Inter-rater Reliability of the Reflux Sign Assessment-10 (RSA-10)

[*,†,‡,§1]Jerome R. Lechien, [¶,1]Chloe Lebrun, [¶,1]Juliette Piquard, [‖]Lisa G. De Marrez, [‡]Laura Bousard, and [¶]Nadine Gallant, *Mons, ‡Baudour, ¶Saint-Ghislain, ‖Brussels, Belgium, †Poitiers, and §Paris, France

ABSTRACT: Objective. To evaluate the inter-rater reliability and internal consistency of the Reflux Sign Assessment-10 (RSA-10) among otolaryngologists and speech therapists with various experiences.

Methods. Six experts (2 otolaryngologists, 2 speech-therapists, and 2 speech-therapist students) rated 300 clinical images of oral, laryngeal, and pharyngeal signs from patients with laryngopharyngeal reflux disease diagnosis at the 24-hour hypopharyngeal-esophageal multichannel intraluminal impedance-pH monitoring. Inter-rater reliability and internal consistency were evaluated with Intraclass Correlation (ICC) and Cronbach-α. The severity of scores was compared between judges. The intra-rater (test-retest) reliability was evaluated with the Spearman correlation coefficient.

Results. The pictures of 40 patients were included. There were 18 females and 22 males. The mean age was 52.6 ± 13.9 years. The Cronbach-α was 0.854, which indicates a high internal consistency between judges. The overall ICC was 0.787 (95% CI: 0.715–0.845; $P = 0.001$). The ICC varied among judges with the highest value for students (ICC = 0.960) and SLP seniors versus students (ICC = 0.805). The severity of RSA-10 rating scores was influenced by the number of reflux patients seen ($r_s = -0.941$; $P = 0.001$) and the number of fiberscope examinations performed ($r_s = -0.812$; $P = 0.049$). The RSA-10 was more severely scored by speech therapists with the least experience compared to otolaryngologists with the most experience in fiberscope/reflux patient assessment.

Conclusion. The RSA-10 demonstrated adequate global ICC and internal consistency among otolaryngologists and speech therapists with various degrees of experience. The assessment of RSA was influenced by the fibroscopy experience, and the number of reflux patients seen.

Key Words: Laryngopharyngeal–Reflux–Gastroesophageal reflux–Consistency–Inter-rater–Reliability–Reflux Sign Assessment–Signs–Findings.

INTRODUCTION

Laryngopharyngeal reflux disease (LPRD) is a prevalent disorder in otolaryngology-head and neck surgery, and laryngology offices.[1,2] The clinical diagnosis of LPRD is challenging because most laryngeal, pharyngeal, and oral symptoms and findings are non-specific[3,4] and they can be found in common inflammatory conditions of the upper aerodigestive tract such as allergy,[5] chronic rhinosinusitis,[6] or tobacco-induced laryngopharyngitis.[7] The non-specificity of symptoms and findings led some authors to develop patient-reported outcome questionnaires and clinical instruments documenting symptoms and findings, respectively.[8–10] Reflux Finding Score (RFS)[11] and Reflux Sign Assessment (RSA)[12] are the two most popular validated instruments for documenting and rating LPRD findings. RFS was developed to rate the severity of laryngeal findings.[11] RFS is easily completed in the office but studies revealed that it is associated with low inter-rater reliability,[13,14] which can be attributed to the subjective evaluation of some items (mild, moderate, or severe).[10] RSA and the short version, RSA-10,[15] reported higher inter-rater reliability compared to RFS[12,15,16] because their items are rated as descriptively as possible. However, only otolaryngologists were involved in RFS and RSA studies, while in many Western countries, speech therapists can carry out stroboscopy in their clinical practice.

The objective of the present study was to evaluate the inter-rater reliability and internal consistency of RSA-10 among otolaryngologists and speech therapists with various experiences.

METHODS

Reflux Sign Assessment-10

The development, validity, and reliability assessments of the RSA-10 were described in a recent paper (Figure 1).[15] In sum, some closely related RSA items reporting significant associations were merged into one item in RSA-10 to reduce the number of findings. RSA-10 was found to be completed in less than 1 min.[15] As for the RSA, the RSA-10 items are rated with a scoring system designed to minimize subjective ratings, such as “mild”, “moderate”, or

Accepted for publication September 16, 2024.

From the *Department of Surgery, UMONS Research Institute for Health Sciences and Technology, University of Mons (UMons), Mons, Belgium; †Department of Otolaryngology, Elsan Polyclinic of Poitiers, Poitiers, France; ‡Division of Laryngology and Bronchoesophagology, Department of Otolaryngology-Head and Neck Surgery, UMONS, Baudour, Belgium; §Department of Otolaryngology-Head Neck Surgery, Foch Hospital, School of Medicine, UFR Simone Veil, Université Versailles Saint-Quentin-en-Yvelines (Paris Saclay University), Paris, France; ¶Haute Ecole Condorcet, Division of Speech Therapy, Saint-Ghislain, Belgium; and the ‖Department of Otolaryngology-Head & Neck Surgery, CHU Brugmann, Brussels, Belgium

[1] Authors have similarly contributed to the study and can be joined as cofirst authors..

Address correspondence and reprint requests to Jerome R. Lechien, Department of Surgery, Faculty of Medicine and Pharmacy, UMONS Research Institute for Health Sciences and Technology, Avenue du Champ de mars, 6, B7000 Mons, Belgium. E-mail: Jerome.Lechien@umons.ac.be

Journal of Voice, Vol xx, No xx, pp. xxx–xxx

0892-1997

https://doi.org/10.1016/j.jvoice.2024.09.025

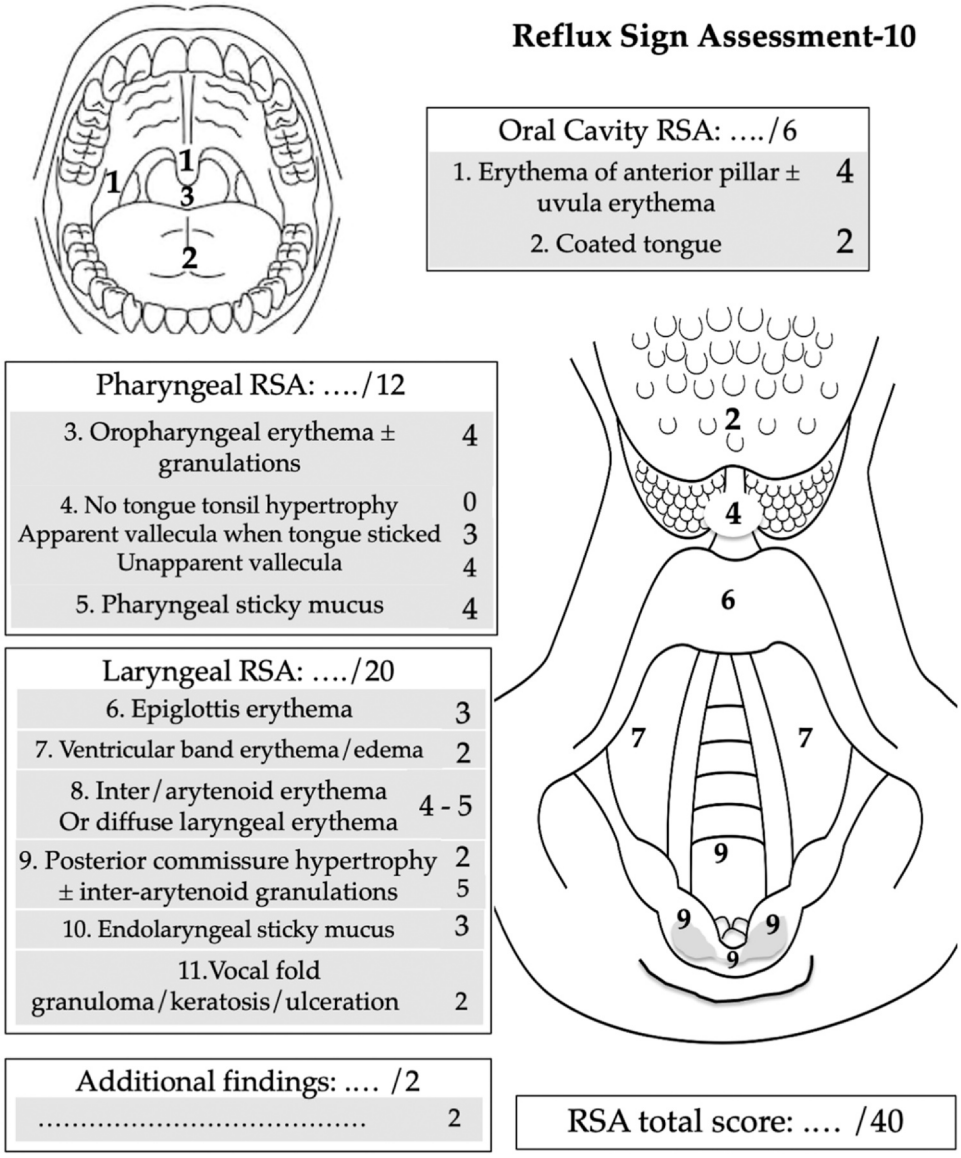

FIGURE 1. Reflux Sign Assessment-10 (RAS-10). RSA-10 is composed of the most prevalent oral, pharyngeal and laryngeal signs associated with laryngopharyngeal reflux disease (LPRD). Additional signs (uncommon) may be added, including nasal findings (Mulberry turbinate or dry mucosa).

"severe" signs. The RSA-10 reported high internal consistency reliability (α = 0.822), test-retest reliability (r_s =0.725), internal validity, and adequate responsiveness to change when used by experienced otolaryngologists.[15] The inter-rater reliability was adequate for sub- and total RSA-10 scores (k = 0.708) but it was exclusively evaluated among experienced board-certified otolaryngologists.[15]

Subjects and setting

The present study included data from 40 patients with LPRD symptoms and findings who were consecutively recruited from January 2020 to December 2023 in the Departments of Otolaryngology-Head and Neck Surgery of three European hospitals: CHU Saint-Pierre (Brussels, Belgium), EpiCURA Hospital (Baudour, Belgium), and Elsan Polyclinique of Poitiers (Poitiers, France). The diagnosis of LPRD was based on the Dubai criteria,[17] which consist of the occurrence of more than one pharyngeal reflux events during the 24-hour hypopharyngeal-esophageal multichannel intraluminal impedance-pH monitoring (HEMII-pH). Additional information about the HEMII-pH placement and analyses is reported in a previous study.[15] The symptoms of patients were evaluated with the French version of the RSS.[18] The exclusion criteria included alcohol dependence, smoking, patients with upper respiratory tract infection within the last month, asthma, inhaled corticosteroid-induced laryngitis, neurological or psychiatric conditions, previous history of neck surgery or trauma, malignancy, history of head and neck radiotherapy and active seasonal allergies. The local ethics committee approved the study protocol (no BE076201837630).

Expert features

Six experts with various degrees of experience were recruited to rate upper aerodigestive tract images with the RSA-10. The features of the experts are described in Appendix 1. The group included a board-certified otolaryngologist with a fellowship in laryngology (J.R.L.) 12 years of experience as an otolaryngologist (considering the residency in otolaryngology-head and neck surgery), and a fellow otolaryngologist (L.D.M.) with 6 years of experience since the first PGY year in otolaryngology. Both otolaryngologists have an academic practice. They have performed a mean of 12 000 and 6000 nasofibroscopy for 1500 to 4000 patients with LPRD, respectively (Appendix 1). The speech therapist (CCC-SLP) group included two board-certified CCC-SLPs with 7 and 24 years of experience since the CCC-SLP degree and two students. The CCC-SLP fellow was active in laryngology and performed more laryngoscopy than the experienced CCC-SLP. Students were trained to analyze laryngostroboscopy with the first author of the paper (J.R.L.) for 6 months.

Statistical methods

Statistical analyses were performed using the Statistical Package for the Social Sciences for Windows (SPSS version 29,0; IBM Corp, Armonk, NY, USA). The clinical images were collected from the consultations of the main investigator (J.R.L.). Clinical pictures of patients with LPRD seen at baseline, 3-month, or 6-month posttreatment were collected for rating the RSA-10 by the experts. The experts were blinded regarding the time of the picture (pretreatment versus post-treatment). Photos included pictures of the mouth, pharynx, and larynx. The distribution of photos was 40% of photos of the pretreatment time, 40% of the 3-month post-treatment time, and 20% of the 6-month post-treatment time, respectively. Figure 2 shows some examples of RSA-10 evaluations. Because the responsiveness to change, internal and convergent validities were previously reported among otolaryngologists,[15] the authors only focused on the evaluation of the inter-rater reliability with intraclass correlation coefficient (ICC, concordance reliability), and internal consistency (Cronbach-α). Potential

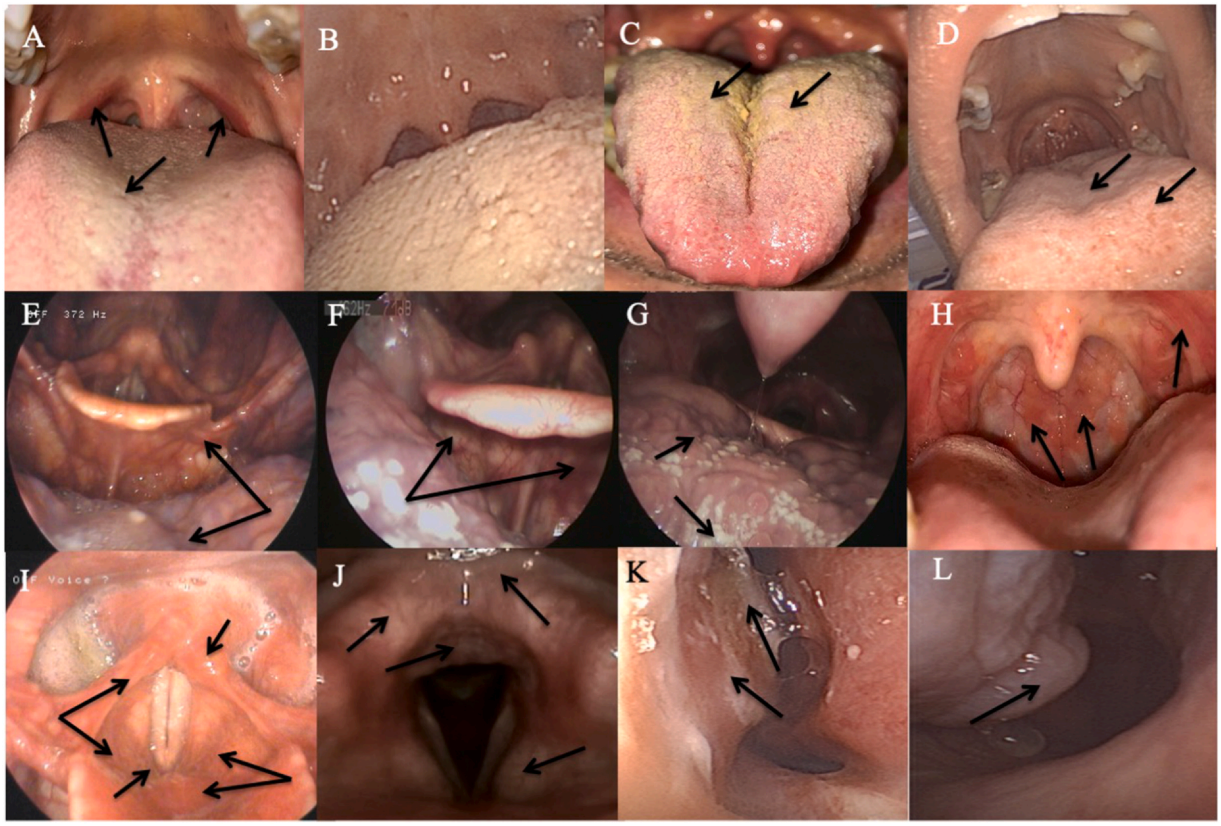

FIGURE 2. Findings of Reflux Sign Assessment-10. Anterior pillar erythema can be present (**A**) or absent (**B**). Patients can have coated tongue (**A**, **C**) or not (**D**). The tongue tonsil hypertrophy may be absent (apparent vallecula at a rest position of the tongue, **E**); mild-to-moderate when the vallecula are apparent only when the tongue is protruded (**F**); or unapparent (**G**). Note that the coated tongue can be observed in the posterior part of tongue (**G**). Granulation and erythema of the oropharynx are commonly observed through the tongue (**H**) or endoscopically. Laryngeal findings include erythema of posterior commissure, ventricles, and epiglottis (**I**). The hypertrophy of the posterior commissure can be absent (**I**) or present (**J**) with granulation tissue between both arytenoids (**J**). Nasal mucosa dryness/crusts (**K**) and mulberry inferior turbinate (**L**) are atypical findings associated with laryngopharyngeal reflux disease.

differences in rating RSA-10 were additionally explored by a comparison of the severity of RSA-10 items and total scores among judges with the Kruskal-Wallis test. The intra-rater (test-retest) reliability was evaluated with the Spearman correlation coefficient. Experts were invited to evaluate photos twice with an interval of two months. The association between the judge features (eg, years of experience, number of fiberscope examinations, and age) and the RSA-10 score was investigated with the Spearman correlation coefficient. A level of significance of $P < 0.05$ was used.

RESULTS

Three hundred clinical images of oral, laryngeal, and pharyngeal cavities related to 40 patients with LPRD were independently rated by the 6 judges in a blinded manner regarding the symptoms of patients. The features of patients are described in Table 1. There were 18 (55%) females and 22 (45%) males. The mean age was 52.6 ± 13.9 years. The demographics, HEMII-pH, and clinical features are reported in Table 1. Patients had primarily acid pharyngeal reflux events.

The internal consistency of RSA-10 is reported in Table 2. Cronbach-α was 0.854, which indicates a high internal consistency across judges of various specialties. The ICC for the 6 raters was 0.787 (95% CI: 0.715–0.845; $P = 0.001$). The inter-rater reliability of all subgroups was statistically significant (Table 3). The highest ICC values were found between CCC-SLP students (ICC = 0.960) and between the students and the senior CCC-SLP (ICC = 0.805). The fellow and the board-certified otolaryngologist reported the lowest ICC values (Table 3). The comparison of RSA-10 items and total scores between judges is available in Table 4. The mean sub- and total RSA-10 scores significantly differed between judges. The scores of the two otolaryngologists were overall lower compared to the scores of CCC-SLP (trained) students, fellows, and seniors. The ICC data are reported in Appendix 2. The test-retest evaluations reported a high Spearman rho coefficient for senior otolaryngologist/CCC-SLP and students. The senior otolaryngologist had the lowest coefficient ($r_s = 0.667$) and one student had the highest coefficient ($r_s = 0.889$). There was a strong negative association between the number of patients seen in the career and the RSA-10 severity ($r_s = -0.941$; $P = 0.001$). The association between the number of fiberscopes performed and the RSA-10 score was strongly negative ($r_s = -0.812$; $P = 0.049$).

TABLE 1.
Demographics and Clinical Features

Characteristics	Patients (N = 40)
Mean age (range, years)	52.6 ± 13.9
Body mass index (mean, SD)	24.9 ± 4.5
Gender (N, %)	
Male	22 (55)
Female	18 (45)
HEMII-pH feature (mean, SD)	
Pharyngeal events	
Pharyngeal acid reflux events	22.1 ± 19.3
Pharyngeal nonacid reflux events	8.9 ± 8.8
Total number of pharyngeal events	31.2 ± 24.2
Clinical data (mean, SD)	
Otolaryngological RSS	56.2 ± 42.0
Digestive RSS	54.4 ± 42.6
Respiratory RSS	24.5 ± 23.4
Total RSS	135.1 ± 84.3
Quality-of-life RSS	35.6 ± 20.6

Abbreviations: HEMII-pH, hypopharyngeal-esophageal multichannel intraluminal impedance-pH monitoring; N, number; RSS, reflux symptom score; SD, standard deviation.

TABLE 2.
Internal Consistency

RSA-10 items	Cronbach- α
All judges	0.854
Senior otolaryngologist	0.569
Fellow otolaryngologist	0.607
Fellow CCC-SLP	0.620
Senior CCC-SLP	0.658
Student CCC-SLP 1	0.548
Student CCC-SLP 2	0.561

Abbreviations: SLP, speech language pathologist; RSA-10, Reflux Sign Assessment-10.

DISCUSSION

The high prevalence of LPRD[1,2] and the increased use of validated clinical instruments[9] make important the assessment of the consistency and the inter-rater reliability of these instruments among specialties. To date, several clinical instruments have been validated for documenting clinical findings associated with LPRD, including the RFS,[11] RSA,[12] RSA-10,[15] laryngoscopic grading scale (LGS),[19] laryngeal reflux grade (LRG),[20] laryngopharyngeal reflux disease index (LRDI),[21] and the Hick's instrument.[22] When assessed by a panel of judges with several experiences, the consistency of RSA-10 (α = 0.854) was close to the consistency values found for the RSA (α = 0.821)[12], and the RSA-10 (α = 0.822)[15] when evaluated by experienced otolaryngologists. To the best of our knowledge, the internal consistency was not evaluated for the other clinical instruments.[9,15] In the present study, the overall internal consistency is moderate to high but the evaluation per judge is low. The internal consistency is defined as the extent to which items within each domain are interrelated. The low value found for RSA-10 could be related to the fact that the instrument evaluates pharyngeal—oral—and laryngeal findings, which are not interrelated in terms of inflammation and clinical findings. Moreover, the variability across judges can be attributed to

TABLE 3.
Inter-rater Reliability

RSA-10	N	ICC	95% CI	P value
All raters	6	0.787	0.715 - 0.845	0.001
CCC-SLP students	2	0.960	0.940 - 0.973	0.001
Otolaryngologists	2	0.588	0.388 - 0.723	0.001
CCC-SLP graduated	2	0.632	0.454 - 0.753	0.001
Otolaryngologists—senior CCC-SLP	4	0.690	0.578 - 0.779	0.001
Senior—students CCC-SLP	4	0.805	0.728 - 0.863	0.001

Abbreviations: 95% CI, 95% confident interval; ICC, interclass correlation; N, number; SLP, speech language pathologist; RSA-10, reflux sign assessment-10.

the clinical experience of judges in the evaluation of several anatomical regions. Indeed, otolaryngologists are more trained in the evaluation of the larynx, pharynx, and oral cavity and they can consequently more easily detect normal versus pathological findings than CCC-SLP. Better knowledge about normal or pathological tissue findings leads to an increased number of differences in ratings of anatomical pathological and normal regions, which can be interpreted as a low Cronbach-α.

In the past few decades, a primary objective of many teams was to develop a reliable clinical instrument documenting the LPRD findings with high inter-rater reliability.[11,12,15,19–22] Most clinical instruments reported high intra-rater reliability but the inter-rater reliability was variable according to studies.[11–15,19–22] The RSA-10 reports moderate-to-high global inter-rater reliability across all judges with various profiles (ICC = 0.787), which corroborates the values found in the initial studies validating the RSA-10 (0.708),[15] RSA (0.66),[12] and LGS (0.75–0.93).[19] When considering the ICC evaluation within the judge groups (students; otolaryngologists), the values substantially varied. The moderate inter-rater reliability of RSA-10 within some judges is probably associated with their differences in terms of experience, which especially concerns the otolaryngologist group. In the literature, the studies investigating clinical instruments (eg, 'mild', 'moderate', or 'severe'), for example, RFS,[13,14] the Hicks's instrument (0.32–0.58),[22] LRG (0.43),[20] LRDI (0.30),[21] reported lower global inter-rater reliabilities, which can be related to the use of subjective score (mild, moderate, severe) rather than more descriptive scores. However, the RSA-10 ICC results can be encouraging because in most other studies,[12–15,20,22] the judges had a similar profile (laryngologists or otolaryngologist potentially sharing the same experience).

In the present study, the intra-rater reliability (test-retest) was adequate for all judges with the highest values for the students and the lowest value for the senior otolaryngologist. The high intra-rater reliability data are an important point, which supports the use of RSA-10 by the same practitioners in the follow-up process of patients. The assessment of clinical findings, such as fiberscope images, is influenced by the judge's experience and the last evaluations performed by the judges, for example, on the day prior to the evaluations.[24] In the present study, the assessment of the senior otolaryngologist could be influenced by its last clinical examinations performed in the hours before the RSA-10 evaluations, while the students evaluated the pictures without consultations.

The variability of inter-rater reliability among practitioners with various degrees of clinical experience was not previously investigated. Clinical experience is an important factor influencing the severity of judgments (RSA-10). Young speech pathologists/students evaluate the LPRD findings more severely than the otolaryngologists, which can be related to their lack of experience in the evaluation of normal and pathological laryngopharyngeal and oral mucosa. The influence of experience on the severity of scores was corroborated in previous studies showing a relationship between the judge's experience in the fibroscopy[23] or perceptual voice quality[25] evaluations. The intra-rater (test-retest) reliability values found in this study were similar to those found for the initial versions of RSA,[12] RSA-10,[15] RFS,[11] LRG,[20] and LRDI,[21] which were all validated among board-certified otolaryngologists. The adequate intra-rater reliability supports the use of RSA-10 in clinical practice in evaluations performed by the same practitioners over time despite their clinical experience.

In most European countries, speech therapists cannot perform fiberscope examinations, including stroboscopy, while in the United States, they commonly can. In this study, we observed that speech therapists report good consistency and inter-rater reliability in using RSA-10, which can contribute to the debate[26] about the evolution of the practice of speech therapists in Europe.

The primary strength of the present study is the consideration of patients with a demonstrated diagnosis of LPRD at the HEMII-pH. Indeed, the non-specificity of symptoms and signs can bias the diagnosis of LPRD and, consequently, the validation of a clinical instrument dedicated to the documentation of LPRD signs. The lack of consideration of a control group (asymptomatic

TABLE 4.
RSA-10 Scores of Judges

	Otolaryngologists		CCC-SLP				
RSA-10	Senior	Fellow	Student 1	Student 2	Fellow	Senior	P value
1. Anterior pillar/uvula erythema	3.04 ± 1.72	3.24 ± 1.58	3.88 ± 0.69	3.80 ± 0.88	4.00 ± 0.01	4.00 ± 0.4	0.001
2. Coated tongue	1.46 ± 0.89	1.82 ± 0.58	1.86 ± 0.51	1.88 ± 0.48	2.00 ± 0.01	1.94 ± 0.34	0.001
3. Posterior oropharyngeal wall erythema/granulation	1.36 ± 1.90	2.79 ± 1.84	4.00 ± 0.01	4.00 ± 0.01	4.00 ± 0.01	3.96 ± 0.40	0.001
4. Tongue tonsil hypertrophy	3.33 ± 1.13	3.00 ± 1.45	2.89 ± 1.61	2.88 ± 1.65	3.50 ± 1.16	3.71 ± 10.57	0.001
5. Pharyngeal sticky mucus	0.96 ± 1.72	2.82 ± 1.83	3.84 ± 0.79	3.89 ± 0.69	2.28 ± 2.00	3.11 ± 1.67	0.001
6. Epiglottis erythema	2.01 ± 1.42	1.15 ± 1.47	2.97 ± 0.30	2.97 ± 0.30	3.00 ± 0.01	2.94 ± 0.42	0.001
7. Ventricular band erythema/edema	1.40 ± 0.94	0.57 ± 0.91	2.00 ± 0.01	2.00 ± 0.01	2.00 ± 0.01	1.92 ± 0.39	0.001
8. Arytenoid, posterior commissure erythema	2.31 ± 2.12	3.90 ± 1.46	4.94 ± 0.24	4.94 ± 0.24	4.92 ± 0.27	4.58 ± 0.92	0.001
9. Posterior commissure/retrocricoid hypertrophy	3.55 ± 2.28	2.92 ± 2.48	4.49 ± 1.51	4.60 ± 1.38	3.95 ± 2.05	4.75 ± 1.10	0.001
10. Endolaryngeal sticky mucus	1.20 ± 1.48	1.84 ± 1.47	1.59 ± 1.50	1.59 ± 1.50	1.86 ± 1.46	1.41 ± 1.50	0.013
11. Vocal cord granuloma, keratosis, or ulceration	0.12 ± 0.48	0.16 ± 0.55	0.64 ± 0.94	0.65 ± 0.94	0.14 ± 0.51	1.20 ± 0.98	0.001
RSA-10	20.17 ± 5.82	23.67 ± 6.20	32.69 ± 3.51	32.49 ± 3.59	31.65 ± 4.07	33.51 ± 4.10	0.001

The comparison between judges was carried out with Kruskal-Wallis test. *Abbreviations:* SLP, speech language pathologist; RSA-10, reflux sign assessment-10.

individuals) in the test-retest reliability is the primary limitation of the study. We did not include a control group because the objective of the study was to investigate the inter-rater reliability of RSA-10 among several specialists. The internal validity of RSA-10 scores was compared in a recent study where LPRD patients had significantly higher RSA-10 compared to controls ($p = 0.001$).[15] In the RSA-10 validation study,[15] the minimal clinically important difference score was not calculated, which is an additional limitation. The authors did not include asymptomatic individuals because the study aimed to investigate the inter-rater reliability of RSA-10 among speech therapists and otolaryngologists; the RSA-10 has already been validated in a study including both patients and asymptomatic individuals.[15] However, the consideration of normal examination should influence the intrarater or inter-rater reliability values of inexperienced practitioners.

CONCLUSION

The development of clinical instruments with high inter-rater reliability is an important issue in otolaryngology-head and neck surgery regarding the prevalence of LPRD. The RSA-10 reported high inter-rater reliability and internal consistency among otolaryngologists and speech therapists with various degrees of experience. The intra-rater reliability data support the use of RSA-10 in practitioners despite their clinical experience. The results of the present study support a significant influence of the clinical experience in the assessment of the findings, consisting of a new element that can potentially explain the low inter-rater reliability findings of some LPRD clinical instruments.

Ethic committee

The local IRB approved the study protocol (CHU Saint-Pierre, Brussels, no BE076201837630).

Funding: Vesale Grant, IRIS Grant.

Declaration of Competing Interest

Authors have no conflict of interest.

Acknowledgments

None.

Informed consent

Patients consented to participate.

Sponsorships

None.

Appendix 1: Judge Features

	Age	Gender	Laryngoscopy	Reflux patients	Years	RSA-10
Senior laryngologist	36	Male	12000	4000	12	20.17
Fellow laryngologist	29	Female	6000	1500	6	23.67
Senior CCC-SLP	45	Female	130	100	24	33.51
Fellow CCC-SLP	32	Female	300	250	7	31.65
Student 1	20	Female	100	100	1	32.69
Student 2	20	Female	100	100	1	32.49

Experts evaluated the number of laryngoscopy and reflux patients according to these annual data found in the hospital. *Abbreviations*: RSA-10, reflux sign assessment-10; SLP, speech language pathologist.

Appendix 2: Intra-Rater Reliability

RSA-10 items	Spearman rho	*P* value
Senior otolaryngologist	0.667	0.001
Fellow CCC-SLP	0.845	0.001
Senior CCC-SLP	0.702	0.001
Student CCC-SLP 1	0.889	0.001
Student CCC-SLP 2	0.862	0.001

The fellow otolaryngologist did not perform the retest reliability analysis.

References

1. Kamani T, Penney S, Mitra I, Pothula V. The prevalence of laryngopharyngeal reflux in the English population. *Eur Arch Otorhinolaryngol.* 2012;269:2219–2225. https://doi.org/10.1007/s00405-012-2028-1.
2. Printza A, Kyrgidis A, Oikonomidou E, Triaridis S. Assessing laryngopharyngeal reflux symptoms with the Reflux Symptom Index: validation and prevalence in the Greek population. *Otolaryngol Head Neck Surg.* 2011;145:974–980. https://doi.org/10.1177/0194599811425142.
3. Chen M, Hou C, Chen T, et al. Reflux symptom index and reflux finding score in 91 asymptomatic volunteers. *Acta Otolaryngol.* 2018;138:659–663.
4. Lechien JR. Sensitivity, specificity, and predictive values of laryngopharyngeal reflux symptoms and signs in clinical practice. *Otolaryngol Head Neck Surg.* 2023;169:97–104. https://doi.org/10.1177/01945998221121822.
5. Eren E, Arslanoğlu S, Aktaş A, et al. Factors confusing the diagnosis of laryngopharyngeal reflux: the role of allergic rhinitis and inter-rater variability of laryngeal findings. *Eur Arch Otorhinolaryngol.* 2014;271:743–747. https://doi.org/10.1007/s00405-013-2682-y.
6. Lechien JR, Saussez S, Hopkins C. Association between laryngopharyngeal reflux, gastroesophageal reflux and recalcitrant chronic rhinosinusitis: a systematic review. *Clin Otolaryngol.* 2023;48:501–514. https://doi.org/10.1111/coa.14047.
7. Kayalı Dinc AS, Cayonu M, Sengezer T, Sahin MM. Smoking cessation improves the symptoms and the findings of laryngeal irritation. *Ear Nose Throat J.* 2020;99:124–127. https://doi.org/10.1177/0145561319881559.
8. 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. https://doi.org/10.1177/0194599816664330.
9. Lechien JR, Schindler A, De Marrez LG, et al. Instruments evaluating the clinical findings of laryngopharyngeal reflux: a systematic review. *Laryngoscope.* 2018;129:720–736.
10. Lechien JR, Akst LM, Hamdan AL, et al. Evaluation and management of laryngopharyngeal reflux disease: state of the art review. *Otolaryngol Head Neck Surg.* 2019;160:762–782. https://doi.org/10.1177/0194599819827488.
11. Belafsky PC, Postma GN, Koufman JA. The validity and reliability of the reflux finding score (RFS). *Laryngoscope.* 2001;111:1313–1317.
12. Lechien JR, Bobin F, Muls V, et al. Validity and reliability of the Reflux Sign Assessment (RSA). *Ann Otol Rhinol Laryngol.* 2019;129:313–325. https://doi.org/10.1177/0003489419888947.
13. Vance D, Alnouri G, Shah P, et al. The validity and reliability of the reflux finding score. *J Voice.* 2023;37:92–96. https://doi.org/10.1016/j.jvoice.2020.11.008.
14. Jahshan F, Ronen O, Qarawany J, et al. Inter-rater variability of reflux finding score amongst otolaryngologists. *J Voice.* 2022;36:685–689. https://doi.org/10.1016/j.jvoice.2020.07.021.
15. Lechien JR, De Marrez LG, Finck C, Saussez S. Validity and reliability of the Reflux Sign Assessment-10 (RSA-10). *Laryngoscope.* 2024;134:3981–3988. https://doi.org/10.1002/lary.31420.
16. Rezaei Fard H, Khoddami SM, Maaroufizadeh S, et al. The persian version of Reflux Sign Assessment scale: validity and reliability for the examination of patients with laryngopharyngeal reflux disease. *J Voice.* 2023. https://doi.org/10.1016/j.jvoice.2023.06.019.
17. Lechien JR, Vaezi MF, Chan WW, et al. The Dubai definition and diagnostic criteria of laryngopharyngeal reflux: the IFOS consensus. *Laryngoscope.* 2024;134:1614–1624. https://doi.org/10.1002/lary.31134.
18. Lechien JR, Bobin F, Muls V, et al. Validity and reliability of the reflux symptom score. *Laryngoscope.* 2020;130:E98–E107. https://doi.org/10.1002/lary.28017.
19. Williams RB, Szczesniak MM, Maclean JC, et al. Predictors of outcome in an open label, therapeutic trial of high-dose omeprazole in laryngitis. *Am J Gastroenterol.* 2004;99:777–785.
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. Beaver ME, Stasney CR, Weitzel E, et al. Diagnosis of laryngopharyngeal reflux disease with digital imaging. *Otolaryngol Head Neck Surg.* 2003;128:103–108.
22. Hicks DM, Ours TM, Abelson TI, et al. The prevalence of hypopharynx findings associated with gastroesophageal reflux in normal volunteers. *J Voice.* 2002;16:564–579.
23. Alvi S, Harsha P. *Flexible Nasopharyngoscopy.* StatPearls; 2023.
24. Lechien JR, Morsomme D, Finck C, et al. The effect of the speech task characteristics on perceptual judgment of mild to moderate dysphonia: a methodological study. *Folia Phoniatr Logop.* 2018;70:156–164. https://doi.org/10.1159/000492219.
25. Alves JDN, Almeida AAF, Yamasaki R, Lopes LW. The influence of listener experience, measurement scale and speech task on the reliability of auditory-perceptual evaluation of vocal quality. *Codas.* 2024;36:e20230175. https://doi.org/10.1590/2317-1782/20232023175.
26. Hans S, Vialatte de Pemille G, Baudouin R. Post-laryngectomy voice prosthesis changes by speech-language pathologists: preliminary results. *J Clin Med.* 2022;11:4113. https://doi.org/10.3390/jcm11144113.