Validation of the Restless Legs Syndrome Quality of Life Questionnaire

Linda Abetz, MA,1 Susan M. Vallow, BSc, MA, MBA,2 Jeff Kirsch, PhD, MRPhS,3 Richard P. Allen, PhD,4 Tinna Washburn, BA,4 Christopher J. Earley, MD, PhD4
1Mapi Values Ltd., Macclesfield, Cheshire, UK; 2Janssen Pharmaceutica, Inc., Titusville, NJ, USA; 3GlaxoSmithKline Pharmaceuticals, Harlow, Essex, UK; 4Johns Hopkins University, Baltimore, MD, USA

ABSTRACT

Objectives: The Restless Legs Syndrome Quality of Life questionnaire (RLSQoL) assesses the impact of RLS on daily life, emotional well-being, social life, and work life. This study investigates its validity and reliability.

Methods: The RLSQoL was tested in 85 American adults with primary RLS. Patients were also asked to rate symptom severity with the International Restless Legs Scale (patient-reported version) and report on changes in symptoms over the 2-week period.

Results: The RLSQoL summary scale score (range: 0–100) demonstrated acceptable internal consistency reliability (Cronbach’s alpha = 0.92) and test–retest reliability (intraclass correlation coefficient = 0.84). All items indicated acceptable item-convergent validity. The RLSQoL distinguished between groups with mild, moderate, and severe symptoms (F = 52.22, P < 0.0001). It demonstrated preliminary responsiveness to changes in RLS status over 2 weeks (effect size: improvement, 0.25; deterioration, −0.32), indicating moderate scale changes consistent with the small clinical change over this time.

Conclusions: These findings support the conceptual framework of the RLSQoL. It is a valid and reliable measure of the impact of RLS on QoL and is responsive to short-term changes in symptom severity. The RLSQoL appears to be an appropriate tool for trial-based assessments of treatments for RLS.

Keywords: health status, quality of life, questionnaire, restless legs syndrome.

Introduction

Restless legs syndrome (RLS) is a movement disorder in which a person experiences a strong urge to move the legs or other extremities while at rest; symptoms are temporarily at least partly relieved by movement. Symptoms have a strong circadian pattern: they are worse in the evening and at night and are often diminished in the morning at the end of the sleep period. The urge to move is usually accompanied by an unpleasant sensation in the affected limb; the sensation may be described as creeping, crawling, tingling and pulling or painful, and commonly affects sleep [1,2].

Idiopathic RLS can begin at any time of life, even in early childhood, though the symptoms often become more frequent, noticeable, and severe with age [3,4]. When the condition starts early in life (before the age of 45 years), there is usually an insidiously slow progression in symptom severity, but when it starts later in life (over the age of 45 years), the symptoms tend to progress rapidly before reaching a plateau. Therefore, the most severely affected individuals are middle-aged and elderly adults, though young adults are also very distressed by the condition [5]. Studies of prevalence suggest that there may be a significant fraction of older individuals with symptoms of RLS, and the prevalence is somewhat greater in women than in men [6]. RLS frequently occurs for the first time in pregnancy or is exacerbated by it. It is also associated with iron deficiency and end-stage renal disease [6,7].

RLS suffers commonly report loss of sleep, with more severely affected individuals sleeping no more than 4 or 5 hours every night [5] and therefore suffering complications in daily functioning, including problems with concentrating. Sufferers also report problems with functioning in sedentary situations, particularly in physically constraining places, and also in the evenings, when symptoms are
usually exacerbated. As a result of these problems, sufferers may have difficulties with their jobs, social life, and recreational activities. Very little information is available on the impact of RLS on patient-reported quality of life (QoL), aside from limited studies indicating that it is associated with depression [8,9].

The primary morbidities of RLS involve sleep loss, extreme discomfort, and disruption of normal life activities [1]. QoL is therefore reduced and a primary goal of treatment should be its restoration. Thus, for this syndrome, assessing QoL becomes critical for evaluating the clinical significance of the disorder and treatment benefits. Despite the central role of QoL in RLS, the development of a method for its assessment has not previously been undertaken. Reductions in patient’s QoL result partly from the distinctive symptoms of this disorder which result in: chronic sleep loss, disruptions to circadian pattern, and disruptions to sedentary activities. Thus physical mobility or functioning would be less likely to be impacted than sleep or daily activities. The distinctive and somewhat unique disruption of QoL begs the development of a responsive RLS-specific scale. More general scales will probably detect the disruption of QoL by RLS, but may also fail to be responsive to the range of life disruptions and to the remarkable benefits of treatment reported by patients [10]. The lack of an RLS-specific QoL scale at the time of this study reflected the nascent development of treatment evaluation for this disorder. The development of the Restless Legs Syndrome Quality of Life questionnaire (RLSQoL; see Appendix) was undertaken to address the existing lack of a disease-specific QoL instrument for this condition. Interest in this issue has since been emphasized by the separate development of other disease-specific QoL instruments [11–13].

Any new instrument must undergo psychometric evaluation to ensure its reliability and validity in the patient population. The process determines whether or not the instrument measures the factors identified as clinically significant, and whether it can be used as planned. The aim of this validation process is to collect convincing evidence that the instrument utilizes the intended constructs, with which the measurements reflect the QoL of the patients and that repeated administrations detect changes in QoL.

The specific objectives of this study were to assess the validity and reliability of the recently developed RLSQoL and its responsiveness to symptom change over short periods of time.

**Methods**

**Participants**

Adults over 21 years old with primary RLS who were attending a large specialty practice focusing on RLS and sleep medicine in the United States were invited to participate. Diagnoses were made by sleep-medicine specialists; each had been certified by the American Board of Sleep Medicine and had extensive experience in treating RLS. Exclusion criteria included secondary RLS due to pregnancy, dialysis, iron deficiency, anaemia, brain stem stroke, and neuropathy.

**Measures**

RLSQoL items were developed with the help of expert clinicians and patients with RLS. An initial 19-item tool was developed based on the expertise of four clinical experts specialized in the field of RLS. Four revisions of the questionnaire were made based on consultation with these clinical experts who informally solicited opinions from RLS patients. After consultation with the experts, 13 of the 19 items were selected to be included in a pilot study of RLS patients identified in a primary care practice who provided a convenience sample of well-diagnosed RLS patients with a wide range of severity. In this study, 185 patients completed the preliminary QoL questionnaire. Ten of the questions were answered by 98% of the patients, two by at least 94%, and one related to the work impact of RLS by 89%. The answers covered the full range of the response levels for all but one item, which had no answers for the highest level (relating to disruption of evening activities, i.e., all of the time). Opinions about the questionnaire were solicited as part of unstructured interviews with 10 of these RLS patients. Based on the information from this study, three items were added to cover more general daily activities and evening social activities. In addition, to reduce the failures to answer the item covering work, a question about work status was added to serve as a qualifier for completing the work-related question. Finally, one item with an exact number answer was supplemented by an additional question involving a subjective estimate of the frequency with which RLS made it difficult to work a full day. This facilitated completion of the quantitative responses. No items were deleted at this stage given the shortness of the questionnaire and the adequacy of the responses. In addition, the wording of the items was adjusted further to reduce some confusion reported by the patients. The final 18-item RLSQoL assesses how RLS...
Validation of the RLSQoL Questionnaire

Procedure

The psychometric properties of the RLSQoL were assessed during an observational, longitudinal 2-week study conducted at a large specialty practice focusing on RLS and sleep medicine in the United States. Those patients who agreed to participate (see Participants section in Methods) signed an Internal Review Board-Approved informed consent statement and were then sent the battery of questionnaires twice by mail over a 2-week period. Patients returned the questionnaires in prepaid postage envelopes. Around the time of the second assessment, a trained clinical investigator telephoned all patients who agreed to participate to conduct an interview about changes in symptoms during the 2-week period. The protocol was reviewed and approved by the Human Subjects Institutional Review Board of the Johns Hopkins School of Medicine, Baltimore, MD, USA.

Psychometric Analysis

All data processing and analyses were performed using Statistical Analysis System (SAS) software (version 8.02). The item-scaling tests were performed using Multitrait Analysis Program for Windows (version 1.0) [17].

The following statistical tests were used: Kruskal–Wallis and ANOVA tests when comparing three groups of patients or more; Mann–Whitney, Wilcoxon and t-tests when comparing two groups of patients; Wilcoxon signed-rank test and paired t-test for paired test comparing a change to 0.

Appropriate statistics, such as Cronbach’s alpha, Pearson, Spearman, and intraclass correlation coefficients (ICCs), were calculated for specific analyses. These are described in the relevant sections below. For all the tests, a significance level of 0.05 was used, unless otherwise indicated.

Construct validity. Exploratory factor analyses through Principle Component Analysis (PCA) with Varimax rotation were used to assess the items with Likert responses in the questionnaire, as we were trying to determine the scale structure of this new questionnaire. Exploratory factor analysis may be conducted if the sample size is at least equivalent to five patients per item included in the analysis [18]. Twelve Likert items were to be included in the analysis; thus, a minimum of 60 people were required for the analysis. If more than 40% of the variance was accounted for by the first unrotated factor, a summary scale score may be calculated.

Item-scaling tests. Given the adequate but relatively small sample, the correlation between each item and the scale, corrected for overlap, was also assessed to ascertain the item-convergent validity of the scale (item-scale correlation ≥ 0.4) [19,20].

impacts patients’ daily activity, morning and evening activity, concentration, sexual activity, and work over the previous 4 weeks. Lower scores indicate lower QoL. The scoring algorithm for the RLSQoL summary scale score (see Appendix) was determined based on the results of this psychometric validation study.

In addition to the RLSQoL, four other questionnaires were administered: demographics, the International Restless Legs Scale-Patient Version (IRLS-PV), the Short Form 36 Health Survey (SF-36), and a brief health status change questionnaire. Patient demographics included patients’ age, gender, work, and marital status.

The IRLS-PV assesses symptom severity and impact and was adapted from the clinician-administered version of the IRLS (IRLS-Investigator Version [IV]). The IRLS-IV contains 10 items and has demonstrated reliability and validity [14]. Because of the fact that our study utilized a mail-out, mail-back design, the IRLS-IV needed to be adapted so that patients could complete it without the aid of a clinician. The adaptation resulted in the production of the 16-item IRLS-PV to meet these objectives. A 1-week recall was used. Lower scores indicate less severe RLS [15]; scores may range from 0 to 50. The data from this study were used to validate the IRLS-PV and found to be both reliable and valid [15]. The IRLS-PV was used to assess RLS severity: a score of 10–25 was considered to be mild; a score of 26–35 was considered to be moderate, and a score of 36–50 was considered to be severe.

The SF-36 is a well-validated and reliable generic measure of health status and was included as a concurrent validity measure [16]. The SF-36 contains 36 items assessing 8 health dimensions: physical functioning, bodily pain, general health perceptions, role limitations due to physical problems, social functioning, role limitations due to emotional problems, vitality, and mental health. A 4-week recall period was used. The SF-36 was scored as per the developer’s instructions [16]. Scores range from 0 to 100 and lower scores indicate poorer health status.

The health status change questionnaire asked the patients to report their perceptions of any changes in health status on a 7-point scale over the 2-week study period. This item was used to assess the stability of the patients’ health status for test-retest reliability and to assess responsiveness.
Reliability. Internal consistency reliability. Using the Cronbach’s alpha coefficient/statistic, the internal consistency reliability was estimated to assess the extent to which individual items are consistent with each other. An alpha value of at least 0.70 has been recommended if the measure is to be considered reliable [21,22], though reliability coefficients are susceptible to the number of items within a scale. Alpha coefficients should be interpreted with greater caution in scales with fewer items.

Reproducibility (test–retest reliability). Reproducibility of scale scores over a short period of time is an important psychometric characteristic. To test this measurement, the RLSQoL was administered on two separate occasions: at baseline and 2 weeks later. This interval was considered to be sufficiently short for patients to remain stable and experience no changes in QoL, while being sufficiently long to avoid memory bias. As recommended in the literature, the ICC was used to compare the test–retest QoL assessments, and should be equal to or greater than 0.70 [21,23].

ICCs were calculated for the total sample and then for patients who reported stable RLS symptoms over the 2-week period. Patients had to completed the RLSQoL at baseline and at week 2 to be included in this analysis. ICCs were calculated for multi-item scores, as well as any single-item measures, which were not included in multi-item scales in the event that the single items are to be used in future studies.

Concurrent validity. The concurrent validity of the RLSQoL was examined by analyzing correlation levels between the RLSQoL summary scale score and the SF-36 summary scales. A correlation of greater than 0.40 was considered a sufficient criterion to determine concurrent validity for the multi-item scales and a criterion of 0.30 was acceptable for the single-item measures. It was hypothesized that the RLSQoL scale would be significantly correlated with the SF-36 mental component summary (MCS) scale because the RLSQoL focuses on the distress the patient feels due to the RLS symptoms. The physical component summary (PCS) scale focuses primarily on mobility limitations and pain, which are not included in the RLSQoL because these are not symptoms consistently reported by most RLS patients.

Known groups validity. The RLSQoL summary scale score was compared with the patients’ self-reported symptom severity based on the overall score of the IRLS-PV. Patient scores on the IRLS-PV were used to determine mild (a score of 10–25), moderate (a score of 26–35), and severe (a score of 36–50) categories for symptoms. This scoring was a departure from the original, clinician-aided, self-administered IRLS-IV severity scoring, given the additional questions that were included in the summary scale score. Three-way median splits were used to determine the severity level for the purposes of this analysis; however, these splits are data-driven and should not be considered clinical cutoffs. It was hypothesized that the more severe the RLS, the worse the QoL scores would be on the RLSQoL.

Responsiveness. QoL changes were compared with direct reports of health status change by patients. The effect sizes (ESs) may be calculated by dividing the change in mean scores from baseline to follow-up by either the SD of the scores at baseline or the SD of the change in scores between baseline and follow-up. The ES measurements recommended in the literature include [23–25]:

1. small change (ES = 0.20);
2. moderate change (ES = 0.50); and
3. large change (ES = 0.80).

The ES in this analysis was calculated as follows: the difference in mean change in score for patients showing a change over time (improvement, deterioration) from baseline to follow-up was divided by the SD of baseline scores for all patients to obtain an ES [21]. This ES was characterized as small, moderate or large following the above guidelines [23–25].

Given the short length of this study, responsiveness was assessed only in a preliminary fashion, using patients’ reports of change in RLS status in the follow-up telephone interview at week 2 as the main criterion.

Results

Response Rates and Acceptability
Of 200 surveys mailed to patients with primary RLS and seen at the clinic in the past year, 85 adults (aged over 21 years) returned baseline assessments (a 42.5% response rate). Of these 85 patients, 62 (72.9%) returned the 2-week assessments. For the RLSQoL, the average amount of missing data per patient was 0.9 ± 1.1 items. Forty-eight percent of the patients had no missing data in the RLSQoL; 25.9% had one item missing; 16.5% had two items missing; 5.9% had three items missing; 3.5% had four items missing. Because 68.2% of the patients did not work, they were asked to omit items related only to work and these items were not
Validation of the RLSQoL Questionnaire

included in the analysis of missing data. Items 5 and 6, however, which asked how often patients were late for work or first appointments of the day and how many days they were late, could mostly only be answered by the patients who were working. These two items accounted for a significant amount of missing data, with 8 (9.4%) patients not responding to item 5 and 25 (29.4%) patients not responding to item 6.

For the two sexual activity items (items 11 and 12), two patients did not respond to item 11 and six patients did not respond to item 12 (i.e., missing data). In addition, 15 (17.6%) patients ticked the “prefer not to answer” box for item 11 and 18 (21.2%) patients ticked this box for item 12; for those patients that chose the response option “prefer not answer,” their responses were not included in the calculation of missing data because they did actually respond to the question. The majority (≥75%) of patients who omitted answers or indicated they preferred not to answer the sexual items were aged over 65 years. These results were expected for those people for whom sexual activity may not be applicable or relevant.

Demographic and Clinical Characteristics of the Sample

Baseline socio-demographic data for the sample are shown in Table 1. The majority of the sample (63.5%) were women and the mean age (± SD) was 62.4 ± 14.0 years. The majority of patients (71.6%) lived with a partner. Approximately 36% of the population

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total (N = 85)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (year)</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>62.4</td>
</tr>
<tr>
<td>SD</td>
<td>14.0</td>
</tr>
<tr>
<td>Range</td>
<td>26–87</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>31 (36.5)</td>
</tr>
<tr>
<td>Female</td>
<td>54 (63.5)</td>
</tr>
<tr>
<td>Marital status, n (%)</td>
<td></td>
</tr>
<tr>
<td>Living alone</td>
<td>14 (17.3)</td>
</tr>
<tr>
<td>Living with partner</td>
<td>58 (71.6)</td>
</tr>
<tr>
<td>Other</td>
<td>9 (11.1)</td>
</tr>
<tr>
<td>Work status, n (%)</td>
<td></td>
</tr>
<tr>
<td>Full-time</td>
<td>23 (30.3)</td>
</tr>
<tr>
<td>Part-time</td>
<td>4 (5.3)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>2 (2.6)</td>
</tr>
<tr>
<td>Volunteer</td>
<td>1 (1.3)</td>
</tr>
<tr>
<td>Retired</td>
<td>33 (43.4)</td>
</tr>
<tr>
<td>Homemaker</td>
<td>5 (6.6)</td>
</tr>
<tr>
<td>Other</td>
<td>8 (10.5)</td>
</tr>
<tr>
<td>Work shift, n (%)</td>
<td></td>
</tr>
<tr>
<td>Day shift</td>
<td>26 (86.7)</td>
</tr>
<tr>
<td>Second shift</td>
<td>1 (3.3)</td>
</tr>
<tr>
<td>Other</td>
<td>3 (10.0)</td>
</tr>
</tbody>
</table>

*Patients were asked to identify the time when symptoms started/stopped; missing time periods indicate that no patients identified these periods as times when symptoms started/stopped.

RLS, restless legs syndrome.

patients were in paid employment (full- or part-time) and, of those, 86.7% worked day shifts.

The baseline clinical characteristics of the sample are summarized in Table 2. The mean age (± SD) when RLS symptoms first appeared was 36.6 ± 19.6 years. On average, daily symptoms began to occur approximately 10.5 ± 10.3 years after first noticing the feelings or movements of RLS. Within the 4 weeks before questionnaire completion, 95.0% reported experiencing RLS feelings daily. For those

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total (N = 85)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age when first noticed symptoms (year)</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>36.6</td>
</tr>
<tr>
<td>SD</td>
<td>19.6</td>
</tr>
<tr>
<td>Range</td>
<td>0–80</td>
</tr>
<tr>
<td>When symptoms began to occur daily (years after symptoms first occurred)</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>10.5</td>
</tr>
<tr>
<td>SD</td>
<td>10.3</td>
</tr>
<tr>
<td>Range</td>
<td>0–41</td>
</tr>
<tr>
<td>RLS symptoms experienced daily, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>57 (95.0)</td>
</tr>
<tr>
<td>No</td>
<td>3 (5.0)</td>
</tr>
<tr>
<td>RLS symptoms experienced in the last 4 weeks, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>57 (79.2)</td>
</tr>
<tr>
<td>No</td>
<td>15 (20.8)</td>
</tr>
<tr>
<td>Number of days symptoms occurred per month</td>
<td></td>
</tr>
<tr>
<td>(for those without daily symptoms and responding to question)</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>7.5</td>
</tr>
<tr>
<td>SD</td>
<td>6.3</td>
</tr>
<tr>
<td>Range</td>
<td>0–20</td>
</tr>
<tr>
<td>Missing data</td>
<td>59</td>
</tr>
<tr>
<td>Time when RLS symptoms started, n (%)</td>
<td></td>
</tr>
<tr>
<td>11 AM–noon</td>
<td>9 (13.2)</td>
</tr>
<tr>
<td>1 PM–3:30 PM</td>
<td>15 (22.1)</td>
</tr>
<tr>
<td>4 PM–5:30 PM</td>
<td>15 (22.1)</td>
</tr>
<tr>
<td>6 PM–7 PM</td>
<td>16 (23.5)</td>
</tr>
<tr>
<td>8 PM–9 PM</td>
<td>9 (13.3)</td>
</tr>
<tr>
<td>2 AM–6 AM</td>
<td>4 (5.9)</td>
</tr>
<tr>
<td>Time when RLS symptoms stopped, n (%)</td>
<td></td>
</tr>
<tr>
<td>Midnight–5 am</td>
<td>21 (41.2)</td>
</tr>
<tr>
<td>6 AM–noon</td>
<td>11 (21.6)</td>
</tr>
<tr>
<td>1 PM–2 PM</td>
<td>10 (19.6)</td>
</tr>
<tr>
<td>9 PM–11 PM</td>
<td>9 (17.6)</td>
</tr>
<tr>
<td>Number of hours per day with RLS symptoms, n (%)</td>
<td></td>
</tr>
<tr>
<td>None of the time</td>
<td>7 (8.3)</td>
</tr>
<tr>
<td>&lt; 1 hours</td>
<td>8 (9.5)</td>
</tr>
<tr>
<td>1–2 hours</td>
<td>20 (23.8)</td>
</tr>
<tr>
<td>≥ 3 hours</td>
<td>41 (48.8)</td>
</tr>
<tr>
<td>≥ 9 hours</td>
<td>8 (9.5)</td>
</tr>
<tr>
<td>Patients’ general health perceptions, n (%)</td>
<td></td>
</tr>
<tr>
<td>Excellent</td>
<td>5 (6.0)</td>
</tr>
<tr>
<td>Very good</td>
<td>24 (28.6)</td>
</tr>
<tr>
<td>Good</td>
<td>33 (39.3)</td>
</tr>
<tr>
<td>Fair</td>
<td>17 (20.2)</td>
</tr>
<tr>
<td>Poor</td>
<td>5 (6.0)</td>
</tr>
<tr>
<td>Medication use, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>73 (92.4)</td>
</tr>
<tr>
<td>No</td>
<td>6 (7.6)</td>
</tr>
</tbody>
</table>
who did not experience RLS feelings daily, RLS symptoms were experienced on an average of 7.5 ± 6.3 days per month. Twenty-nine (34.6%) patients reported either excellent or very good health, 33 (39.3%) good health, and 22 (26.2%) either fair or poor health. Seventy-three (92.4%) patients reported taking treatment for their RLS, and six (7.6%) reported taking no medication.

Psychometric Results

Construct validity. The factor analysis (PCA) was performed on the items with Likert responses in the RLSQoL. Three factors were identified (daily activities, morning activities, and sexual activities).

The eigenvalue was 6.1, with a cumulative variance of 51% in the first unrotated factor, indicating that the RLSQoL can be calculated as a summary scale score with the possibility of reporting the three factors separately. The eigenvalues and cumulative variances for the second and third factors were 1.54 (13%) and 1.01 (8%), respectively. However, given that the first factor was so large and that some clinically relevant items loaded on two factors, we decided to focus our analysis and interpretation on the summary score.

Based on the missing data and factor analysis results, it was decided that the RLSQoL summary scale would include items 1–5, 7–10, and 13. Given the large number of patients who were found to be not currently working, the work-specific items were excluded (items 14–18) from the summary scale score. In addition, item 6, an item related to work activities, was excluded from the summary scale because of the high rates of missing data. Interest and disturbance in sexual activity (items 11 and 12) were excluded because many patients preferred not to answer these questions. If these sexual items had been included in the summary scale score, it would have artificially inflated the potential for missing data (if a “prefer not to respond” answer was considered missing) or artificially inflated the score (if a “prefer not to respond” answer was considered as “no limitations”). Nevertheless, it was decided that each item excluded from the summary score may still provide useful information in larger scale studies. For this reason, the items excluded from the summary score were still retained as part of the questionnaire as a whole, until further analyses of their usefulness in larger studies, with a broader population of RLS sufferers, including sexually active or working patients, could be assessed.

Item-scaling tests. For the RLSQoL summary scale, all items met or exceeded the test for item-convergent validity (r ≥ 0.4). Item-scale correlations ranged from 0.3 to 0.9.

For the RLSQoL summary scale, 1.2% of the patients reported the lowest or worst possible scores and 3.5% reported the highest or best QoL scores. Thus, the floor and ceiling effects were minimal and acceptable for the summary scale score (< 5%).

Reliability

1. Internal consistency. For the RLSQoL summary scale, the Cronbach’s alpha reliability coefficient was 0.92, exceeding the minimum criterion of 0.70.

2. Test–retest reliability. The mean difference between baseline and week 2 was −0.1 (SD = 15.7) for the total sample. This difference was minimal and not statistically significant (t = −0.05, P = 0.96). Similarly, in patients with stable symptoms (n = 33), as reported by the patients in the health status change questionnaire, the mean difference between baseline and week 2 was −0.1 (SD = 15.0). This difference was also minimal and not statistically significant (t = −0.04, P = 0.97). The ICCs were above the minimum standard of 0.70 for test–retest reliability for the RLSQoL summary scale (0.79) in the total sample and for patients with stable symptoms (0.84).

ICCs for those items not included in the summary scale score met or exceeded the tests for reproducibility (ICC range: 0.7–0.97), with the exception of items 15 (ICC = 0.69) and 18 (ICC = 0.5). The former is just below the prespecified test limit criterion and the latter item may be expected to vary over time as it relates to a potentially variable measure (i.e., hours of work cut back).

Concurrent validity. The RLSQoL summary scale score was correlated most highly with the SF-36 MCS scale, indicating that the better the QoL, as rated by the RLSQoL, the better the mental health status was, as rated by the SF-36 (r = 0.5, P ≤ 0.0001). Correlations between the RLSQoL summary scale and the SF-36 PCS scale were not significant (r = 0.1, P = 0.3).

Of those items not included in the RLSQoL summary scale (items 6, 11, 12, 14–18), only item 17 (“on average, how many hours per day did you work?”) was significantly correlated to the SF-36 PCS scale (r = 0.4, P = 0.04). Five items were significantly correlated to the SF-36 MCS scale: items 6 (r = 0.5, P = 0.0002), item 11 (r = 0.3, P = 0.003),
Validation of the RLSQoL Questionnaire

item 15 ($r = 0.4, P = 0.0001$), item 16 ($r = 0.5, P = 0.004$), and item 18 ($r = 0.6, P = 0.0002$).

Known groups validity. Figure 1 compares RLSQoL summary scale scores with the patients’ reports of RLS severity, based on the patients’ scores for the IRLS-PV. The RLSQoL summary scale score was able to distinguish between patients whose symptoms were mild, moderate, or severe, indicating the known groups validity of the RLSQoL ($F = 52.22, P < 0.0001$).

Responsiveness. For the RLSQoL summary scale, preliminary responsiveness was measured by assessing the distribution of scores at baseline and at week 2, and then assessing the mean difference between the two scores for patients whose symptoms were worse, stable, or improved, as defined by the coadministered health status change questionnaire item (Fig. 2). Patients who reported that their RLS symptoms had worsened over the 2-week period had slightly higher baseline QoL scores than patients who reported stability or improvement in RLS symptoms over the 2-week period. At the 2-week assessment, however, QoL scores were worse for those patients whose symptoms had worsened over the 2-week period compared with those patients whose symptoms had improved or remained stable. Indeed, though the sample sizes were too small to detect statistically significant differences, the initial results were promising, as the trends were in the correct direction. That is, patients reporting worse symptoms at the 2-week assessment reported poorer QoL than those patients who had stable or improved symptoms. Conversely, improved patients reported better QoL scores.

When examining mean difference scores, trends were in a similar direction, with patients whose symptoms had become worse demonstrating a decrease in scores (mean change $= -7.92$, SD $= 9.2$, indicating worse QoL), patients with stable symptoms demonstrating similar scores (mean change $= -0.09$, SD $= 2.6$, indicating fairly stable QoL), and patients with symptoms that improved reporting an increased QoL score (mean change $= 6.18$, SD $= 4.8$, indicating better QoL; Fig. 2). Finally, the ESs for both the improved and worsened groups indicated a small-to-moderate change (0.25 and $-0.32$, respectively) (Fig. 2). Despite the small sample sizes within each group, these results demonstrated preliminary responsiveness for the RLSQoL scale.

This test was also performed for the single-item assessments not included in the summary scale score. These results also suggested small-to-moderate ESs (data available from authors upon request).

Discussion

Establishing the psychometric properties of a QoL measure is an essential part of its development. If such measures are to be useful tools in both the clinical and research settings, they must have good reliability and validity, and be responsive to change. The work reported here confirms the psychometric integrity of the RLSQoL, a newly developed patient-reported measure for assessing QoL specific to RLS.

Based on the results of this evaluation, scaling assumptions are satisfactorily met for the summary scale of the RLSQoL based on 10 of the 18 items. While the RLSQoL proved to be psychometrically robust, further analyses of potential subscales could be examined in a larger study. The results provide evidence of the psychometric properties of the RLSQoL within the RLS population studied, and at least support its use in RLS patients living in the USA. Its use in other populations needs to be determined by further empiric investigation. Findings from this study offer support for the conceptual framework of the instrument and underscore the
value of measures that evaluate relevant issues surrounding the QoL of RLS sufferers.

In this analysis, we have focused primarily on the multi-item summary scale score of the RLSQoL. The sexual activity items were consistently problematic; however, the variance in response would suggest that it is useful to include these questions in the final version of the questionnaire, particularly for those patients for whom the topic is relevant and applicable. It may also be useful to assess each item to ascertain which distinguish best among RLS severity levels (e.g., using Rasch analysis and structural equation modeling, though a much larger sample size would be needed for such an analysis). Given the extensive number of tests performed on the data and the relatively small sample size, it was most appropriate to use the multi-item summary scale score to initially examine the psychometric properties of the RLSQoL. In this way, maximum power is achieved by using the minimal number of variables. Nevertheless, test–retest reliability and concurrent validity results for the single items not included in the summary scale have been provided in the event that these items may prove useful in a more heterogenous population. Further testing of the single-item measures, in terms of responsiveness, will be carried out at a later date on a larger sample.

All of the items included in the RLSQoL summary scale score were found to be psychometrically robust. The RLSQoL summary scale score had high item-convergent validity, acceptable levels of internal consistency reliability, and all had negligible floor and ceiling effects. In addition, test–retest reliability, concurrent validity, known group's validity, and preliminary responsiveness were demonstrated for the questionnaire.

The RLSQoL was more highly correlated with the SF-36 MCS score than the PCS score; this might have been expected because the PCS scale focuses primarily on mobility limitations and pain, which are not included in the RLSQoL. In addition, in the RLSQoL, some items focused on symptom-related distress, which could explain this finding.

A clear pattern emerged with regard to the known group's validity of the RLSQoL. The RLSQoL was able to distinguish between patients whose symptoms varied in severity based on the IRLS-PV scores. The degree of differences between the three severity groups was quite substantial, with an average of 20 points or more. However, it is important to note that the IRLS-PV, while assessing RLS severity, also includes some items that assess the impact of symptoms. Thus, the symptom-impact items of the IRLS-PV may contribute substantially to the differences between scores.

The fact that the RLSQoL was able to distinguish not only between severity groups but also between those patients who reported improved and worsened symptoms over a 2-week period, suggests that differences may be assessed using this measure in long-term treatment trials. In considering our findings, it is important to bear in mind that the RLSQoL was developed as a longitudinal measure for use in clinical trials. We would argue that patients’ responses to these scales are likely to change over time depending on the clinical course of their RLS problems and, potentially, depending on their age. In this preliminary psychometric validation, no extensive longitudinal data were available, though the preliminary responsiveness testing over the 2-week period was extremely encouraging. Further testing in a controlled study of intervention methods is warranted.

Based on this initial validation of the RLSQoL, we recommend that the final version be used in conjunction with continued research. We propose concerted efforts be made in five complementary areas:

1. evaluating the potential for subscales from the RLSQoL;
2. evaluating the responsiveness of the RLSQoL scale and item scores to changes over time in larger sample sizes over a longer period of time in intervention studies;
3. using a causal indicator model to assess appropriate constructs in heterogeneous populations with this chronic condition;
4. assessing the psychometric properties of the RLSQoL in foreign cultures; and
5. incorporating use of the RLSQoL into routine clinical assessments and trials to better understand the meaning and interpretation of individual patient scores.

Addressing these issues represents a tremendous challenge and will require a timely, well-orchestrated effort. Potentially, however, the RLSQoL could lead to an appreciably greater understanding of RLS, its diagnosis, and its treatment.

Source of financial support: This study was supported in part by NIH grant 1R01 AG-16362 and by GlaxoSmithKline.

References

Validation of the RLSQoL Questionnaire

4 Lavigne GJ, Montplaisir JY. Restless legs syndrome and sleep bruxism: prevalence and association among Canadians. Sleep 1994;17:739–43.
Appendix: RLS Quality of Life Questionnaire*

The following are some questions on how your Restless Legs Syndrome might affect your quality of life. Answer each of the items below in relation to your life experience in the past 4 weeks. Please mark only one answer for each question.

In the past four weeks:

1. How distressing to you were your restless legs?
   - Not at all
   - A little
   - Some
   - Quite a bit
   - A lot

2. How often in the past 4 weeks did your restless legs disrupt your routine evening activities?
   - Never
   - A few times
   - Sometimes
   - Most of the time
   - All the time

3. How often in the past 4 weeks did restless legs keep you from attending your evening social activities?
   - Never
   - A few times
   - Sometimes
   - Most of the time
   - All the time

4. In the past 4 weeks how much trouble did you have getting up in the morning due to restless legs?
   - None
   - A little
   - Some
   - Quite a bit
   - A lot

5. In the past 4 weeks how often were you late for work or your first appointments of the day due to restless legs?
   - Never
   - A few times
   - Sometimes
   - Most of the time
   - All the time

6. How many days in the past 4 weeks were you late for work or your first appointments of the day due to restless legs?
   - Write in number of days:

7. How often in the past 4 weeks did you have trouble concentrating in the afternoon?
   - Never
   - A few times
   - Sometimes
   - Most of the time
   - All the time

8. How often in the past 4 weeks did you have trouble concentrating in the evening?
   - Never
   - A few times
   - Sometimes
   - Most of the time
   - All the time

9. In the past 4 weeks how much was your ability to make good decisions affected by sleep problems?
   - None
   - A little
   - Some
   - Quite a bit
   - A lot

10. How often in the past 4 weeks would you have avoided traveling when the trip would have lasted more than two hours?
    - Never
    - A few times
    - Sometimes
    - Most of the time
    - All the time

11. In the past 4 weeks how much interest did you have in sexual activity?
    - None
    - A little
    - Some
    - Quite a bit
    - A lot
    - Prefer not to answer

12. How much did restless legs disturb or reduce your sexual activities?
    - None
    - A little
    - Some
    - Quite a bit
    - A lot
    - Prefer not to answer

13. In the past 4 weeks how much did your restless legs disturb your ability to carry out your daily activities, for example carrying out a satisfactory family, home, social, school or work life?
    - Not at all
    - A little
    - Some
    - Quite a bit
    - A lot

14. Do you currently work full or part time (paid work, unpaid or volunteer)?
    (mark one box)
    - YES If Yes please answer questions #15 through #18
    - NO, because of my RLS – Please go to the next page
    - NO, due to other reasons – Please go to the next page

15. How often did restless legs make it difficult for you to work a full day in the past 4 weeks?
    - Never
    - A few times
    - Sometimes
    - Most of the time
    - All the time

16. How many days in the past 4 weeks did you work less than you would like due to restless legs?
    - Write in number of days:

17. On the average, how many hours did you work in the past 4 weeks?
    - Write in number of hours per day:

18. On days you worked less than you would like, on average about how many hours less did you work due to your restless legs?
    - Write in number of hours per day:

*Copyright 2002, Richard Allen; reproduced with kind permission.
**Scoring**

A summary score can be calculated for the RLS quality of life questionnaire based on the following items: 1–5, 7–10 and 13. All items must be recoded such that 1 equals most severe and 5 equals least severe, so that lower scores indicate worse quality of life. The score is then transformed to a 0–100 score using the following algorithm:

\[(\text{Actual raw score} - \text{lowest possible raw score}) / \text{Possible raw score range}) \times 100.\]

If more than two items are missing from the summary scale, the summary scale score cannot be calculated and is set to missing. If one or two items from the summary scale are missing, then a person-specific estimate is substituted for that missing item. This person-specific estimate is the average score, across the completed items in the summary scale, for that respondent.

Items 6 and 16–18 are scored as continuous variables, as written by the patient. For items 6 and 16, the minimum number of days is 0 and the maximum number of days is 28. For items 17 and 18, the minimum number is 0 hours and the maximum number is 24 hours. If the response to one of these items is missing or out of range, than that item is set to missing. Items 14–18 are work-related items, thus if patients reply “2” or “3” to item 14, they are not expected to reply to items 15–18. Thus, the missing data rates for items 15–18 will be artificially inflated.

Items 11, 12 and 15 should be scored as categorical variables. Finally, item 14 can also be treated as a categorical variable as follows: “yes” = 1; “no, because of my RLS” = 2; “no because of other reasons” = 3. If a response to one of these items is missing, then no score can be calculated for that item.