 mg, 2/d)
			GERD	VR, PP, SP, SR				Gr2: Fundoplication
			Resistance to PPIs treatment					Diet: NA
Qua, 2007 (50)	MPUC	Suspected LPR (N = 32)	LPR symptoms and signs	LE, EH, VE, VR	Laryngoscopic Grading Score	Gr1 > Gr2	8 w	Lansoprazole (30 mg, 2/d)
		Gr1: GERD (N = 21)		SE, SU, UC	Gr1–Gr2	t1 > t0; t1 = t0		Diet: –
		Gr2: non-GERD (N = 11)						
Jin, 2008 (51)	MuPUC	LPR (N = 40)	LPR symptoms and signs	SE, VV, EH, VE,	RFS	t2,3,4,5 > t0	2, 4, 8, 12,	Lansoprazole (30 mg/d)+
			Dual-probe pH metry	LE, PH, GG, TM			16, 20 w	Levosulpiride (25 mg, 3/d) or
								Mosapride (5 mg, 3/d)
								Diet: NA

(Continues)

TABLE II.
Continued

References	Design	Characteristics	Inclusion Criteria	Signs	Sign Outcomes	Results	ET	Treatment
Koufman,	MPUC	Resistant LPR (N = 20)	Resistance to PPI+Anti-H2	SE, VV, EH, VE,	RFS	t1 > t0	2 w	1. Strict diet
2011 (52)			Dual-probe pH metry	LE, PH, GG, TM				
Lee, 2011 (4)	MuPUC	Suspected LPR (N = 455)	LPR symptoms and signs	SE, VV, EH, VE,	RFS	t1 > t0	12 w	Rabeprazole (10/20 mg/d)
				LE, PH, GG, TM				Diet: +
Masaany, 2011 (53)	MPUC	Suspected LPR (N = 47)	RSI > 13 and RFS > 7	SE, VV, EH, VE,	RFS	t3 > t0; t2 > t0; t1 > t0	8, 12, 16 w	Pantoprazole (40 mg, 2/d)
				LE, PH, GG, TM				Diet: NA
Naiboglu, 2011 (54)	MPUC	Suspected LPR (N = 50)	Esophagitis	SE, VV, EH, VE,	RFS	t1 > t0	12 w	Lansoprazole (30 mg/d)
			RS1 > 13 and RFS > 7	LE, PH, GG, TM				Diet: +
Patigaroo, 2011 (55)	MPUC	Suspected LPR (N = 50)	RSI > 13 and RFS > 7	SE, VV, EH, VE,	RFS (8, 16 w)	t1 = t0, t2 > t1	8, 16 w	(Es)/omeprazole (20 mg, 2/d)
				LE, PH, GG, TM				Lansoprazole (30 mg, 2/d)
								Pantoprazole (40 mg, 2/d)
								Diet: NA
Habermann, 2012 (56)	MuPUC	Suspected LPR (N = 1044)	RSI > 9 and RFS > 7	SE, VV, EH, VE,	RFS	t1 > t0	8–12 w	Pantoprazole (20 or 40 mg, 2/d)
				LE, PH, GG, TM				Diet: NA
Park, 2012 (57)	MPUC	Gr1: suspected LPR (N = 50)	RSI > 13 and RFS > 7	SE, VV, EH, VE,	RFS (4 and 8 w)	Gr2 = Gr1	4, 8, 12 w	Gr1: Omeprazole (20 mg, 2/d)
		Gr2: suspected LPR (N = 50)		LE, PH, GG, TM	RFS (4, 8, 12 w)	Gr2 > Gr1		Gr2: Omeprazole +
								Voice therapy
								Diet: –
Vailati, 2013 (74)	MPUC	Suspected LPR (N = 22)	RSI > 13	SE, VV, EH, VE,	RFS	t1 > t0	12 w	Pantoprazole (40 mg, 2/d)
				LE, PH, GG, TM				Diet: NA
Lee, 2014 (58)	MuPUC	Suspected LPR (N = 180)	LPR symptoms and signs	SE, VV, EH, VE,	RFS	t1 > t0	12 w	Lansoprazole (15 mg, 2/d)
				LE, PH, GG, TM				Diet: +

(Continues)

TABLE II.
Continued

References	Design	Characteristics	Inclusion Criteria	Signs	Sign Outcomes	Results	ET	Treatment
Semmanaselvan, 2015 (59)	MPUC	Suspected LPR (N = 50)	RSI > 13 and RFS > 7	SE, VV, EH, VE, LE, PH, GG, TM	RFS	t1 > t0	12 w	1. Rabeprazole (20 mg/d) 2. Domperidone (30 mg/d) Diet: NA
Batioglu, 2016 (60)	MPUC	Suspected LPR (N = 84)	RSI > 13 and RFS > 7	SE, VV, EH, VE, LE, PH, GG, TM	RFS	t1 > t0	12 w	Lansoprazole (30 mg, 2/d) Diet: NA
Dulery, 2016 (61)	MPUC	Suspected LPR (N = 24)	LPR symptoms	SE, VV, EH, VE, LE, PH, GG, TM	RFS	t0 = t1	8 w	Esomeprazole (40 mg, 2/d) Diet: NA
Joshi, 2017 (62)	MPUC	Suspected LPR (N = 100)	LPR symptoms LPR signs: RFS > 7	SE, VV, EH, VE, LE, PH, GG, TM TT	RFS	t3 > t2 > t0	4, 8, 24 w	Omeprazole (20 mg, 2/d) Diet: +
Pullarat, 2017 (63)	MPUC	Suspected LPR (N = 30)	LPR symptoms Hoarseness since > 3 w	SE, VV, EH, VE, LE, PH, GG, TM TT	RFS	t1 > t0	8 w	Pantoprazole (40 mg/d) Diet: NA
Lechien, 2018 (5)	MuPUC	LPR (N = 80)	RSI > 13 and RFS > 7 pH metry (resistant patients)	SE, VV, EH, VE, LE, PH, GG, TM TT	RFS	t1 > t0	12 w	Pantoprazole (20 mg, 2/d) Diet: +
Gupta, 2016 (11)	Monocentric Retrospective	Suspected LPR (N = 188)	LPR symptoms LPR signs	SE, VV, EH, VE, LE, PH, GG, TM TT	RFS	t1 > t0	10 w	PPIs (2/d) Diet: NA

AN = anterior pillars erythema/edema; CPLI = chronic posterior laryngitis index; d = day; ET = evaluation time; EH = laryngeal/arytenoids erythema; GG = interarytenoid granulation and/or granuloma; Gr = group; KT = laryngeal keratosis; LE = laryngeal edema; LGS = laryngoscopic grading scale; LL = leukoplakia; LO = loss light reflect; LPR = laryngopharyngeal reflux; LRDI = laryngopharyngeal reflux disease index; LRG = laryngeal reflux grade; MPC = monocentric prospective controlled study; M(Mu)PUC = monocentric (multicentric) prospective uncontrolled study; NC = nasal congestion; ND = nodules; PH = posterior commissure hypertrophy; PI = mucous pooling in the pyriform sinus; PO = posterior oropharyngeal wall erythema; PP = polyp/Reinke edema; PW = posterior pharyngeal wall erythema; PY = postpharyngeal cobblestoning; RCT = randomized controlled trial; RFS = reflux finding score; RSI = reflux symptom index; SE = subglottic edema/pseudosulcus/stenosis; SP = supraglottis edema; SR = supraglottis erythema; SU = subglottic erythema; TM = thick endolaryngeal mucus; t = time; TT = tongue tonsil hypertrophy; UC = laryngeal ulcerations; UV = uvula erythema/edema; VE = vocal fold edema; VR = vocal fold erythema; VV = ventricular obliteration; w = week; WW = vocal web.

TABLE III.
Objectives, Targeted Population, Population Characteristics, Setting, and Instrument Characteristics for LPRD Sign Validated Instruments

Instrument	Year	Language/ Country	Validation/ Standard	Objective	Target and Patient Characteristics	Setting	Scale Characteristics					
							Signs	Type	Item (N)	Item Response	Calculation	Subscales
RFS (42)	2001	U.S.	+	Diagnosis	Suspected LPR	Monocenter	SE, VV, EH, VE,	PRI	8	Severity: 0–4	Sum of items	0
				Therapeutic	LPR	Tertiary center	LE, PH, GR, TM			or 0–2	Total score: 26	
				outcome	Age: 52 y	Controlled						
					Gender: 29 F/11 M	Prospective						
Vaezi	2002	U.S.	+	Diagnosis	Uncured LPR	Monocenter	PY, PW, GG, EH,	Yes/ No	12	Presence: yes/no	Sign prevalence	0
Instrument (64)				Therapeutic	Gender: 18 F, 7 M	Sec/ter center	PH, KT, LE, VE,				Total score: NA	
				outcome	Age: 39.9 y		VR, PP, SP, SR					
LRDI (9)	2003	U.S.	+	Therapeutic	Suspected LPR	Monocenter	PH, SP, SE, VR,	VAS	12	Severity: 0–3	Sum of items	0
				outcome	Gender: NA	Tertiary center	SR, SU, ND, PP,				Total score: 36	
					Age: NA		LL, GG, WW					
LGS (48)	2004	Australia	+	Therapeutic	Suspected LPR	Monocenter	LE, EH, VE, VR	PRI	4	Laryngitis grade: 0–4	–	0
				outcome	Gender: 11 F, 9 M	Sec/ter center	SE, SU, UC			Each grade is defined		
					Age: 55 y							
LRG (20)	2004	U.S.	+	Therapeutic	LPR	Monocenter	EH, VE, LE, PH,	Likert	Signs: 6	Severity: 0–4	Sum of items	Sign scale
				outcome	Gender: 30 F, 12 M	Sec/ter center	VR, GG, ND, UC,	Scale	VC wave:		Total score: 24 + 16	VC wave
					Age: 49.3 y	Placebo–RCT	SE		4			
CPLI (21)	2006	U.S.	+	Therapeutic	Suspected LPR	Monocenter	EH, GG, LE, PW	VAS	10	Severity: 0–3	Sum of items	0
				outcome	Gender: 74 F, 71 M	Sec/ter center	PH, VR, VE				Total score: 30	
					Age: 51 y	Placebo–RCT						

+ = criterion met; AN = anterior pillars erythema/edema; CPLI = chronic posterior laryngitis index; EH = Laryngeal/arytenoids erythema; F = female; GG = interarytenoid granulation and/or granuloma; KT = laryngeal keratosis; LE = laryngeal edema; LGS = laryngoscopic grading scale; LL = leukoplakia; LO = loss light reflect; LPR = laryngopharyngeal reflux; LRDI = laryngopharyngeal reflux disease index; LRG = laryngeal reflux grade; M = male; NA = not available; NC = nasal congestion; ND = nodules; N.E. = not evaluable; PH = posterior commissure hypertrophy; PI = mucous pooling in the pyriform sinus; PO = posterior oropharyngeal wall erythema; PP = polyp/Reinke edema; PRI = predefined item; PW = posterior pharyngeal wall erythema; PY = postpharyngeal cobblestoning; RCT = randomized controlled trial; RFS = reflux finding score; sec = secondary; SE = subglottic edema/pseudosulcus/stenosis; SP = supraglottis edema; SR = supraglottis erythema; SU = subglottic erythema; TM = thick endolaryngeal mucus; TT = tongue tonsil hypertrophy; UC = laryngeal ulcerations; UV = uvula erythema/edema; VAS = visual analog scale; VC = vocal cords; VE = vocal fold edema; VR = vocal fold erythema; VV = ventricular obliteration; WW = vocal web; y = year.

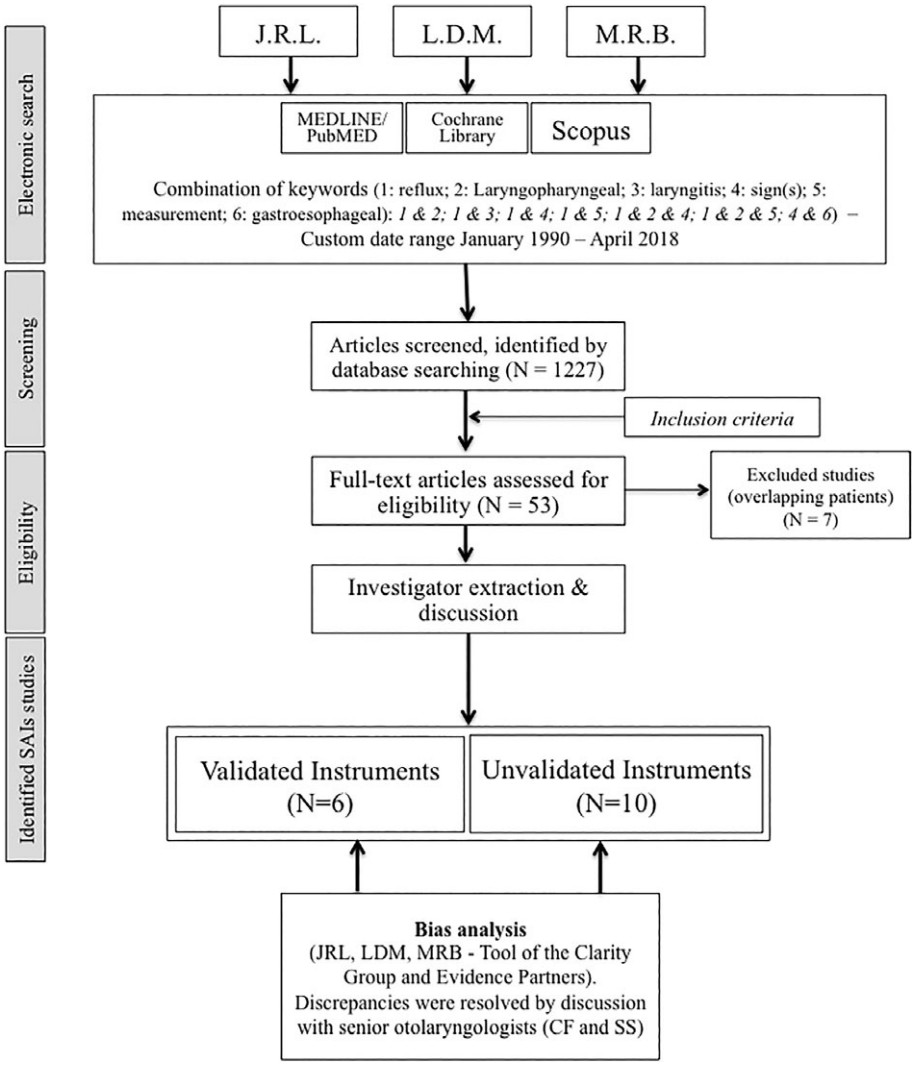

Fig. 1. Flow chart shows the process of article selection for this study.

Analysis of Validated ICFs

RFS[42] is the most common validated ICF in the literature (N = 31 of 53), followed by the composite score developed by the group of Vaezi (N = 3 of 53), which we refer to as the *Vaezi instrument.* Notably, this instrument was validated over four studies.[10,31,32,64] The remaining validated ICFs were used only in their validation study. We did not identify other authors who used these ICFs in their studies. Thus, these ICFs are limited in scope. The measurement properties of the six validated ICFs are described in Table V. Overall, no ICF met all the validated criteria described in Table I.

CONCEPTUAL MODEL AND CONTENT VALIDITY. The conceptual model was adequately described in all references describing ICF validation. Descriptions of the item development and the appropriation of all the LPRD instruments were provided in all the publications. Concerning instrument development, no instrument was based on a clinical study that assessed the prior prevalence of all ENT findings among patients with LPRD. The content of all ICFs was based on the opinions of experts (i.e., otolaryngologists, gastroenterologists, and/or speech therapists). However, some laryngeal and many extralaryngeal signs were not considered in the elaboration of all ICFs, particularly the LGS, CPLI, and RFS.

RELIABILITY. The reliabilities of four instruments were assessed via test–retest reliability. Of these instruments, the RFS and LRG exhibited consistent overall test–retest reliability. Concerning individual items, RFS showed initially consistent reliability and LRG had moderate-to-high reliability. However, additional studies did not confirm the high reliability of the RFS.[71] The reliabilities of the LRDI and Vaezi instruments were lower. Test–retest reliability analyses were not reported for either LGS or CPLI. Based on interrater reliability (i.e., concordance), the RFS and LGS exhibited higher concordance than did the LRDI and Vaezi instrument. No instrument was tested for internal consistency.

CONSTRUCT VALIDITY. Known-groups validity was statistically demonstrated with regard to the RFS,

TABLE IV.
Assessed Signs in Validated, Unvalidated Instruments, and in the Literature

Signs in Instruments (N = 16)		N	Signs in Number of Studies: N = 53		
Laryngopharyngeal			**Laryngopharyngeal**		
Laryngeal/arytenoids erythema	EH	15	Laryngeal/arytenoids erythema	EH	51
Laryngeal edema	LE	14	Granuloma/granulation (interarytenoid nodularity)	GG	49
Posterior commissure hypertrophy	PH	11	Laryngeal edema	LE	45
Vocal fold edema	VE	10	Posterior commissure hypertrophy	PH	44
Thick endolaryngeal mucous	TM	7	Vocal fold edema	VE	38
Subglottic edema/pseudosulcus/stenosis	SE	6	Thick endolaryngeal mucous	TM	37
Granuloma/granulation (interarytenoid nodularity)	GG	12	Subglottic edema/pseudosulcus/stenosis	SE	49
Ventricular obliteration	VV	2	Ventricular obliteration	VV	34
Vocal fold erythema	VR	11	Vocal fold erythema	VR	14
Vocal cord epithelium thickening	TI	2	Laryngeal ulcerations	UC	1
Laryngeal ulcerations	UC	5	Supraglottis erythema	SR	9
Posterior pharyngeal wall erythema	PW	3	Posterior pharyngeal wall erythema	PW	7
Supraglottis edema	SP	4	Supraglottis edema	SP	7
Supraglottis erythema	SR	6	Polyp/Reinke edema	PP	8
Leukoplakia	LL	1	Postpharyngeal cobblestoning	PY	1
Polyp/reinke edema	PP	4	Subglottic erythema	SU	5
Postpharyngeal cobblestoning	PY	3	Loss light reflect	LO	5
Loss light reflect	LO	2	Nodules	ND	2
Nodules	ND	2	Leukoplakia	LL	2
Vocal web	WW	1	Vocal cord epithelium thickening	TI	1
Laryngeal keratosis	KT	1	Vocal web	WW	2
Mucous pooling in the pyriform sinus	PI	1	Laryngeal keratosis	KT	1
Subglottic erythema	SU	2	Mucous pooling in the pyriform sinus	PI	3
Extralaryngopharyngeal			**Extralaryngopharyngeal**		
Tongue tonsil hypertrophy	TT	1	Tongue tonsil hypertrophy	TT	5
Nasal congestion	NC	1	Posterior oropharyngeal wall erythema	PO	1
Uvula erythema/edema	UV	0	Nasal congestion	NC	1
Anterior pillars erythema/edema	AN	0	Uvula erythema/edema	UV	1
Posterior oropharyngeal wall erythema	PO	0	Anterior pillars erythema/edema	AN	2
Dull tympanic membrane	DT	1	Dull tympanic membrane	DT	1

AN = anterior pillars erythema/edema; EH = laryngeal/arytenoids erythema; GG = interarytenoid granulation and/or granuloma; KT = laryngeal keratosis; LE = laryngeal edema; LL = leukoplakia; LO = loss light reflect; NC = nasal congestion; ND = nodules; PH = posterior commissure hypertrophy; PI = mucous pooling in the pyriform sinus; PO = posterior oropharyngeal wall erythema; PP = polyp/Reinke edema; PW = posterior pharyngeal wall erythema; PY = postpharyngeal cobblestoning; SE = subglottic edema/pseudosulcus/stenosis; SP = supraglottis edema; SR = supraglottis erythema; SU = subglottic erythema; TM = thick endolaryngeal mucus; TT = tongue tonsil hypertrophy; UC = laryngeal ulcerations; UV = uvula erythema/edema; VE = vocal fold edema; VR = vocal fold erythema; VV = ventricular obliteration; WW = vocal web.

LRDI, and Vaezi instruments. Other instruments did not provide comparisons between LPR patients and controls. Responsiveness to change, especially as assessed throughout the therapeutic course, was satisfactory with regard to all ICFs. In the current literature, the RFS significantly improved over treatment according to 25 studies (Table II). Concerning the Vaezi instrument, only Qaader et al.[10] found substantial improvement of the score throughout treatment. No instrument was tested for convergent validity.

INTERPRETABILITY AND SCORING. All the publications provided scoring (i.e., calculation) details. The scoring system substantially varied across instruments. The majority of the instruments had a total score that corresponded to the sum of the item scores where higher total scores indicated more severe LPRD signs. Thus, the addition of each item score from the RFS, LRDI, CPLI, and LRG led to a total score ranging from a minimum of 0 to a maximum of 4 to 40 depending on the instrument. The Vaezi instrument does not have a total score because it is based on the presence or absence of signs. In addition, the Vaezi instrument was not constructed to sum all item scores. Regarding interpretability, only the RFS determined a diagnosis cutoff (RFS = 7) for an abnormal score that could be combined with a reflux symptom index > 13 to represent a high probability of LPRD. With regard to the other ICFs, we did not find cutoff or severity thresholds with clinical significance.

Bias Analysis

Many studies included patients suspected as having LPR without a formal diagnosis. The heterogeneity of the

TABLE V.
Instrument Analysis

Instrument	Studied Population	Construct Definition	Content Validity	Internal Consistency	Test–Retest Reliability	Concordance	Convergent Validity	Known-Groups Validity	Responsiveness to Change	Interpretability and Scoring
RFS	LPR patients	+	±	NP	Total: 0.95	0.90	NP	+	+	+
					Item: > 0.90	2 physicians				
LRG	LPR patients	+	+	NP	Total: 0.83–0.90	0.43	NP	NP	+	±
						3 physicians				
CPLI	LPR patients	+	±	NP	NP	NP	NP	NP	+	±
Vaezi Instrument	LPR patients	+	+	NP	0.26–0.78	0.32–0.58	NP	+	+	NA
						1 physician 1 speech th.				
LRDI	Suspected LPR	+	+	NP	0.42–0.78	0.30	NP	+	+	±
						3 physicians				
LGS	Suspected LPR	+	±	NP	NP	0.75–0.93 3 physicians	NP	NP	+	±

CPLI = chronic posterior laryngitis index; LGS = laryngoscopic grading scale; LPR = laryngopharyngeal reflux; LRDI = laryngopharyngeal reflux disease index; LRG = laryngeal reflux grade; NA = not available; NP = not provided (the analysis was not made); RFS = reflux finding score; + = the conducted analysis was completely consistent with our definition; ± = the conducted analysis was partly consistent with our definition; – = the analysis was inconsistent.

inclusion criteria across studies can lead to differences in the profiles of the patients with LPR who comprised the validation studies. This point is consistent with a selection bias.

In a large majority of the studies, physicians were not blinded to the description of LPRD signs based on symptoms. Moreover, one study[21] proposed specialized training to improve/standardize the evaluation of signs among physicians before they could be assessed with an ICF. The training of judges can bias certain measurement properties such as concordance (i.e., interrater reliability). These two points represent the possibility of an evaluation/detection bias.

In addition, we identified various therapeutic regimens including the use of PPIs in association with other drugs (i.e., alginate, antihistamine, or gastroprokinetic) or diet. Because the changes in signs over treatment depend on the therapy, the therapeutic variability and the lack of a treatment demonstrated as superior to placebo might bias the assessment of responsiveness to change. The lack of inclusion of many signs related to LPRD in ICFs might also bias an overall patient response to treatment that is consistent with an evaluation bias. Supporting Table SII). provides risk of bias assessment according to studies.

DISCUSSION

The use of patient-reported outcome measures, ICFs, or both has become standard in studies assessing the efficiency of a treatment.[65] Measurement instruments are useful for collecting precise data for the initial evaluation of patient symptoms and signs and the assessment of treatment effectiveness. This point is particularly relevant with regard to LPRD and the low specificity of its signs and symptoms given the existing controversies regarding the superiority of PPIs over placebo.[66]

In this systematic review, we identified 16 ICFs, which included 10 unvalidated tools and six instruments that met at least one developmental measurement property criterion. The usefulness of an outcome measurement instrument is related to its intent and the quality of its development. Of the 6 ICFs that underwent psychometric evaluation, only RFS met the following validation criteria: construct definition, content validity, reliability, concordance, known-group validity, responsiveness to change, and interpretability and scoring.[42] However, the initial version of the RFS was not evaluated for internal consistency or convergent validity. Four other instruments (i.e., the LRG, LRDI, LGS, and Vaezi instrument) provided partial validation of two to six validated conceptual properties. The CPLI does not have documented reliability or construct validity. Of these six ICFs, only RFS and the Vaezi instrument were conceived for both diagnosis and therapeutic outcomes.[42,64]

Input from the targeted subject population is crucial establish an instrument's content validity because it ensures that all relevant signs are included and that the instrument accurately assesses the concepts of interest for physicians. The target population of the included studies consisted of patients suspected[9,48] or confirmed as

having LPR.[20,21,42,64] Our review showed that the diagnosis criteria substantially varied across studies regarding symptoms, signs considered for a LPRD diagnosis, or the diagnostic thresholds used for pH impedance monitoring. Of the studies that applied a clinical diagnosis, the majority did not provide information about the exclusion criteria or did not exclude major cofounding factors such as smoking or alcoholism,[64] infections within the last month, or active allergies,[9,42] which all represent a selection bias. In fact, the inclusion or exclusion criteria of a specific study population can have a dramatic effect on the conclusions regarding the effectiveness of a treatment.[67] Thus, the lack of sensitivity of pH metry/impedance, the inclusion of certain confounds, and the disparity in the LPRD diagnosis methods undeniably affect how current measures define their targeted populations. In addition, the number of patients included in validation studies was low (< 100), although it is recommended to have at least 100 participants to optimize component/factor-analysis–based methods.[68]

Experienced otolaryngologists have examined the content of all validated ICFs with regard to the specific signs that are usually treated in their practice. No study based the content elaboration on the prevalence of LPR signs. Because of the definition of the target population, the content of an ICF is important for its psychometric validation. Our analysis of the signs available in the validated tools reported an important level of heterogeneity among the signs assessed for LPR diagnosis or as a therapeutic outcome. We also observed an overreliance on the same laryngeal signs, especially those described in RFS, and a lack of consideration of other laryngeal (i.e., vocal fold erythema, leukoplakia, and keratosis) and extralaryngeal signs (i.e., tongue tonsil hypertrophy, posterior pharyngeal wall and anterior pillars erythema, edema, coated tongue), although they seemed to concern a considerable number of patients according to previous clinical reports.[7,8,39,45,69] The lack of consideration of these signs might significantly affect the development of an instrument and its usefulness as a diagnostic tool (i.e., patient inclusion) or therapeutic outcome.

Reliability reflects an instrument's stability over time and between different investigators. Our analysis identified high test–retest reliabilities for the RFS[42] and LRG,[20] as well as moderate-to-high reliability for the LRDI[9] and Vaezi instrument.[64] In all studies, the test–retest procedures were consistent with the recommendations of the current literature.[70] The reliability of an instrument also involves a high degree of rater agreement. The initial publications describing the RFS and LGS reported higher interrater agreements, ranging from 0.75 to 0.93.[42,48] However, Chang et al. independently assessed the reliability of RFS among general otolaryngologists and found only a fair level of concordance (correlation coefficient = 0.586) among raters.[71] Belafsky et al. studied two physicians who provided the analysis, came from the same center, and benefited from the same training, explaining the high concordance values compared with other studies.[42] Moreover, a key point that might explain the better concordance for LGS concerns the rating system of their items because it corresponded to a clear definition of an identified sign. In other words, the assessment of the severity of a sign using a VAS is more subject to interrater differences because it is more subjective and related to the knowledge and experience of the physician[71] than an instrument in which each rating is well-defined. However, the concordances of the LRDI[9] and Vaezi instrument[64] based on their sign evaluation on a VAS were lower than those of the RFS[42] and LGS,[48] even though the authors of these two studies followed the usual recommended procedures. In addition, concordance seems to depend on the signs evaluated. As such, Beaver et al. showed that certain signs, especially leukoplakia, nodules or prenodules, and contact granulomas, had the highest levels of agreement compared with edema and erythema of laryngeal spaces.[9] Thus, evaluating the interrater reliability is important for the total score and each item.

The validity of an instrument depends on its ability to detect differences between independent groups using between-participant statistics, which corresponds to known-groups validity. The LRDI, RFS and Vaezi instrument provided consistent results in their respective case-control studies comparing healthy participants and patients.[21,42,64] Nevertheless, the numbers of healthy participants included in these case-control studies were low, and the selection criteria for these participants were rarely described. Concerning the interpretability of a score, we found critical thresholds for RFS only because Belafsky et al. determined that a combination of RFS > 7 and RFS > 13 is highly correlated with an LPRD diagnosis.[42,72]

Finally, the construct validity of an instrument involves its ability to measure responsiveness to change that strongly depends on the content, reliability, and validity of the outcome measurement instrument. Hence, the lack of consideration of certain typical findings related to LPRD might negatively affect responsiveness to change. In fact, we and other authors[66] suspect that the controversial results of placebo-randomized controlled trials are partially due to the lack of complete, reliable tools evaluating both the signs and symptoms of LPRD. This point concerns all the ICFs described in the present review, including RFS. Thus, RFS does not include extralaryngeal signs or laryngeal signs (i.e., leukoplakia, vocal fold erythema, subglottic erythema, and thickening of the vocal fold epithelium). However, as the first and most well-known clinical instrument regarding LPRD, RFS remains commonly used around the world. Moreover, patients with leukoplakia and granuloma were excluded during the development of RFS,[42] although they might be related to LPR.[8] Responsiveness to change is based on a physician's perception of signs, which involves an unbiased evaluation of clinical findings. Our analysis revealed that many physicians assessed laryngoscopic signs with knowledge of the patient's symptoms.[20,21,42,64] Only Beaver et al. performed blinded evaluations via clinical trials involving a laryngoscopic examination.[9] The unblinded assessment of laryngoscopic signs increases the risk of misjudging the sign score because LPR studies have demonstrated that the perception of a sign can be influenced by the physician's personal knowledge, its understanding

within the clinical area, and knowledge of the patient's complaints.[71,73] For these reasons, the assessment of signs must be blinded according to both the patient's state (LPR vs. healthy) and their complaints to provide reliable instrument development. In clinical practice, with regard to the subjectivity in the assessment of signs, the poor specificity of signs, and low interjudge reliability, future ICFs could comprise scores with precise descriptions of the item grade of each sign. In other words, future ICFs could avoid the grading of signs with a VAS that remains objective and related to physician abilities. The development of software that can objectify the degree of inflammation of the mucosa on the basis of redness intensity could also aid physicians in the clinical evaluation.

Finally, no ICFs provided internal consistency or convergent analyses. Regarding the limitations of this review, first, we limited our research to publications in the English and French languages and excluded unpublished ICFs. Second, judgments of instrument characteristics were based on the analysis of three independent physicians. Even if this procedure minimizes the risk of the biases related to the subjectivity of the task, these biases remain possible. Third, the low number of patients in all included studies and their methodological biases might have affected our psychometric analysis. Four, many authors selected their LPR patients on the basis of symptoms and signs without additional examination. Potentially, these cohorts of suspected LPR patients included subjects without LPRD, which can also negatively impact the treatment efficiency and related responsiveness to change.

CONCLUSION

With patient-reported questionnaires, ICFs have an important place in LPRD diagnosis and evaluation of treatment effectiveness. To date, only a few ICFs have been developed, and none currently meet all clinical and regulatory requirements. The main impairments related to their psychometric validation include variable construct validity, disparate and uncertain reliabilities, and lack of interpretability. Moreover, no current instrument takes into account a full assessment of laryngeal and extralaryngeal signs related to LPRD, hindering content and construct validities. The development of further complete, reliable ICFs is particularly important with regard to the low specificity of both signs and symptoms related to LPRD and the lack of a gold standard for the diagnosis.

BIBLIOGRAPHY

1. Koufman JA, Aviv JE, Casiano RR, Shaw GY. Laryngopharyngeal reflux: position statement of the committee on speech, voice, and swallowing disorders of the American Academy of Otolaryngology–Head and Neck Surgery. *Otolaryngol Head Neck Surg* 2002;127:32–35.
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(suppl 53):1–78.
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. Lee YS, Choi SH, Son YI, Park YH, Kim SY, Nam SY. Prospective, observational study using rabeprazole in 455 patients with laryngopharyngeal reflux disease. *Eur Arch Otorhinolaryngol* 2011;268:863–869.
5. Lechien JR, Finck C, Khalife M, et al. Change of signs, symptoms and voice quality evaluations throughout a 3 to 6-months empirical treatment for laryngopharyngeal reflux disease. *Clin Otolaryngol* 2018. doi: 10.1111/coa.13140.
6. Ford CN. Evaluation and management of laryngopharyngeal reflux. *JAMA* 2005;294:1534–1540.
7. Lechien JR, Saussez S, Harmegnies B, Finck C, Burns JA. Laryngopharyngeal Reflux and Voice Disorders: A Multifactorial Model of Etiology and Pathophysiology. *J Voice* 2017;31:733–752.
8. Remacle M, Lawson G. Diagnosis and management of laryngopharyngeal reflux disease. *Curr Opin Otolaryngol Head Neck Surg* 2006;14:143–149.
9. Beaver ME, Stasney CR, Weitzel E, et al. Diagnosis of laryngopharyngeal reflux disease with digital imaging. *Otolaryngol Head Neck Surg* 2003; 128:103–108.
10. Qadeer MA, Swoger J, Milstein C, et al. Correlation between symptoms and laryngeal signs in laryngopharyngeal reflux. *Laryngoscope* 2005;115: 1947–1952.
11. Gupta N, Green RW, Megwalu UC. Evaluation of a laryngopharyngeal reflux management protocol. *Am J Otolaryngol* 2016;37:245–250.
12. Francis DO, Patel DA, Sharda R, et al. Patient-reported outcome measures related to laryngopharyngeal reflux: a systematic review of instrument development and validation. *Otolaryngol Head Neck Surg* 2016;155: 923–935.
13. Moher D, Liberati A, Tetzlaff J, Altman DG; The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med* 2009;6:e1000097.
14. Francis DO, McPheeters ML, Noud M, Penson DF, Feurer ID. Checklist to operationalize measurement characteristics of patient-reported outcome measures. *Syst Rev* 2016;5:129.
15. Vakil NB, Halling K, Becher A, Ryden A. Systematic review of patient-reported outcome instruments for gastroesophageal reflux disease symptoms. *Eur J Gastroenterol Hepatol* 2013;25:2–14.
16. Systematic and literature review resources 2011. Available at: http:// distillercer.com/resources. Accessed July 2015.
17. El-Serag HB, Lee P, Buchner A, Inadomi JM, Gavin M, McCarthy DM. Lansoprazole treatment of patients with chronic idiopathic laryngitis: a placebo-controlled trial. *Am J Gastroenterol* 2001;96:979–983.
18. Noordzij JP, Khidr A, Desper E, Meek RB, Reibel JF, Levine PA. Correlation of pH probe-measured laryngopharyngeal reflux with symptoms and signs of reflux laryngitis. *Laryngoscope*. 2002;112:2192–2195.
19. Eherer AJ, Habermann W, Hammer HF, Kiesler K, Friedrich G, Krejs GJ. Effect of pantoprazole on the course of reflux-associated laryngitis: a placebo-controlled double-blind crossover study. *Scand J Gastroenterol* 2003;38:462–467.
20. Steward DL, Wilson KM, Kelly DH, et al. Proton pump inhibitor therapy for chronic laryngo-pharyngitis: a randomized placebo-control trial. *Otolaryngol Head Neck Surg* 2004;131:342–350.
21. Vaezi MF, Richter JE, Stasney CR, et al. Treatment of chronic posterior laryngitis with esomeprazole. *Laryngoscope* 2006;116:254–260.
22. Wo JM, Koopman J, Harrell SP, Parker K, Winstead W, Lentsch E. Double-blind, placebo-controlled trial with single-dose pantoprazole for laryngopharyngeal reflux. *Am J Gastroenterol* 2006;101:1972–1978; quiz 2169.
23. Reichel O, Dressel H, Wiederanders K, Issing WJ. Double-blind, placebo-controlled trial with esomeprazole for symptoms and signs associated with laryngopharyngeal reflux. *Otolaryngol Head Neck Surg* 2008; 139:414–420.
24. Vashani K, Murugesh M, Hattiangadi G, et al. Effectiveness of voice therapy in reflux-related voice disorders. *Dis Esophagus* 2010;23:27–32.
25. McGlashan JA, Johnstone LM, Sykes J, Strugala V, Dettmar PW. The value of a liquid alginate suspension (Gaviscon Advance) in the management of laryngopharyngeal reflux. *Eur Arch Otorhinolaryngol* 2009;266:243–251.
26. Fass R, Noelck N, Willis MR, et al. The effect of esomeprazole 20 mg twice daily on acoustic and perception parameters of the voice in laryngopharyngeal reflux. *Neurogastroenterol Motil* 2010;22:134–141.
27. Lam PK, Ng ML, Cheung TK, et al. Rabeprazole is effective in treating laryngopharyngeal reflux in a randomized placebo-controlled trial. *Clin Gastroenterol Hepatol* 2010;8:770–776.
28. Ezzat WF, Fawaz SA, Fathey H, El Demerdash A. Virtue of adding prokinetics to proton pump inhibitors in the treatment of laryngopharyngeal reflux disease: prospective study. *J Otolaryngol Head Neck Surg* 2011;40: 350–356.
29. Tseng WH, Tseng PH, Wu JF, et al. Double-blind, placebo-controlled study with alginate suspension for laryngopharyngeal reflux disease. *Laryngoscope* 2018. doi: 10.1002/lary.27111.
30. Siupsinskiene N, Adamonis K. Diagnostic test with omeprazole in patients with posterior laryngitis. *Medicina (Kaunas)* 2003;39:47–55.
31. Park W, Hicks DM, Khandwala F, et al Laryngopharyngeal reflux: prospective cohort study evaluating optimal dose of proton-pump inhibitor therapy and pretherapy predictors of response. *Laryngoscope* 2005;115: 1230–1238.
32. Swoger J, Ponsky J, Hicks DM, et al. Surgical fundoplication in laryngopharyngeal reflux unresponsive to aggressive acid suppression: a controlled study. *Clin Gastroenterol Hepatol* 2006;4:433–441.

33. Chung JW, Chun HJ, Lee MS, et al. Effect of stellate ganglion block on laryngopharyngeal reflux disease. *Korean J Anesthesiol* 2013;64:439–442.
34. Oridate N, Tokashiki R, Watanabe Y, Taguchi A, Kawamura O, Fujimoto K. Endoscopic laryngeal findings in Japanese patients with laryngopharyngeal reflux symptoms. *Int J Otolaryngol* 2012;2012:908154.
35. Chun BJ, Lee DS. The effect of itopride combined with lansoprazole in patients with laryngopharyngeal reflux disease. *Eur Arch Otorhinolaryngol* 2013;270:1385–1390.
36. Chappity P, Kumar R, Deka RC, Chokkalingam V, Saraya A, Sikka K. Proton pump inhibitors versus solitary lifestyle modification in management of laryngopharyngeal reflux and evaluating who is at risk: scenario in a developing country. *Clin Med Insights Ear Nose Throat* 2014;7:1–5.
37. Wan Y, Yan Y, Ma F, et al. LPR: how different diagnostic tools shape the outcomes of treatment. *J Voice* 2014;28:362–368.
38. Ozturan O, Dogan R, Yenigun A, Veyseller B, Yildirim YS. Photographic objective alterations for laryngopharyngeal reflux diagnosis. *J Voice* 2017; 31:78–85.
39. Hanson DG, Kamel PL, Kahrilas PJ. Outcomes of antireflux therapy for the treatment of chronic laryngitis. *Ann Otol Rhinol Laryngol* 1995;104: 550–555.
40. Shaw GY, Searl JP. Laryngeal manifestations of gastroesophageal reflux before and after treatment with omeprazole. *South Med J* 1997;90: 1115–1122.
41. Habermann W, Eherer A, Lindbichler F, Raith J, Friedrich G. Ex juvantibus approach for chronic posterior laryngitis: results of short-term pantoprazole therapy. *J Laryngol Otol* 1999;113:734–739.
42. Belafsky PC, Postma GN, Koufman JA. The validity and reliability of the reflux finding score (RFS). *Laryngoscope* 2001;111:1313–1317.
43. Rodriguez-Tellez M, Galera-Ruiz H, Arguelles-Arias F, et al. Posterior laryngitis: effects of treatment with omeprazole alone. *Rev Esp Enferm Dig* 2002;94:123–130.
44. Habermann W, Kiesler K, Eherer A, Friedrich G. Short-term therapeutic trial of proton pump inhibitors in suspected extraesophageal reflux. *J Voice* 2002;16:425–432.
45. DelGaudio JM, Waring JP. Empiric esomeprazole in the treatment of laryngopharyngeal reflux. *Laryngoscope* 2003;113:598–601.
46. Bilgen C, Ogut F, Kesimli-Dinc H, Kirazli T, Bor S. The comparison of an empiric proton pump inhibitor trial vs 24-hour double-probe Ph monitoring in laryngopharyngeal reflux. *J Laryngol Otol* 2003;117:386–390.
47. Garrigues V, Gisbert L, Bastida G, et al. Manifestations of gastroesophageal reflux and response to omeprazole therapy in patients with chronic posterior laryngitis: an evaluation based on clinical practice. *Dig Dis Sci* 2003; 48:2117–2123.
48. Williams RB, Szczesniak MM, Maclean JC, Brake HM, Cole IE, Cook IJ. Predictors of outcome in an open label, therapeutic trial of high-dose omeprazole in laryngitis. *Am J Gastroenterol* 2004;99:777–785.
49. Sereg-Bahar M, Jansa R, Hocevar-Boltezar I. Voice disorders and gastroesophageal reflux. *Logoped Phoniatr Vocol* 2005;30:120–124.
50. Qua CS, Wong CH, Gopala K, Goh KL. Gastro-oesophageal reflux disease in chronic laryngitis: prevalence and response to acid-suppressive therapy. *Aliment Pharmacol Ther* 2007;25:287–295.
51. Jin bJ, Lee YS, Jeong SW, Jeong JH, Lee SH, Tae K. Change of acoustic parameters before and after treatment in laryngopharyngeal reflux patients. *Laryngoscope*. 2008;118:938–941.
52. Koufman JA Low-acid diet for recalcitrant laryngopharyngeal reflux: therapeutic benefits and their implications. *Ann Otol Rhinol Laryngol* 2011; 120:281–287.
53. Masaany M, Marina MB, Sharifa Ezat WP, Sani A. Empirical treatment with pantoprazole as a diagnostic tool for symptomatic adult laryngopharyngeal reflux. *J Laryngol Otol* 2011;125:502–508.
54. Naiboglu B, Durmus R, Tek A, Toros SZ, Egeli E. Do the laryngopharyngeal symptoms and signs ameliorate by empiric treatment in patients with suspected laryngopharyngeal reflux? *Auris Nasus Larynx* 2011;38:622–627.
55. Patigaroo SA, Hashmi SF, Hasan SA, Ajmal MR, Mehfooz N. Clinical manifestations and role of proton pump inhibitors in the management of laryngopharyngeal reflux. *Indian J Otolaryngol Head Neck Surg* 2011;63: 182–189.
56. Habermann W, Schmid C, Neumann K, Devaney T, Hammer HF. Reflux symptom index and reflux finding score in otolaryngologic practice. *J Voice* 2012;26:e123–e127.
57. Park JO, Shim MR, Hwang YS, et al. Combination of voice therapy and antireflux therapy rapidly recovers voice-related symptoms in laryngopharyngeal reflux patients. *Otolaryngol Head Neck Surg* 2012;146:92–97.
58. Lee JS, Lee YC, Kim SW, Kwon KH, Eun YG. Changes in the quality of life of patients with laryngopharyngeal reflux after treatment. *J Voice* 2014; 28:487–491.
59. Semmanaselvan K, Mukaddam QI, Naik M. An open label, prospective, single centre study to evaluate the efficacy and safety of fixed dose combination of rabeprazole (enteric-coated, EC) 20 mg + domperidone (sustained release, SR) 30 mg capsule in treatment of patients with laryngopharyngeal reflux disease. *J Assoc Physicians India* 2015;63:27–32.
60. Batıoglu-Karaaltin A, Develioglu ON, Tarhan O, Kulekci M. The importance of voice analysis in evaluating the effectiveness of reflux treatment. *Kulak Burun Bogaz Ihtis Derg* 2016;26:207–212.
61. Dulery C, Lechot A, Roman S, et al. A study with pharyngeal and esophageal 24-hour pH-impedance monitoring in patients with laryngopharyngeal symptoms refractory to proton pump inhibitors. *Neurogastroenterol Motil*. 2017. doi: 10.1111/nmo.12909.
62. Joshi AA, Chiplunkar BG, Bradoo RA. Assessment of treatment response in patients with laryngopharyngeal reflux. *Indian J Otolaryngol Head Neck Surg* 2017; 69:77–80.
63. Pullarat AN, Attakkil A, Sibin. MA, Nayak DR, Pillai S. Hoarseness and laryngopharyngeal reflux: a prospective study. *IOSR J Dent Med Sci* 2017;16:41–50.
64. 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.
65. Prinsen CA, Vohra S, Rose MR, et al. Core outcome measures in effectiveness trials (COMET) initiative: protocol for an international Delphi study to achieve consensus on how to select outcome measurement instruments for outcomes included in a 'core outcome set'. *Trials* 2014;15:247.
66. Karkos PD, Wilson JA. Empiric treatment of laryngopharyngeal reflux with proton pump inhibitors: a systematic review. *Laryngoscope* 2006;116: 144–148.
67. Yang LJ, Chang KW, Chung KC. Methodologically rigorous clinical research. *Plast Reconstr Surg* 2012;129:979e–988e.
68. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.
69. Harris MS, Rotenberg BW, Roth K, Sowerby LJ. Factors associated with lingual tonsil hypertrophy in Canadian adults. *J Otolaryngol Head Neck Surg* 2017;46:32.
70. Rousson V, Gasser T, Seifert B. Assessing intrarater, interrater and test-retest reliability of continuous measurements. *Stat Med* 2002;21: 3431–3446.
71. Chang BA, MacNeil SD, Morrison MD, Lee PK. The Reliability of the reflux finding score among general otolaryngologists. *J Voice* 2015;29:572–577.
72. Belafsky PC, Postma GN, Koufman JA. Laryngopharyngeal reflux symptoms improve before changes in physical findings. *Laryngoscope* 2001;111: 979–981.
73. Musser J, Kelchner L, Neils-Strunjas J, Montrose M. A comparison of rating scales used in the diagnosis of extraesophageal reflux. *J Voice* 2011;25: 293–300.
74. Vailati C, Mazzoleni G, Bondi S, Bussi M, Testoni PA, Passaretti S. Oropharyngeal pH monitoring for laryngopharyngeal reflux: is it a reliable test before therapy? *J Voice* 2013;27:84–89.