Validity and Reliability of the Persian Version of Reflux Symptom Score-12 in Patients with Laryngopharyngeal Reflux Disease

Azin Kohansal, Seyyedeh Maryam Khoddami, Noureddin Nakhostin Ansari, Jerome R. Lechien, and Kayvan Aghazadeh

Summary: Objective. Cross cultural adaptation of the reflux symptom score-12 (RSS-12) into Persian language and to evaluate its validity and reliability in the assessment of patients with laryngopharyngeal reflux disease (LPRD).

Study design. A cross-sectional and prospective cohort design.

Methods. A standard forward and backward translation was followed to cross-culturally adapt the RSS-12 into Persian language. To study discriminative validity, the RSS-12p was administrated to 63 patients with LPRD (40 men and 23 women; mean age: 39.26 ± 9.79 years) and 50 healthy volunteers (31 men and 19 women; mean age: 37.24 ± 10.28 years). The patients completed the reflux symptom index (RSI) to assess construct validity. The test-retest reliability was investigated in 31 patients (time interval = 7 days).

Results. There were no missing responses and floor or ceiling effects. The assessing of discriminative validity showed that the questionnaire was able to discriminate between patients with LPRD and healthy participants (P<0.001). Construct validity was confirmed by the Pearson correlation between the RSS-12p and the RSI (r_p = 0.87; P<0.00). The internal consistency was confirmed with Cronbach α 0.85 and 0.72 for the RSS-12p and quality of life (QoL), respectively. Test–retest reliability was excellent (ICC_agreement = 0.98 for the RSS-12p and 0.94 for QoL).

Conclusions. The Persian version of RSS-12 is a valid and reliable self-administered questionnaire for assessing LPRD in Persian-speaking patients.

Key Words: Laryngopharyngeal—Reflex—Persian language—Validity—Reliability.

INTRODUCTION
Retrograde flow of stomach content to the throat and larynx at the site of contact with upper respiratory tract tissues is known as laryngopharyngeal reflux disease (LPRD). LPRD symptoms are found in 10% to 15% of outpatients in otolaryngology clinics. The patients who receive the diagnosis of LPRD commonly report symptoms such as coughing, throat clearing, hoarseness, globus sensation, and throat pain. There are also pathologic findings in laryngoscopic examination of patients with LPRD.

Regarding the high prevalence of the symptoms in LPRD and lack of definitive diagnostic tests, self-assessment methods have been developed for use in the clinics. Various patient-reported outcome (PRO) questionnaires are designed based on the symptoms of patients with LPRD. The reflux symptom index (RSI) was the first tool for assessment of symptoms in patients with LPRD which was published by Belafsky et al in 2002. After that, the laryngopharyngeal reflux health-related quality of Life (LPR-HRQL) questionnaire, 34-item laryngopharyngeal reflux (LPR-34) questionnaire, and pharyngeal reflux symptom questionnaire (PRSQ) were designed to assess LPRD. In addition to the mentioned questionnaires, there are other questionnaires such as the throat questionnaire (TQ), the glasgow-edinburgh throat scale (GETS), and the supraesophageal reflux questionnaire (SERQ) targeting the patients with globus or other pharyngeal complaints and do not specifically evaluate patients with LPRD. These questionnaires while are useful because they require no special equipment or examinations, and also they are inexpensive compared to instrumental diagnostic methods for LPRD, they have some disadvantages. Most of the questionnaires which specifically designed for LPRD do not consider all common symptoms (e.g. throat pain, odynophagia, halitosis, regurgitations, nausea, and burps) and respiratory symptoms (e.g. cough, respiratory problems, and wheezing) in this pathological condition. Lechien et al (2019) developed a self-assessment 22-item questionnaire entitled reflux symptom score (RSS) for LPRD. This questionnaire examines all self-reported symptoms as well as respiratory symptoms that can be clinically observed in patients with LPRD. In addition, the original RSS evaluates quality of Life (QoL) in patients with LPRD. A Korean version of the RSS (K-RSS) was developed and validated in 2021. However, the original RSS was time-consuming to complete and therefore lacks clinical

ARTICLE IN PRESS
utility. Consequently, Lechien et al (2020) developed and validated an abbreviated version of the original RSS by removing and combining some items named as reflux symptom score–12 (RSS-12). The RSS-12 questionnaire contains 12 common items or symptoms in LPRD and consists of two scores, a total score and a QoL score which takes less time rather than the original RSS so it is easy to use in the clinical and research settings. In comparison to the RSI which is a common self-reported questionnaire in LPRD, the RSS-12 includes greater number of LPRD symptoms. Also, the RSS-12 evaluates both frequency and severity of the symptoms as well as QoL related to the symptoms in while the RSI only surveys the severity of LPRD symptoms. There is no Persian validated measure of the RSS-12 and therefore the purpose of the current study was to translate it into Persian language and evaluates its validity and reliability in patients with LPRD.

MATERIALS AND METHODS

Study design

This study had two phases of adaptation of the RSS-12 and validation. First phase followed a cross-sectional design to convert the RSS-12 into Persian language using backward and forward translation. Second phase followed a prospective design to assess the psychometric properties of the Persian RSS-12. All participants signed the written informed consent form before study initiation. The study protocol was approved by the institutional review board, School of Rehabilitation, and the Ethics Committee of Tehran University of Medical Sciences (Code number: IR.TUMS.FNM.REC.1399.241).

First phase: Cross cultural adaptation of the RSS-12p

After getting permission from developer of the original version of RSS-12, Jerome R. Lechien, cultural adaptation of the English version into Persian language was performed using the forward and backward translation procedure according to the recommended guideline as used previously. Two translators, one bilingual English translator and one speech and language pathologist (SLP), both of whom were native Persian speakers and fluent in English, translated the English version RSS-12 into Persian. Two translators, along with members of the expert panel, including two SLPs, an otorhinolaryngologist, and a methodologist discussed two Persian translations and produced an integrated Persian version of the questionnaire. At the next stage, the consensus Persian version of the questionnaire was translated into English by two other translators, separately. Each discrepancy was carefully analyzed in the meeting of expert panel. After confirmation of the English version by the developer of questionnaire, the pre-final Persian version was provided. In validity pilot testing, the pre-final Persian version of RSS-12 was given to 20 patients with LPRD (ranged from 18 to 60 years old) referred to the otorhinolaryngology clinic of Amir-A’laam Hospital in Tehran, Iran. The authors asked them to report any ambiguity in the questions or items of the questionnaire. The patients faced no problem in understanding any part of the questionnaire. Eventually, the final Persian version of the RSS-12 (RSS-12p) was produced (Appendix).

Second phase: Validity and reliability of the final RSS-12p

Participants and procedure

The current study consisted of 113 participants, 63 patients and 50 healthy people. The patients were enrolled from May 2021 to September 2021 from the otorhinolaryngology clinics of Amir-A’laam hospital and speech therapy clinics of Tehran University of Medical Sciences in Tehran, Iran. Inclusion criteria were: 1) aged ≥ 18 years; 2) presenting with common symptoms of LPRD (throat clearing, persistent cough, foreign body sensation in the throat, and altered voice quality); 3) a positive diagnosis of LPRD based on the reflux finding score (RFS) and RSI (RFS ≥ 7 or RSI > 13); 7); and 4) able to read and write in Persian language. The exclusion criteria were: smoking history, alcohol dependence, pregnancy, neurologic or psychiatric illness, upper respiratory tract infection within the last month, current use of antireflux treatment (ie, proton pump inhibitors, H2 blockers, alginate, and/or magaldrate), history of neck surgery or trauma, benign or malignant vocal fold lesions, history of ear, nose and throat radiotherapy, and active seasonal allergies or asthma. An otorhinolaryngologist and three SLPs who were experienced in the evaluation of voice disorders assessed all patients for eligibility. The patients included to the study completed both the RSS-12p and the RSI for construct validity.

Moreover, fifty healthy volunteers completed the RSS-12p for discriminant validity. The study population in the healthy group was selected based on the results of medical history (without symptoms of LPRD), self-reporting (RSI < 13), and laryngoscopy examination (RFS < 7). They also were homogeneous with the patient group in terms of age and gender. The exclusion criteria for the healthy group were similar to the patient group. To measure the test-retest reliability, 31 patients completed the RSS-12p again with an interval of one week.

Statistical analysis

The data were analyzed using SPSS statistics software for windows, version 23.0 (SPSS, Inc, Chicago, IL). Floor and ceiling effects were computed for RSS-12p total scores. If more than 15% of respondents scored the lowest or highest possible score, this would indicate floor or ceiling effects. The discriminant validity of the RSS-12p was analyzed by using the independent t test. The construct validity was tested by calculating a Pearson correlation coefficient between the RSS-12p and the RSI. The correlation values were interpreted as 0.81–1.0 (Excellent), 0.61–0.80 (Very strong), 0.41–0.60 (Good), 0.21–0.40 (Fair),
and 0.00 to 0.20 (Poor).\textsuperscript{29} Internal consistency reliability was assessed using Cronbach’s alpha ($\alpha$); the cut-off value of 0.70 was interpreted as acceptable.\textsuperscript{27} The intraclass correlation coefficient, absolute measure (ICC \textit{agreement}) (two-way random effects model, single measure) was calculated to measure test-retest reliability. The ICC values were interpreted as follows: $\geq 0.75$ (Excellent); 0.60–0.75 (Good), and 0.40 to 0.59 (Fair).\textsuperscript{30} Absolute reliability refers to the degree to which repeated measures vary for individual.\textsuperscript{31} It is obtained by calculating the standard error of measurement (SEM).\textsuperscript{32} The SEM has been introduced as “the determination of the amount of variation or spread in the measurement errors for a test”.\textsuperscript{33} The smallest detectable change (SDC) is defined as the change in the instrument’s score beyond measurement error.\textsuperscript{34} The SDC value demonstrates the minimum required change in order to be confident that the observed change is real and not a product of measurement error.\textsuperscript{36} The SEM and SDC were calculated with following formulas: $\text{SEM} = \alpha \times 3 \times \sqrt{1-\text{ICC}}$, $\text{SDC} = 1.96 \times \sqrt{2} \times \text{SEM}$.

Excellent test-retest reliability was found for both RSS-12p total score $0.98$ (95\% CI $0.97$–$0.99$) and QoL score $0.94$ (95\% CI $0.87$–$0.97$), respectively (Table 3).

The absolute reliability measures of the SEM and the SDC for RSS-12p total score were $6.51$ (CI $95\% = \pm 12.75$) and $7.07$, respectively. For RSS-12p QoL score, the SEM and the SDC were $2.56$ (CI $95\% = \pm 5.01$) and $4.43$, respectively.

**RESULTS**

In total, 63 patients and 50 healthy volunteers participated (Table 1). This sample size was based on the criteria proposed by Terwee et al.\textsuperscript{27}

There were no missing data for individual items of the RSS-12p. The RSS-12p scores ranged between $20$ and $263$ ($\text{mean} \pm \text{SD}: 115.11 \pm 51.53$). There were no floor or ceiling effects for the RSS-12p questionnaire. Also, the QoL scores were well distributed (range $7$–$55$; $\text{mean} \pm \text{SD} = 26.04 \pm 10.13$). No patient achieved the minimum or maximum possible score for the RSS-12p and QoL.

The means and standard deviations of RSS-12p and QoL scores in the patient and healthy groups are summarized in Table 2. Discriminative validity analysis showed that the differences of RSS-12p scores for both total score and total score of QoL were statistically significant between patients and healthy individuals ($P<0.001$) (Table 2).

An excellent significant correlation was observed between the RSS-12p and the RSI scores ($r_p = 0.87$; $P<0.00$).

The measurement of the internal consistency reliability for RSS-12p (0.85) and QoL (0.72) indicated a high Cronbach’s alpha coefficients.

**DISCUSSION**

The purpose of the present study was to produce the Persian version of RSS-12 and investigate the psychometric properties in patients with LPRD. The current study showed that the RSS-12p is a valid and reliable questionnaire in assessing Persian speaking patients with LPRD. The results of current study are in line with those of the original version of RSS, the original version of RSS-12, and the $\text{K-RSS.}$\textsuperscript{20,22}

There is no other language version of RSS-12 available to compare our findings.

**Acceptability**

The process of translation and cultural adaptation of RSS into Persian language conducted without any problems. The participants completed the RSS-12p easily that indicates all participants understood the questions and responded to them with no problems. The patients with LPRD reported the RSS-12p questions were clear and relevant for LPRD which demonstrates acceptability, face, and content validity of the RSS-12p in line with the original version.\textsuperscript{22}

**Floor and ceiling effects**

Floor or ceiling effects occur when patients score at the extremes on a questionnaire. It means that the target tool is not responsive to change and cannot be able to show any clinical worsening or improving conditions for patients.\textsuperscript{35} In this study, absence of floor or ceiling effects confirms the content validity of the RSS-12p and hence the RSS-12p may be sensitive to detect changes clinically in patients with LPRD.\textsuperscript{36} The floor and ceiling effects were not reported for the original RSI,\textsuperscript{7} the original RSS,\textsuperscript{20} and the original RSS-12.\textsuperscript{22}

**TABLE 1. Characteristics of the Participants (n=113)**

<table>
<thead>
<tr>
<th>Group</th>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients (n=63) Male=40 Female=23</td>
<td>Age (year)</td>
<td>39.26</td>
<td>9.79</td>
<td>20</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Education (year)</td>
<td>12.22</td>
<td>3.68</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Duration of disease (Month)</td>
<td>11.06</td>
<td>9.31</td>
<td>2</td>
<td>48</td>
</tr>
<tr>
<td>Healthy subjects (n=50) Male=31</td>
<td>Age (year)</td>
<td>37.24</td>
<td>10.28</td>
<td>20</td>
<td>58</td>
</tr>
<tr>
<td>Female=19</td>
<td>Education (year)</td>
<td>15.74</td>
<td>3.95</td>
<td>8</td>
<td>23</td>
</tr>
</tbody>
</table>
Discriminant validity
The RSS-12 scores and QoL scores were significantly higher in the patients with LPRD in comparison to healthy individuals; this indicates the discriminant validity of the RSS-12p. Discriminant validity of the RSS-12p is in agreement with the original RSS-12 in diagnosing patients with LPRD with those without.22 For the original RSI, similar to our study, the mean of RSI score of asymptomatic individuals was significantly less than that of persons with LPRD.7

Construct validity
Construct validity was tested by examining the correlation between the RSS-12p and RSI. There was a significant correlation between the RSS-12p total score and the RSI total score demonstrating the construct convergent validity of the RSS-12p in agreement with the original RSS-12 and the Korean version of RSS.21 The total score of the K-RSS was correlated with the RSI score ($r = 0.90$) and was a little better than correlation which we found between the RSS-12p and the RSI ($r_p = 0.87$).21 Furthermore, the previous studies found significant positive correlations between the original RSS and the original RSS-12 with the RSI ($r = 0.84$ and $r_p = 0.83$, respectively) and also with the voice handicap index (VHI) ($r = 0.57$ and $r_p = 0.49$, respectively).20,22 Although the correlations between the RSS and VHI were significant, the values were lower than the correlation found between the RSS and RSI. The possible reason for the low correlation might be that both RSS and RSI are self-assessment questionnaires developed for patients with LPRD while VHI is a questionnaire developed to assess the quality of life related to voice.

Internal consistency reliability
The internal consistency analysis yielded a Cronbach’s alpha = 0.85 for RSS-12p score and Cronbach’s alpha = 0.72 for QoL score which are larger than cut-off value of 0.70 indicating high reliability. The RSS-12p’s internal consistency was better than the original RSS-12 ($\alpha = 0.74$) while less than the original RSS ($\alpha = 0.97$) and the Korean version ($\alpha = 0.89$).20,22 The internal consistency reliability has not been calculated for the original RSI.7

Test-retest reliability
The results of test-retest reliability showed that the RSS-12p has a high repeatability and stability over time. The test-retest reliability for the RSS-12p ($ICC_{agreement} = 0.98$) was higher than that of the original RSS ($r = 0.92$), the original RSS-12 ($r = 0.96$), and the K-RSS ($r = 0.88$).20,22 In the original RSI, test-retest reliability was high ($r = 0.81$) but not higher than RSS-12p.7
The authors encountered with some limitations in the process of diagnosis of patients with LPRD. Our first limitation was the insufficient number of LPRD patients. Although there are some disadvantages of the HEMII-pH test, it has been considered as a reliable standard for the diagnosis of LPRD patients. However, due to the high price and the lack of import of the HEMII-pH test equipment, the diagnostic method is not commonly taken in the clinical or research settings in our country. Consequently, the diagnosis of LPRD is performed based on self-reporting and laryngoscopy examination, which are more accessible and have a lower price compared to the HEMII-pH test. In the present study, the test-retest reliability of the RSS-12p was provided by an otolaryngologist and three SLPs based on the self-reported symptoms and laryngoscopic signs by using the RSI and the RFS questionnaires according to the literature and the suggestion taken from the developer of the original version of RSS-12. To exclude any other pathological conditions, any possible pathologies with similar signs and symptoms were assessed and excluded by an initial taking history. Next limitation was related to the number of participants in the process of studying test-retest reliability. Data gathering was conducted during 19-Covid epidemic. Regarding to epidemic conditions, the authors could not complete the questionnaire for the second time in all patients. Therefore, with considering central limited theorem, the test-retest data were gathered only from 31 patients. Moreover, the responsiveness to change was not evaluated in the present study. Therefore, a further study of test-retest reliability and responsiveness of RSS-12p with a larger sample of patients is imperative. The RSS-12p questionnaire can be a useful tool in the research and clinical diagnosis of Persian speakers with LPRD. Responsiveness of the RSS-12p is warranted for future investigation.


