Construct Validity and Test-Retest Reliability of the Questionnaire Rising and Sitting Down in Lower-Limb Amputees

Fred A. de Laat, MD, Gerardus M. Rommers, MD, PhD, Jan H. Geertzen, MD, PhD, Leo D. Roorda, MD, PT, PhD


Objective: To investigate the construct validity and test-retest reliability of the Questionnaire Rising and Sitting Down (QR&S), a patient-reported measure of activity limitations in rising and sitting down, in lower-limb amputees.

Design: Cross-sectional study.

Setting: Outpatient department of a rehabilitation center.

Participants: Lower-limb amputees (N = 171; mean age ± SD, 65 ± 12y; 71% men; 83% vascular cause) participated in the study, 33 of whom also participated in the reliability study.

Interventions: Not applicable.

Main Outcome Measures: Construct validity was investigated by testing 8 hypotheses: limitations in rising and sitting down according to the QR&S would be: (1) greater in lower-limb amputees who are older, (2) independent of level of amputation, (3) greater in lower-limb amputees with a bilateral amputation, and (4) greater in lower-limb amputees who had rehabilitation treatment in a nursing home. Furthermore, limitations in rising and sitting down will be positively related to activity limitations according to (5) the Locomotor Capabilities Index (LCI), (6) the questions about rising and sitting down in the LCI, (7) the Climbing Stairs Questionnaire, and (8) the Walking Questionnaire. Construct validity was quantified with an independent t test and Pearson correlation coefficient. Test-retest reliability was assessed with a 3-week interval and quantified with the intraclass correlation coefficient (ICC), standard error of measurement, and smallest detectable difference (SDD).

Results: Construct validity (7 of 8 null hypotheses not rejected) and test-retest reliability were good (ICC = .84; 95% confidence interval, .65–.93; standard error of the measurement = 6.7%; SDD = 18.6%).

Conclusions: The QR&S has good construct validity and good test-retest reliability in lower-limb amputees.

Key Words: Amputation; Disability evaluation; Questionnaires; Rehabilitation.

© 2011 by the American Congress of Rehabilitation Medicine

An important goal in the rehabilitation of lower-limb amputees is that they learn to walk with a prosthesis and regain functional independency. For mobility in walking, lower-limb amputees have to be able to rise, stand and maintain balance, initiate gait, walk, terminate gait, and to sit down. Therefore, rising and sitting down are important aspects of mobility in lower-limb amputees, and the assessment of limitations in rising and sitting down is of great value in prosthetic training. Studies focusing on limitations in rising and sitting down in lower-limb amputees are scarce. Moreover, these studies often use performance-based measurements to assess the limitations, and therefore, they do not provide information about the patient’s perspective of limitations in rising and sitting down. Furthermore, these performance-based measurements are often part of a test battery, and as a consequence, these studies do not report in detail on limitations in rising and sitting down. The only study that specifically addressed standing up from a chair in transfemoral amputees is a biomechanical study, but it provided no detailed information about the patient’s perspective of limitations in rising and sitting down.

Therefore, when the aim is to assess the patient’s perception of limitations in rising and sitting down, a self-report questionnaire is more appropriate than a performance test. Questions about these limitations in lower-limb amputees are mostly included in a more comprehensive questionnaire, for instance the Locomotor Capabilities Index (LCI) or the Prosthesis Evaluation Questionnaire, and therefore only assess global limitations in transfers. As far as we know, the only questionnaire that provides a detailed assessment of limitations in rising and sitting down is the Questionnaire Rising and Sitting Down (QR&S).

The QR&S is a patient-reported questionnaire that measures activity limitations in rising and sitting down. It contains 39 items formulated in behavioral terms (eg, I have [some] difficulty getting up from a high-seated chair) with dichotomous response options (yes box marked/yes box not marked). The sum score is based on the 1-parameter logistic model and standardized (range, 0–100) with higher scores indicating less limitation. The selection of items to be included in the QR&S was based on an extensive literature review, and the first draft version was subjected to the opinions of experts. The improved version was then tested in 759 patients with lower-extremity disorders (including 230 lower-limb amputees) living at

List of Abbreviations

| CI | confidence interval |
| GRCQ | Global Rating of Change Questionnaire |
| ICC | intraclass correlation coefficient |
| LCI | Locomotor Capabilities Index |
| QR&S | Questionnaire Rising and Sitting Down |
| SDD | smallest detectable difference |
METHODS

Participants

Participants were recruited between 1998 and 2008 in the Outpatient Department of the Rehabilitation Center Tolbrug, ’s-Hertogenbosch, in The Netherlands. The first group consisted of lower-limb amputees at the end of their outpatient rehabilitation treatment in this center (rehabilitation center group). These lower-limb amputees were assessed just before the start of their follow-up in the outpatient department. The second group consisted of lower-limb amputees directly after discharge from inpatient or outpatient rehabilitation treatment in nursing homes in the region of Tolbrug, ’s-Hertogenbosch (nursing home group). These lower-limb amputees were assessed at the start of their follow-up at the outpatient department of the rehabilitation center. The 2 groups together encompassed all lower-limb amputees undergoing rehabilitation treatment in this region. Only lower-limb amputees who were wearing a prosthesis after their rehabilitation treatment were selected. For the test-retest reliability study, a subgroup of lower-limb amputees who had finished their rehabilitation treatment between June 2003 and November 2004 was recruited from the rehabilitation center group. All participants provided informed consent.

Procedure

The rehabilitation center group received the first questionnaire from the therapists on their penultimate day of treatment. They were asked to fill in the questionnaire at home and to bring it with them on the last day of treatment. The nursing home group received the questionnaire during their first appointment in the rehabilitation center. They were asked to fill in the questionnaire at home, and to return the completed questionnaire by mail. The first questionnaire consisted of the QR&S, LCI,12,17 Climbing Stairs Questionnaire 15 items,18 and the Walking Questionnaire.19

For the test-retest reliability study, patients received a second questionnaire by mail 3 weeks later. This period was considered to be long enough to ensure that the participants would not remember their first responses (recall bias). They were asked to fill in the second questionnaire at home and return it by mail. Patients who returned questionnaires with missing data were contacted by telephone by an independent physician and asked to provide the missing data. This second questionnaire consisted of the QR&S and a self-constructed Global Rating of Change Questionnaire (GRCQ). The GRCQ was used to exclude patients whose limitations had changed significantly in the 3-week period after treatment. Patient instructions and items of the GRCQ can be found in Appendix 1. Participants were considered to be stable with respect to their limitations in rising and sitting down if they rated themselves between 6 and 10 on the GRCQ.

Measurements

Data on sociodemographic characteristics, diagnosis, and prosthesis prescription were extracted from medical records.

To assess construct validity, we selected the following patient-reported measurement instruments with a good conceptual framework,10,22 measuring mobility or aspects of mobility: the LCI,11,12,17 the Climbing Stairs Questionnaire,18 and the Walking Questionnaire.16

The LCI11,12,17 is a patient-reported questionnaire that measures activity limitations in climbing stairs. It consists of 15 items with dichotomous response options. The sum score is calculated by adding the scores of the 15 items. This sum score is subsequently standardized (range, 0–100), with higher scores indicating less limitation in climbing stairs. Patients can mark a 16th item if they do not climb stairs at all, due to their health, and these patients are given the minimum score. The Climbing Stairs Questionnaire showed to have: (1) good fit with the monotonicity model18 (or scalability), indicating that the items form a scale; (2) good fit with the double monotonicity model,18 indicating invariant (hierarchical) item ordering; (3) good intratest reliability, indicating good repeatability of the sum score; (4) good robustness, indicating both stability of scalability and invariant item-ordering in subgroups of patients; and (5) some differential item functioning (4 items for amputees compared with nonamputees). Furthermore the Climbing Stairs Questionnaire has good construct validity and test-retest reliability in lower-limb amputees.23

The Walking Questionnaire19 is a patient-reported questionnaire that measures activity limitations when walking inside and outside the house. It contains 35 items with dichotomous response options (yes box marked/yes box not marked). The sum score is calculated by adding the scores for the 35 items. Subsequently, the sum score is standardized (range, 0–100), with higher scores indicating less limitation in walking. Patients can mark a 36th item if they do not walk inside the house at all, and these patients are given the minimum scores. Patients can mark a 37th item if they do not walk outside the house at all due to their health, and these patients are treated as if they had marked the yes box for all the items concerning walking outside the house. The Walking Questionnaire was tested in 981 patients with lower-extremity disorders, (including 239 lower-limb amputees) who were living at home. It had: (1) good fit with the monotonicity model, (2) good fit with the double monotonicity model, (3) good intratest reliability, (4) good robustness, and (5) some differential item functioning (6 items for amputees compared with nonamputees).19

The study protocol was approved by the Research Ethics Committee of the Jeroen Bosch Hospital, ’s Hertogenbosch.

Analysis

Construct validity. Construct validity indicates the degree to which the scores on a measurement instrument are consistent with theoretically derived hypotheses (eg, with regard to internal relationships, relationships with the scores of other instruments, or differences in scores between relevant groups), based...
on the assumption that the instrument validly measures the construct to be measured. Construct validity is considered to be good if at least 75% of the hypotheses are not rejected in a study group of at least 50 participants.\textsuperscript{24} There are, as far as we know, no available scales that specifically assess rising and sitting down in lower-limb amputees. Therefore, before examining our data, we formulated 8 hypotheses based on the available literature concerning the relationship between transfers of general mobility limitations and patient-related factors in lower-limb amputees, or based on clinical experience. We hypothesized that limitations in rising and sitting down according to the QR&S would be:

1. Greater in lower-limb amputees who are older.\textsuperscript{22,25,26}
2. Equal in lower level of amputation (transfemoral or syme amputation) and higher level of amputation (transfemoral or knee disarticulation)\textsuperscript{13,27,28}
3. Greater in lower-limb amputees with a bilateral amputation than in lower-limb amputees with a unilateral amputation\textsuperscript{25,29}
4. Greater in lower-limb amputees who had rehabilitation treatment in a nursing home than in lower-limb amputees who had this treatment in an outpatient department of a rehabilitation center.\textsuperscript{29}
5. Greater in lower-limb amputees who have more limitations in locomotor capabilities, according to the LCI.\textsuperscript{22,27}
6. Greater in lower-limb amputees who have more limitations in the 3 items concerning rising and sitting down in the LCI.\textsuperscript{22,27}
7. Greater in lower-limb amputees who have more limitations in climbing stairs, according to the Climbing Stairs Questionnaire.\textsuperscript{16,23}
8. Greater in lower-limb amputees who have more limitations in walking, according to the Walking Questionnaire.\textsuperscript{19}

Hypotheses addressing relationships (hypotheses 1, 5, 6, 7, and 8) were quantified with Pearson correlation coefficient, and hypotheses addressing the presence or absence of differences were quantified with the independent t test (hypotheses 2, 3, and 4; 2-tailed \( P<.05 \)).

Test-retest reliability. Test-retest reliability refers to the reproducibility of measurements with the same instrument over time. To assess the reproducibility of the QR&S, we used the QR&S data from the first and second questionnaires of the patients who rated themselves stable on the GRQ. Reproducibility includes reliability and agreement.\textsuperscript{30} Reliability refers to how well individuals can be distinguished from each other; whereas, agreement indicates how close the repeated measurements are to the original measurements. The most frequently used reliability parameter is the intraclass correlation coefficient (ICC), which is calculated as the ratio of the variance between participants and the total variance. To estimate the test-retest reliability of the QR&S, we calculated the ICC with 95% confidence interval (CI), using a 2-way random model. Patients and measures were considered to be random effects. An ICC of at least .70 was considered to be satisfactory for group comparisons, whereas an ICC of at least .90 was considered to be satisfactory for individual comparisons.\textsuperscript{24} Agreement was quantified by the standard error of measurement, the square root of the within-subject variance, which indicates how close the scores for repeated measurements are. The smallest detectable difference (SDD) can be derived from the SEM, where SEM is defined as the standard error of measurement: \( \text{SDD}=1.96 \times \sqrt{2} \times \text{SEM} \).\textsuperscript{31} The SDD is the smallest difference in measurement that can be interpreted as a real difference between 2 measurements in an individual. Standard errors of measurement and SDDs are expressed in the units of the measurement scale. To our knowledge, there are no generally accepted criteria for satisfactory standard error of measurement and SDD values for group or individual comparisons. To visualize the agreement, we represented the data graphically in a Bland-Altman plot.\textsuperscript{32} All statistics were calculated with SPSS 15.0 for Windows.\textsuperscript{a}

RESULTS

Patient Characteristics

Of the 175 lower-limb amputees who fulfilled the selection criteria, 171 were willing to participate in the construct validity study. The lower-limb amputees who were unwilling to participate were: 2 with a transfemoral amputation, 1 with a knee disarticulation, and 1 with a transfemoral amputation, all from the rehabilitation center group. The characteristics of the 171 lower-limb amputees are presented in table 1.

For the test-retest reliability study, 35 of the 171 lower-limb amputees who participated in the construct validity study met the additional selection criteria. Of these 35 lower-limb amputees, 2 were unwilling to fill in the second questionnaire: 1 with a transfemoral amputation and 1 with a transfemoral amputation. The resulting data therefore concerned 33 lower-limb amputees, only 22 of whom considered their condition to be stable with regard to their limitations in rising and sitting down.

Construct Validity

The results of the 8 hypotheses that we tested are presented in table 2. Only hypothesis 3 (bilateral vs unilateral amputation) was rejected.

Table 1: Patient Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>Mean years ± SD 65±11</td>
</tr>
<tr>
<td></td>
<td>Range (minimum to maximum) 37–92</td>
</tr>
<tr>
<td>Sex</td>
<td>Women 50 (29)</td>
</tr>
<tr>
<td></td>
<td>Men 121 (71)</td>
</tr>
<tr>
<td>Setting</td>
<td>Rehabilitation center 154 (90)</td>
</tr>
<tr>
<td></td>
<td>Nursing home 17 (10)</td>
</tr>
<tr>
<td>Amputation etiology</td>
<td>Vascular 142 (83)</td>
</tr>
<tr>
<td></td>
<td>Infection 13 (8)</td>
</tr>
<tr>
<td></td>
<td>Traumatic 13 (8)</td>
</tr>
<tr>
<td></td>
<td>Oncologic 3 (2)</td>
</tr>
<tr>
<td>Amputation level</td>
<td>Hip disarticulation 3 (2)</td>
</tr>
<tr>
<td></td>
<td>Transfemoral 54 (32)</td>
</tr>
<tr>
<td></td>
<td>Knee disarticulation 8 (5)</td>
</tr>
<tr>
<td></td>
<td>Transtibial 93 (54)</td>
</tr>
<tr>
<td></td>
<td>Syme 1 (1)</td>
</tr>
<tr>
<td></td>
<td>Amputation unilateral 159 (93)</td>
</tr>
<tr>
<td></td>
<td>Transfemoral and transfibial 2 (1)</td>
</tr>
<tr>
<td></td>
<td>Transtibial and transfibial 7 (4)</td>
</tr>
<tr>
<td></td>
<td>Syme and transfibial 3 (2)</td>
</tr>
<tr>
<td></td>
<td>Amputation bilateral 12 (7)</td>
</tr>
</tbody>
</table>

NOTE. \( N=171 \). Values are expressed as \( n (\%) \) unless noted otherwise.
Test-Retest Reliability

Mean ± SD scores for the first and second QR&S were 42±13 and 39±19, respectively. The 3-week test-retest reliability of the QR&S was good, with an ICC of .83 (95% CI, .65–.93). The agreement of the QR&S was good with a standard error of measurement of 6.7% and an SDD of 18.6%. Agreement is presented graphically in a Bland-Altman plot (fig 1). Although overall agreement between measurements was acceptable, we found large differences for 2 lower-limb amputees in the lower range of the mean sum score.

DISCUSSION

The objective of this study was to assess the construct validity and test-retest reliability of the QR&S in lower-limb amputees. As far as we know, this is the only patient-reported questionnaire that assesses items related to rising and sitting down. The results of our study showed that the QR&S has good construct validity and good test-retest reliability for group comparisons in lower-limb amputees.

Of the 175 lower-limb amputees who fulfilled the selection criteria, only 4 were unwilling to participate in the study. With regard to the cause and level of amputation, our study was similar to lower-limb amputees in general in The Netherlands.33,34 With regard to the construct validity assessment, it was difficult to formulate hypotheses a priori, because there is hardly any literature that specifically reports on the relationship between patient-reported perceptions of limitations in rising and sitting down and clinical factors in lower-limb amputees. We found no relationship between limitations in rising and sitting down, according to the QR&S, and unilateral versus bilateral amputation (hypothesis 3), possibly due to the small number of bilateral amputees (n=12) who participated and the fact that we selected only lower-limb amputees who wore a prosthesis. Therefore, the bilateral amputees may have had fewer activity limitations.

The test-retest reliability of the QR&S was good. The SDD resulting from the found test-retest agreement was 18.6%, indicating that to detect a true difference, the difference between the 2 measurements has to be at least 19 (on a scale from 0–100). This value is quite high, but for application in a group of lower-limb amputees (eg, for research purposes), smaller differences can be detected, because then the SDD has to be divided by √n.24,30 Thus, for example, in a group of 25 lower-limb amputees, a difference of only 3.7 can be considered as a true difference in limitations in rising and sitting down.
Study Limitations

One limitation of our study is that the nonresponse rate of the lower-limb amputees treated in nursing homes is unknown, because only those lower-limb amputees who had a first follow-up appointment at the outpatient department of the rehabilitation center after their rehabilitation treatment in a nursing home were invited to participate. Nevertheless, all the lower-limb amputees who kept this appointment were willing to participate. Possibly, those who were unwilling to attend the follow-up appointment in the rehabilitation center were the worst performers in the nursing home group. So the difference between the 2 groups might have been even greater.

Another limitation of our study is that the lower-limb amputees were selected at the end of their multidisciplinary rehabilitation treatment, when only 22 of the 33 participants of the reliability study considered their condition to be stable with regard to their limitations in rising and sitting down. The reason for this was not investigated in the present study. For nonstable participants, the limitations in rising and sitting down. The reason for this was not investigated in the present study. For nonstable participants, the reason could be that the socket was fitting less well because of atrophy of the stump in the meantime, which can continue for up to 2 years after amputation. In future studies, the reliability of the QR&S should therefore be reassessed in lower-limb amputees without stump atrophy who have experience in wearing their prosthesis. Furthermore, recent standards recommend at least 50 participants for a test-retest reliability study. Therefore, we recommend that future research should replicate our study in a much larger sample.

Finally, we used only patient-reported measurement instruments to assess the construct validity of the QR&S, because the results of a performance test are not necessarily strongly related to perceived limitations. However, we recommend that the construct validity of the QR&S should be further assessed with data from biomechanic and performance-based studies of lower-limb amputees.

CONCLUSIONS

The QR&S provides a detailed assessment of patient-reported limitations in rising and sitting down, and it has good construct validity and good test-retest reliability in lower-limb amputees directly after their multidisciplinary rehabilitation treatment. Based on the results of the reliability study, the QR&S can be recommended for group comparisons of lower-limb amputees, but not for individual comparisons.

APPENDIX 1: HOW DO YOU RATE YOUR ABILITY TO RISE AND SIT DOWN NOW, COMPARED WITH THE FIRST TIME YOU FILLED IN THE QUESTIONNAIRE?

(1) Extremely good.
(2) Very much better.
(3) Much better.
(4) Better.
(5) Somewhat better.
(6) Slightly better.
(7) Almost the same, marginally better.
(8) No change.
(9) Almost the same, marginally worse.
(10) Slightly worse.
(11) Somewhat worse.
(12) Worse.
(13) Much worse.
(14) Very much worse.
(15) Extremely bad.

References


Supplier

a. SPSS Inc, 233 S Wacker Dr, 11th Fl, Chicago, IL 60606.