

Questionnaire for Persons with a Transfemoral Amputation (Q-TFA): Initial validity and reliability of a new outcome measure

Kerstin Hagberg, PT; Rickard Brånemark, MD, PhD; Olle Hägg, MD, PhD

Department of Prosthetics and Orthotics, Sahlgrenska University Hospital, Göteborg, Sweden; Department of Orthopaedics, Göteborg University, Göteborg, Sweden

Abstract—The Questionnaire for Persons with a Transfemoral Amputation (Q-TFA) is a new self-report measure developed for nonelderly transfemoral amputees using a socket- or osseointegrated prosthesis to reflect use, mobility, problems, and global health, each in a separate score (0–100). This paper describes the initial measurement properties of the Q-TFA as completed by 156 persons with a transfemoral amputation using a socket prosthesis (67% male, 92% nonvascular cases, mean age 51 years). Criterion validity was determined by associations between scores of the Q-TFA and the Short-Form 36 (SF-36)-Item Health Survey. Reliability was assessed by retest ($n = 48$) and by determination of the internal consistency. Correlations between Q-TFA and SF-36-Item Health Survey scales matched hypothesized patterns. Intraclass correlations were between 0.89 and 0.97, and measurement error ranged from 10 to 19 points. Cronbach's alpha revealed good internal consistency, with no values less than 0.7. This study shows that the Q-TFA, applied to persons using a transfemoral socket prosthesis, has adequate initial validity and reliability.

Key words: amputation, artificial limbs, lower limb, osseointegration, outcome assessment, prosthesis and implants, questionnaires, reliability and validity, walking.

INTRODUCTION

Lower-limb amputation is a permanent impairment that has been shown to lead to decreased health-related quality of life (HRQL), activity limitations, and participation restrictions [1–5]. In Sweden, as well as in the rest of the developed world, the majority of lower-limb

amputations are performed because of vascular disease, with the high mortality rate and average age typical of this population [6–8]. The ultimate goal for rehabilitation after a lower-limb amputation is to achieve the best possible quality of life for the individual, which in most cases includes mobility with a prosthetic limb. The success of prosthetic rehabilitation is influenced by both the cause and level of the amputation: vascular cases demonstrate poorer results than nonvascular cases, and transfemoral cases demonstrate poorer results than transtibial cases [8–11]. Among all persons using an artificial limb, socket-related problems and discomfort also constitute major concerns affecting quality of life and prosthetic mobility [5,12–16].

Abbreviations: HRQL = health-related quality of life, ICC = intraclass correlation coefficient, MCS = Mental Component Score, Med = median, PCS = Physical Component Score, PEQ = Prosthesis Evaluation Questionnaire, Q-TFA = Questionnaire for Persons with a Transfemoral Amputation, SD = standard deviation, SF-36 = Short-Form 36.

This material was based on work supported by Greta and Einar Askers Stiftelse and Johan Janssons Stiftelse, all of Sweden.

Address all correspondence to Kerstin Hagberg, Department of Orthopaedics, Bruna Stråket 11b, Plan 5, Sahlgrenska University Hospital, S-413 45 Göteborg, Sweden; +46-31-343-8133; fax: +46-31-342-2630; email: kerstin.hagberg@vgregion.se.

DOI: 10.1682/JRRD.2003.11.0167

A specific subgroup of individuals with amputation is the group of nonelderly persons having an amputation for reasons other than vascular disease. This group is a clear minority in the developed world, but constitutes a substantial portion of those having an amputation due to trauma or tumor [17,18]. Those who survive the trauma or malignancy have a long life expectancy [7].

In our practice, we have a specific need to monitor changes in status for nonelderly individuals having a transfemoral amputation to assess those patients who receive a direct bone-anchored amputation prosthesis through the method of osseointegration [19]. This method includes a two-stage surgery in which a titanium fixture is implanted in the residual bone in the first stage, and an abutment penetrating the skin is inserted at a second surgery. Prosthetic attachment to the residual limb is obtained by connection of the prosthetic limb directly to the abutment. Patients most suitable to be treated with a bone-anchored prosthetic limb have so far been younger or middle-aged transfemoral amputees, without vascular disease, facing severe problems in using an artificial limb suspended with a conventional socket. To optimize the selection of suitable patients and to study outcomes related to the change of prosthetic attachment, we need a useful instrument to reflect changes after treatment with a bone-anchored prosthesis.

Previously described outcome instruments targeted for individuals with lower-limb amputation have not been specifically designed to address the needs of nonelderly persons having a transfemoral amputation [20,21]. The Locomotor Capability Index, which is included in the Prosthetic Profile of the Amputee, has been shown to have a high ceiling effect and is recommended for use on individuals with lower prosthetic mobility capabilities [22]. Therefore, we concluded that development of a new tool that could be expected to be sensitive for changes in mobility and problems in use with a prosthesis for our specific subgroup of nonelderly patients was necessary. Such a tool could be used to monitor changes in status in response to, for example, new prosthetic components or a surgical revision. During the process of development of our new instrument, the Prosthesis Evaluation Questionnaire (PEQ) was published [23]. This is a valid and reliable instrument developed to "measure small differences in prosthesis function and major life domains related to prosthesis function" among persons with a lower-limb amputation. However, no study has reported on the sensitivity of the PEQ within the subgroup of nonelderly indi-

viduals having a transfemoral amputation, and there is no valid translation to Swedish.

The patient's subjective experience is the issue of most importance in measurements of health outcomes [24,25]. Self-report questionnaires are the preferred format for such measurements [26]. Ideally, such a questionnaire should be clinically relevant, consistent, and brief and should cover "those issues that are considered of particular interest to the study" [27]. Moreover, outcome questionnaires should yield reliable results if used repeatedly on the same patient with a stable condition.

The Questionnaire for persons with a Transfemoral Amputation (Q-TFA) is a targeted self-report outcome measure that reflects current prosthetic use, mobility, problems, and health. Primarily designed for nonelderly persons having a transfemoral amputation, the Q-TFA was also developed to study outcomes when changing from a conventional socket prosthesis to a bone-anchored prosthesis. Our objective in this study was to assess the initial validity and reliability of the Q-TFA in individuals using a transfemoral socket prosthesis in Sweden.

METHODS

Study Population

Subjects were selected based on the following inclusion criteria: adults aged 20 to 70 years, with a unilateral transfemoral amputation at least 2 years before the study, using a conventional socket prosthesis, and able to read and understand the Swedish language. A prosthetic user was defined as "a person who wears a prosthesis at least once a week," according to Grisé et al. [20]. Participants in the study were recruited in two stages: first, through national associations for amputees and selected prosthetic workshops and/or rehabilitation units, and second, through all four prosthetic workshops in the county of Västra Götaland.

The Human Research Ethics Committee of the Medical Faculty at Göteborg University approved the study. Eligible persons received written information regarding the study, anonymity, and notification of their right to discontinue their participation at any time, before signing the consent form.

Procedures

Two questionnaires (Q-TFA and the Short-Form 36 [SF-36]-Item Health Survey) were sent by mail, together

with questions regarding demographic characteristics and baseline information. Those who did not answer received two reminders. Participants living in the county of Västra Götaland were also asked to participate in a test-retest reliability study: 2 weeks after the questionnaires were received, we mailed the Q-TFA a second time, along with four additional questions on important changes of health condition and condition of the prosthesis to verify a stable situation since the first mailing. A 2-week interval between test and retest was chosen because this interval has been suggested to be long enough to prevent memory bias but short enough to assure a stable condition in most cases [28]. Of the 62 persons asked to answer the questionnaire a second time, 9 were excluded because of reported changes in condition that were considered clinically significant, and 5 failed to answer, leaving 48 participants for this component of the study.

Q-TFA: Development and Description

The aim of the Q-TFA is to determine current status regarding prosthetic use, prosthetic mobility, problems, and global health of persons using a transfemoral prosthesis. Experts involved in the treatment of such patients, including two orthopedic surgeons, a prosthetist, and a physical therapist, identified items selected for construction of the Q-TFA. Their clinical experience, combined with a review of existing literature and a continuous discussion during rehabilitation of patients, formed the foundation for the questionnaire design. Further, a small group of patients ($n = 4$), not included in the study, with experience using first a conventional socket prosthesis and then a bone-anchored prosthesis, were consulted in the development of the instrument.

The Q-TFA is a self-report questionnaire. The current version consists of 70 questions, and the time to complete it is approximately 20 minutes. Based on the answers to the questionnaire, we constructed a scoring system. Of the 70 questions, 54 are condensed to four separate scores: Prosthetic Use Score (Use), Prosthetic Mobility Score (Mobility), Problem Score (Problem), and Global Score (Global) (**Appendix**, available in the online version only). Each raw score is transformed to a range from 0 to 100 with the standard scoring method [27]. The remaining 16 questions, describing details connected with prosthetic use or mobility (i.e., preference to use the prosthesis or not while cooking, driving, or socializing, and details of reasons to refrain from using the prosthetic limb), are not included in any score

because of statistical redundancy or for other reasons. Those 16 questions are not analyzed in this study and are not included in the **Appendix** (online version only).

Prosthetic Use Score (2 Items)

Prosthetic use is defined as the amount of normal prosthetic wear per week. The number of days per week the prosthesis is normally worn is multiplied by the number of hours per day. A Use score of 100 indicates that the prosthesis normally is worn every day for more than 15 hours a day.

Prosthetic Mobility Score (19 Items)

Prosthetic mobility is defined as the ability and performance of the amputee to move and change and maintain postures when using the prosthesis. The score consists of three subscores, each with a range from 0 to 100: capability (12 items), use of walking aids (2 items), and walking habits (5 items). The average of these three subscores generates the total Mobility score. Capability items consist of questions on the ability to perform locomotor activities, independent of the level of difficulty to perform them. The subscores of walking aids and walking habits are estimates of prosthetic performance rather than capability. A result of 100 in the subscore of walking aids indicates that, in general, the person does not use walking aids at all. With regard to walking habits, the patient is asked to answer how often he/she had walked continuously various distances outdoors during the last 3 months. To summarize, the Mobility score consists of the average of three subscores (capability, walking aids, and walking habits). A score of 100 indicates the best possible prosthetic mobility as measured with this instrument.

Problem Score (30 Items)

Problems are defined as the extent of specific problems related to the amputation and the prosthesis and their impact on the quality of life. Each item consists of a paired question: the first asks about the extent of a specific problem during the last 4 weeks and the second about the impact on quality of life of that specific problem. Ten items concern problems regardless of prosthetic use and twenty concern problems in connection with prosthetic use. Answers are given on a 5-point Likert scale. No "not applicable" alternative is available, and if not relevant, the item could be left out and treated as missing. A minimum number of 15 paired questions must

be answered for the calculation of the total Problem score. A higher score indicates more serious problems.

Global Score (3 Items)

Global health is defined as the perception of function and problems with the current prosthesis and the perception of the current overall amputation situation. The score is a summary of three questions to which answers are given on a 5-point Likert scale. A Global score of 100 indicates the best possible overall situation as measured by this instrument.

SF-36-Item Health Survey

The SF-36-Item Health Survey is a validated generic measure of HRQL assessing the concept in eight separate scales (Physical Functioning, Role Functioning—Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role Functioning—Emotional, and Mental Health) [29]. The first four scales measure physical health, and the following four measure psychological well-being. The results can also be presented in two summary measures, the Physical Component Score (PCS) and the Mental Component Score (MCS) [30]. Each scale, as well as PCS and MCS, can take a value between 0 and 100, where a higher score indicates better health. The validated Swedish version of the instrument was used [31].

Measurement Framework

Two perspectives can be applied when one is assessing measurement properties: a psychometric approach and a clinimetric approach [27]. In psychometrics, a single construct is measured with multiple items. Assessments of internal consistency of the instrument are based on statistical analysis of the relationships between the items. In clinimetrics, the aim is to measure “clinical phenomena that are generally believed to comprise several unrelated patient characteristics or attributes” [32]. In our study, we were influenced by both approaches, consistent with the suggestion of Marx et al. that they could be complementary strategies for the development of health measurement scales [32]. Specifically, we began the ongoing process of validating the Q-TFA by determining its clinical sensibility, criterion validity, test-retest reliability, and internal consistency.

Clinical sensibility, or face validity, implies that the instrument is based on clinical knowledge and that the items included are relevant to the target population [33].

Clinical sensibility is determined qualitatively through expert opinion and more objectively by examining item response patterns.

Criterion validity is determined by an analysis of the extent to which the new tool is related to an existing “gold standard.” In the absence of a gold standard, an already validated instrument measuring the same or similar constructs could be used. For this study, we chose the SF-36-Item Health Survey to assess criterion validity of Q-TFA because of its documented capability to detect changes in health, its overall widespread use, and its use with individuals with amputation [2,12,34,35]. The following hypotheses were made:

1. The Prosthetic Use score should show moderate positive correlations with Physical Functioning and PCS but no or weak correlations with MCS.
2. The Prosthetic Mobility score should show strong positive correlations with Physical Functioning and PCS and no or weak correlations with MCS.
3. The Problem score should be negatively correlated to all dimensions of SF-36-Item Health Survey, but those primarily describing physical function should be more strongly related than the others.
4. The Global score should show positive correlations with all subscores of SF-36-Item Health Survey.

Reliability was assessed in two ways, test-retest reliability and internal consistency. Test-retest reliability was assessed by analysis of the differences between the repeated measurements of the subgroup of 48 participants living in Västra Götaland [36,37].

Statistical Methods

For descriptive purposes, we calculated the mean, standard deviation (SD), median (Med), and range for demographic and questionnaire variables. The Mann-Whitney *U*-test and Fisher’s exact test were used to determine whether there were differences between the retest subgroup and the remaining study sample. Floor and ceiling effects of each score of the Q-TFA were calculated as the percentage of subjects achieving the lowest and highest score possible. Criterion validity was determined with Spearman’s nonparametric correlation coefficient (r_s) between the subscores of the Q-TFA and the SF-36-Item Health Survey.

We determined test-retest reliability by descriptive statistics for each occasion and the differences between them, calculations of intraindividual SD, measurement error [38], and intraclass correlation (ICC). We calculated

intraindividual SD by dividing the within-person variance by the number of participants. ICC is a measure of the strength of agreement between repeated measurements, using a one-way analysis of variance model. An ICC of 0.70 or more is recommended for comparisons between groups and 0.90 for evaluation of individual patients [27]. The Wilcoxon Signed Rank Test was used for hypothesis testing of the differences between the mean of scores for occasions 1 and 2. All tests were two-tailed and conducted at a 5-percent significance level.

Internal consistency was analyzed with Cronbach's alpha. Alpha coefficients higher than 0.70 are generally recommended to reflect relation between items within a score [27]. The corrected item total correlation between each item or subscore and the item's overall score was calculated with Pearson's product-moment correlation (r) for descriptive purposes [28]. A correlation of $r = 0.4$ or higher has been shown to be acceptable [27]. Statistical calculations were performed with SPSS 10.1 for Windows (SPSS Inc., Chicago, IL).

RESULTS

Participants

A total of 156 of 204 (76%) of the eligible persons answered the questionnaires. Demographic information and baseline characteristics of the study population and the subgroup constituting the test-retest sample ($n = 48$) are presented in **Table 1**. The participants in the retest subgroup were older ($p = 0.011$), they had a larger interval since amputation ($p = 0.001$), and a higher proportion had a prosthetic limb with vacuum suspension ($p = 0.039$) than the remaining study sample ($n = 108$). No statistically significant differences between the groups were found regarding sex ($p = 0.36$), age at amputation ($p = 0.084$), presence of other medical problems ($p = 0.16$), and reason for amputation divided into vascular or nonvascular reasons ($p = 1.0$). Descriptive statistics as well as floor and ceiling effects of each score on the Q-TFA are shown in **Table 2**.

Clinical Sensibility

Clinical sensibility of the Q-TFA is supported by the development procedures used, which involved expert opinions, review of the literature, semistructured interviews with experienced prosthetic users, and testing of the questionnaire on the target population. The content of

the Q-TFA is relevant because very few items were left unanswered and each item of the questionnaire was answered with every possible option.

Criterion Validity

Correlations between the scores of the Q-TFA and scores of the SF-36 are shown in **Table 3**. As hypothesized, the Prosthetic Use scores were more highly correlated to the physical subscales and the PCS (range of 0.24 to 0.36) than to the mental subscales and the MCS (range of 0.11 to 0.30). Likewise, the Prosthetic Mobility scores were generally more highly correlated to the physical subscales and the PCS (range of 0.38 to 0.79) than to the mental subscales and the MCS (range of 0.10 to 0.44). The Problem score, as hypothesized, was inversely correlated with all the SF-36 scales (range of -0.30 to -0.68). In accordance with the final hypothesis, the Global score was correlated with all the SF-36 scales (range of 0.27 to 0.62).

Test-Retest Reliability

Agreement between test and retest for each subscale are illustrated with scatter plots in the **Figure**. **Table 4** presents descriptive measures, intraindividual SD, measurement error, ICC, and hypothesis testing of mean differences for each subscale. ICC values ranged from 0.89 to 0.94, and measurement error ranged from 10 to 19 points.

Internal Consistency

The corrected item total correlations ranged from 0.57 to 0.74 in the Mobility score, from 0.39 to 0.83 in Problem, and from 0.66 to 0.73 in the Global score. Internal consistency shown by Cronbach's alpha was 0.80 in Mobility, 0.94 in Problem, and 0.83 in the Global score. Within the Mobility score, the 12 items of the capability subscore showed corrected item total correlations between 0.40 and 0.65 with a total Cronbach's alpha of 0.86; for each of the 2 items of the walking aids subscore, r was 0.65 with a Cronbach's alpha of 0.78; and for the 5 items constituting the walking habits subscore, the corrected item total correlations ranged from 0.55 to 0.83, with a Cronbach's alpha of 0.85.

DISCUSSION

This study demonstrates adequate clinical sensibility, criterion validity, test-retest reliability, and internal

Table 1.Demographic and baseline information of study population ($N = 156$) and subgroup of test-retest sample ($n = 48$).

Demographic/Baseline Information	Study Population		Test-Retest Subgroup	
	N (%)	Mean \pm SD, Med (Range)	n (%)	Mean \pm SD, Med (Range)
Male	104 (67)	—	35 (73)	—
Female	52 (33)	—	13 (27)	—
Age	—	51 \pm 13.1, 53 (20–70)	—	55 \pm 11.6, 58 (32–70)
Years Since Amputation	—	25 \pm 15.6, 24 (2–56)	—	32 \pm 16.2, 30 (2–56)
Age at Amputation	—	26 \pm 15.0, 26 (0–66)	—	23 \pm 15.2, 19.5 (0–63)
Cause of Amputation				
Trauma	86 (55)	—	30 (63)	—
Tumor	48 (31)	—	12 (25)	—
Vascular	13 (8)	—	4 (8)	—
Other	9 (6)	—	2 (4)	—
Vacuum Socket (Current Prosthesis)	141 (90)	—	47 (98)	—
Other Prosthesis	15 (10)	—	1 (2)	—
Presence of Other Medical Problem	71 (46)	—	26 (54)	—

SD = standard deviation Med = median

Table 2.Scores of Questionnaire for Persons with a Transfemoral Amputation (Q-TFA) showing descriptive statistics and percentage of floor and ceiling effects ($N = 156$).

Variable	Prosthetic Use	Prosthetic Mobility	Problem	Global
Mean	79	67	34	60
Median	90	71	30	58
SD	25	21	20	21
Minimum/Maximum	2/100	3/98	1/84	0/100
% Floor	0	0	0	1
% Ceiling	31	0	0	5

SD = standard deviation

Table 3.Spearman's correlation coefficient (r_s) for scales of Short-Form 36 (SF-36)-Item Health Survey and scores of Questionnaire for Persons with a Transfemoral Amputation (Q-TFA).

SF-36-Item Health Survey Scale	n	Q-TFA Score			
		Prosthetic Use	Prosthetic Mobility	Problem	Global
Physical Functioning	156	0.36*	0.79*	-0.65*	0.59*
Role Functioning—Physical	154	0.26 [†]	0.49*	-0.59*	0.53*
Bodily Pain	156	0.24 [†]	0.55*	-0.62*	0.52*
General Health	155	0.27 [†]	0.38*	-0.48*	0.50*
Vitality	156	0.16 (NS)	0.20 [‡]	-0.43*	0.42*
Social Functioning	156	0.30*	0.44*	-0.61*	0.52*
Role Functioning—Emotional	154	0.11 (NS)	0.32*	-0.34*	0.34*
Mental Health	156	0.23 [†]	0.22 [†]	-0.45*	0.40*
Physical Component Score	152	0.34*	0.70*	-0.68*	0.62*
Mental Component Score	152	0.19 (NS)	0.10 (NS)	-0.30*	0.27 [†]

* $p < 0.001$ [†] $p < 0.01$ [‡] $p < 0.05$ NS = not significant

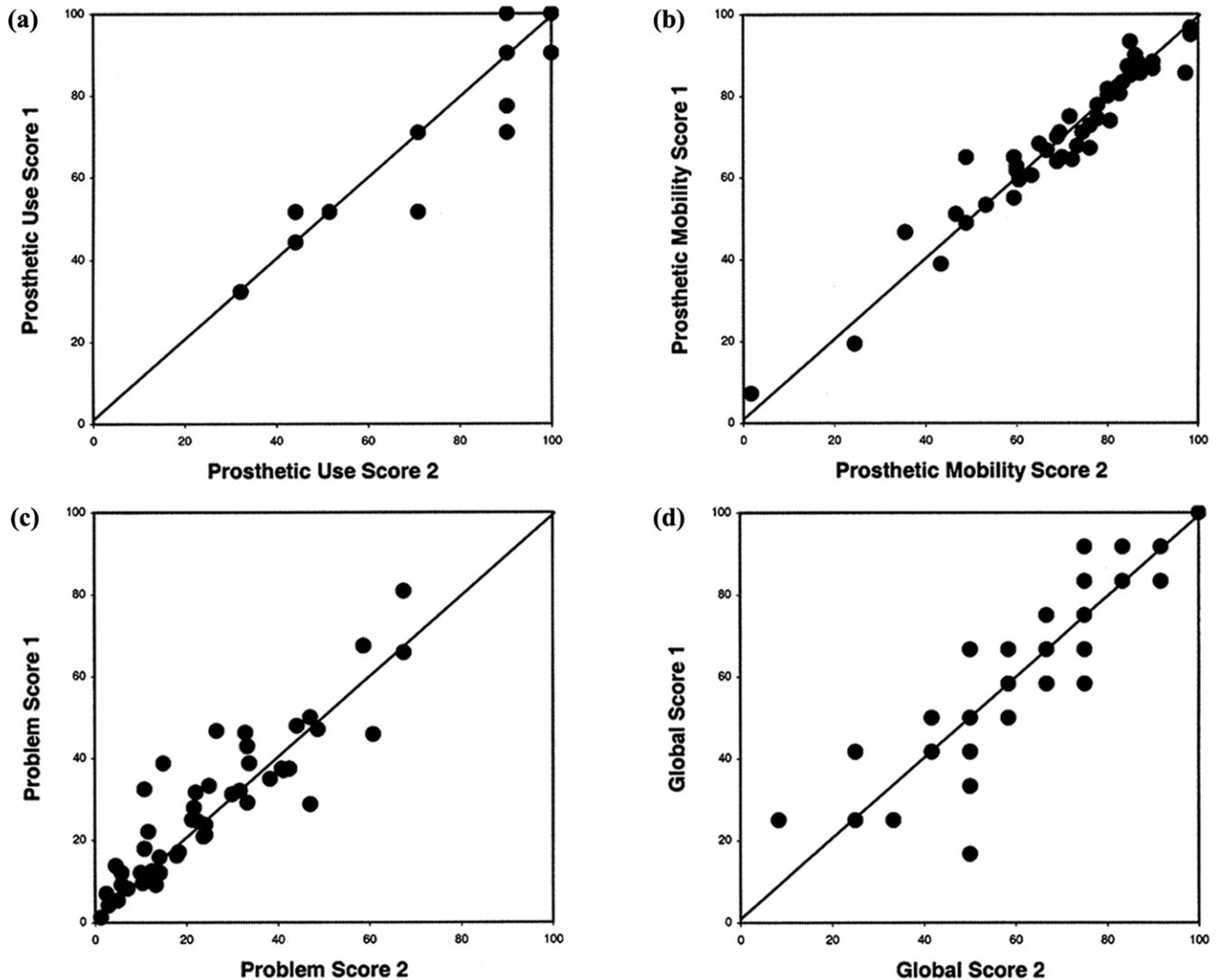


Figure 1.

Scatter plots of test-retest for (a) Prosthetic Use, (b) Prosthetic Mobility, (c) Problem, and (d) Global scores of Questionnaire for Persons with a Transfemoral Amputation (Q-TFA), $n = 48$.

consistency of the scores of Q-TFA for persons using a transfemoral socket prosthesis in Sweden.

Q-TFA is primarily intended to be used with the non-elderly amputee population, among which high prosthetic use has been reported [5,16,39–41]. The age limit of 20 to 70 years in this study was set to follow the current age criteria for treatment with a bone-anchored prosthesis. Sixteen percent ($n = 25$) of the participants were older than 64 years, and eight percent ($n = 13$) had an amputation due to vascular disease. The vascular cases had an age range between 44 and 70 years (mean 61 years), and nine of them used a prosthesis daily, showing that some

individuals having a transfemoral amputation for vascular disease are younger and use their prosthesis extensively.

Extensive problems related to prosthetic use have been reported, suggesting that it is not sufficient to present results solely on the amount of prosthetic use if the purpose is to reflect the situation for the individual [5,12,13,15,16,41,42]. In our development of the Q-TFA, we tried to capture a broad picture of what it is like to have to use a transfemoral prosthetic limb, and the subscores concerning prosthetic mobility, problems, and global health were developed to meet these requirements. The term “mobility” has previously been defined as “the

Table 4.Test-retest reliability of Prosthetic Use, Prosthetic Mobility, Problem, and Global scores when completed on occasions 1 and 2 ($n = 48$).

Score	Occasion 1 Mean \pm SD, Median (Range)	Occasion 2 Mean \pm SD, Median (Range)	Difference, Occasion 2 – Occasion 1 Mean \pm SD, Median (Range)	Intraindividual SD (Measurement Error*)	<i>p</i> -Value Wilcoxon Signed Rank Test	ICC
Prosthetic Use	85.0 \pm 17.5, 90.3 (32.3 – 100)	86.7 \pm 17.6, 90.3 (32.7 – 100)	1.7 \pm 5.8, 0 (29.0)	4.2 (12)	0.028	0.94
Prosthetic Mobility	70.7 \pm 18.2, 71.9 (7.2 – 96.7)	71.0 \pm 19.2, 75.3 (1.7 – 98.3)	0.4 \pm 4.9, 0 (27.8)	3.5 (10)	0.39 (NS)	0.97
Problem	28.0 \pm 17.7, 26.5 (1.3 – 80.8)	25.4 \pm 17.7, 22.3 (1.3 – 67.5)	-2.7 \pm 7.8, -1.3 (42.0)	5.8 (16)	0.026	0.89
Global	62.7 \pm 21.3, 58.3 (16.7 – 100)	63.4 \pm 19.9, 62.5 (8.3 – 100)	0.7 \pm 9.7, 0 (50.0)	6.8 (19)	0.80 (NS)	0.89

*Measurement error = $1.96 \times \sqrt{2} \times$ intraindividual standard deviation (SD)

ICC = intraclass correlation coefficient

NS = not significant

capability and performance of moving oneself and changing and maintaining postures” [26]. We added the wording “when using the prosthesis.” Capability and performance are two different perspectives of mobility, reflecting issues of “*can* do” in the first and “*do* do” in the second [26,43]. The Mobility score of Q-TFA is intended to reflect both perspectives. In the subscore of capability, only those items answered with “Yes” are counted to reflect what the individual really is able to perform.

A main concern among persons using a transfemoral socket prosthesis is socket-related problems. We emphasized these problems in the Problem score to make it sensitive to changes for the assessment of individuals being supplied with a bone-anchored prosthetic limb. The 4-week interval in the Problem score was chosen to be in line with the SF-36-Item Health Survey. However, the items concerning different weather conditions (29 and 30 in the Problem score) and the Walking Habits subscore within the Mobility score had a longer timeframe due to the changing seasons in Sweden, which result in different conditions (**Appendix**, online version only).

Clinical Sensibility

Determination of clinical sensibility involved qualitative analysis rather than statistical testing. When a new questionnaire is developed, the item-generation process should include input from specialists from the area of interest, a review of existing literature, and interviews with

patients to ensure that the content of the questionnaire covers relevant and important issues [27]. We considered these requirements in developing the Q-TFA. The fact that every individual item included in the scores was answered with the full range of options that were presented further supports the relevance of the content toward the target population. Another concern is the ease of answering the questionnaire. One indication of ease is a low number of missing answers. In this study, the number of missing answers of separate items was, in general, very low, and enough of the items were answered to be able to calculate all four scores for each participant. Some individuals, however, needed additional information on how to answer the walking habits subscore, and six individuals (4%) did not answer all five questions of this subscore, which indicates a possible need to change the wording of this specific question.

Criterion Validity

In this study, we have begun the process of criterion validation by assessing the relationship between scores of the SF-36-Item Health Survey and the Q-TFA. Correlation coefficients of $r = 0.4$ to 0.8 are considered adequate between abstract constructs such as those in generic HRQL measures and a new instrument [28]. We achieved these levels of correlation for most of the hypothesized relationships, and the associations between the SF-36-Item Health Survey and Q-TFA were in accordance with

the prior hypotheses made to assure criterion validity in this study. We had, however, expected a somewhat greater correlation between the Physical Function and Prosthetic Use scores ($r_s = 0.36$) (Table 3). The low association suggests that frequent use of a prosthetic limb is not equivalent to high-level prosthetic physical function.

A lower-limb amputation has previously been reported to influence mainly the physical dimensions of the SF-36-Item Health Survey [2,12,34]. The scores of the Q-TFA were also generally more strongly associated with those dimensions primarily reflecting physical health. However, the Problem and Global scores were also associated, to a substantial degree, to scores reflecting physiological well-being (Table 3), indicating that the Q-TFA also captures a broad range of health concepts. This study confirms adequate criterion validity of Q-TFA. Additional research to further assess construct validity, sensibility, and responsiveness of the tool is ongoing and will be presented in future studies.

Reliability

Along with the second mailing of the Q-TFA, we included four additional questions to verify if any important change in condition had occurred since the first mailing. Nine individuals were excluded from the calculations of agreement because of such changes (two volume changes of the residual limb, two sores on the residual limb, one flu, one bronchitis, one acute lumbago, one change to a new prosthesis, one other personal problem). We are, however, not convinced that we really captured a stable situation in all the included cases. The clinical experience is that a true stable condition for persons using a prosthetic limb is rare and that frequent, sometimes daily, changes in the condition of the residual limb could appear, and this could have influenced the test-retest result. The overall agreement between tests reveals that scores of Q-TFA are reliable (Table 4). The very small but statistically significant difference of the mean in the Prosthetic Use and Problems scores (+1.7, -2.7) is not considered to have important clinical relevance. What is more important is the result of the intraindividual SD, the measurement error, and the ICC. The ICC is a measure of the variance between subjects on repeated measures, while the intraindividual SD describes the within-person variation [36]. The interpretation of the measurement error of scores of Q-TFA in this study reveals that a difference larger than 12 in the Prosthetic Use score, 10 in the Mobility score, 16 in the Problem score, and 19 in the Global

score are needed to claim a real difference, above the noise, on repeated assessments with 95-percent confidence (Table 4). The results demonstrate less agreement at retest in the Problem and Global scores than in the Prosthetic Use and Mobility scores. Hays et al. claim that “the well-being part of HRQL is more subjective than the functioning component” [44]. The same circumstances could explain the differences of agreement regarding issues of prosthetic use and mobility on one hand and the more subjective matters of problems and global health on the other. A more subjective item could be expected to depend more on the mood of the person answering the questionnaire on that specific day, leading to larger variation at retest.

In concordance with the clinimetric approach [27,33], we tried to include items that could be expected to be important to the target population, regardless of their influence on the consistency of each score. Nevertheless, we were also interested in assessing whether the scales exhibited sufficient internal consistency. The results revealed good internal consistency with no alpha coefficient below 0.7. Furthermore, only one item was below the recommended $r = 0.4$ when calculations of the corrected item-total correlation were conducted.

Study Limitations

A self-report instrument has the advantages of reflecting the subjective experience of each participant and being easy and inexpensive to administer. The limitation is that misunderstandings of answers to the questionnaire are difficult to control. However, because our primary goal was to describe the current situation from the view of the patient, with minimal influence from an investigator, we chose a self-report tool.

The method by which participants were recruited to the study might be another source of error. Presently, no national registration exists of persons with a lower-limb amputation in Sweden, and potential participants were identified through amputees associations and selected orthopedic workshops. Thus, some potential participants may not have been reached. The number of participants in the study is, however, satisfactory with regard to the inclusion criteria used. Only adults, 20 to 70 years of age, who had been amputated at least 2 years earlier and who currently were using a prosthetic limb, were included. The age limit was set to reflect the population that could be considered potential candidates to be treated with a bone-anchored prosthesis. To reach persons that could be

considered established prosthetic users, we chose a minimum 2-year interval since the amputation.

CONCLUSIONS

We developed a new self-report questionnaire, the Q-TFA, to assess nonelderly persons having a transfemoral amputation and using a prosthesis. This study of 156 individuals with transfemoral amputation living in Sweden documents adequate clinical sensibility, criterion validity, test-retest reliability, and internal consistency for the Q-TFA for persons using a transfemoral socket prosthesis. Further assessment of the construct validity, sensitivity, and responsiveness of the Q-TFA is needed.

REFERENCES

- Pell JP, Donnan PT, Fowkes FG, Ruckley CV. Quality of life following lower limb amputation for peripheral arterial disease. *Eur J Vasc Surg*. 1993;7(4):448–51.
- Pezzin LE, Dillingham TR, MacKenzie EJ. Rehabilitation and the long-term outcomes of persons with trauma-related amputations. *Arch Phys Med Rehabil*. 2000;81(3):292–300.
- Burger H, Marincek C. The life style of young persons after lower limb amputation caused by injury. *Prosthet Orthot Int*. 1997;21(1):35–39.
- van der Schans CP, Geertzen JH, Schoppen T, Dijkstra PU. Phantom pain and health-related quality of life in lower limb amputees. *J Pain Symptom Manage*. 2002;24(4):429–36.
- Schoppen T. Functional outcome after a lower limb amputation. Groningen: University of Groningen, Netherlands; 2002.
- Ebskov LB. Level of lower limb amputation in relation to etiology: an epidemiological study. *Prosthet Orthot Int*. 1992;16(3):163–67.
- Pohjola-Tinikka T, Alaranta H. Ten-year survival of Finnish lower limb amputees. *Prosthet Orthot Int*. 1998;22(1):10–16.
- Pernot HF, de Witte LP, Lindeman E, Cluitmans J. Daily functioning of the lower extremity amputee: an overview of the literature. *Clin Rehabil*. 1997;11(2):93–106.
- Holden JM, Fernie GR. Extent of artificial limb use following rehabilitation. *J Orthop Res*. 1987;5(4):562–68.
- Gauthier-Gagnon C, Grise MC, Potvin D. Enabling factors related to prosthetic use by people with transtibial and transfemoral amputation. *Arch Phys Med Rehabil*. 1999; 80(6):706–13.
- Gautier-Gagnon C, Grise MC, Potvin D. Predisposing factors related to prosthetic use by people with a transtibial and transfemoral amputation. *J Prosthet Orthot*. 1998;10: 99–109.
- Legro MW, Reiber G, del Aguila M, Ajax MJ, Boone DA, Larsen JA, Smith DG, Sangeorzan B. Issues of importance reported by persons with lower limb amputations and prostheses. *J Rehabil Res Dev*. 1999;36(3):155–63.
- Matsen SL, Malchow D, Matsen FA 3rd. Correlations with patients' perspectives of the result of lower-extremity amputation. *J Bone Joint Surg Am*. 2000;82-A(8):1089–95.
- Hagberg K, Brånemark R. Consequences of nonvascular trans-femoral amputation: a survey of quality of life, prosthetic use and problems. *Prosthet Orthot Int*. 2001;25(3): 186–94.
- Gallagher P, Maclachlan M. Adjustment to an artificial limb: a qualitative perspective. *J Health Psychol*. 2001; 6(1):85–100.
- Dillingham TR, Pezzin LE, MacKenzie EJ, Burgess AR. Use and satisfaction with prosthetic devices among persons with trauma-related amputations: a long-term outcome study. *Am J Phys Med Rehabil*. 2001;80(8):563–71.
- Dillingham TR, Pezzin LE, MacKenzie EJ. Limb amputation and limb deficiency: epidemiology and recent trends in the United States. *South Med J*. 2002;95(8):875–83.
- Ebskov LB. Trauma-related major lower limb amputations: an epidemiologic study. *J Trauma*. 1994;36(6):778–83.
- Brånemark R, Brånemark PI, Rydevik B, Myers RR. Osseointegration in skeletal reconstruction and rehabilitation: a review. *J Rehabil Res Dev*. 2001;38(2):175–81.
- Grise MC, Gauthier-Gagnon C, Martineau GG. Prosthetic profile of people with lower extremity amputation: conception and design of a follow-up questionnaire. *Arch Phys Med Rehabil*. 1993;74(8):862–70.
- Day HJ. The assessment and description of amputee activity. *Prosthet Orthot Int*. 1981;5(1):23–28.
- Miller WC, Deathe AB, Speechley M. Lower extremity prosthetic mobility: a comparison of 3 self-report scales. *Arch Phys Med Rehabil*. 2001;82(10):1432–40.
- Legro MW, Reiber GD, Smith DG, del Aguila M, Larsen J, Boone D. Prosthesis evaluation questionnaire for persons with lower limb amputations: assessing prosthesis-related quality of life. *Arch Phys Med Rehabil*. 1998;79(8):931–38.
- Cella D, Nowinski CJ. Measuring quality of life in chronic illness: the functional assessment of chronic illness therapy measurement system. *Arch Phys Med Rehabil*. 2002;83(12 Suppl 2):S10–17.
- Patrick DL, Chiang YP. Measurement of health outcomes in treatment effectiveness evaluations: conceptual and methodological challenges. *Med Care*. 2000;38 Suppl 9:II14–25.
- Bussmann JB, Stam HJ. Techniques for measurement and assessment of mobility in rehabilitation: a theoretical approach. *Clin Rehabil*. 1998;12(6):455–64.
- Fayers PM, David M. Quality of life: assessment, analysis and interpretation. Chichester: John Wiley & Sons, Ltd; 2000.

28. Streiner DL, Norman GR. Health measurements scales: a practical guide to their development and use. 2nd ed. New York: Oxford University Press; 1995.
29. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30(6):473–83.
30. Ware JE Jr, Kosinski M, Bayliss MS, McHorney CA, Rogers WH, Raczek A. Comparison of methods for the scoring and statistical analysis of SF-36 health profile and summary measures: summary of results from the Medical Outcomes Study. *Med Care*. 1995;33 Suppl 4:AS264–79.
31. Sullivan M, Karlsson J, Ware JE. SF-36 Hälsoenkät: Svensk manual och tolkningsguide [Swedish manual and interpretation guide]. Gothenburg: Sahlgrenska University Hospital; 1994.
32. Marx RG, Bombardier C, Hogg-Johnson S, Wright JG. Clinimetric and psychometric strategies for development of a health measurement scale. *J Clin Epidemiol*. 1999;52(2):105–11.
33. Feinstein AR. *Clinimetrics*. New Haven and London: Yale University Press; 1987.
34. Smith DG, Horn P, Malchow D, Boone DA, Reiber GE, Hansen ST Jr. Prosthetic history, prosthetic charges, and functional outcome of the isolated, traumatic below-knee amputee. *J Trauma*. 1995;38(1):44–47.
35. Beaton DE, Hogg-Johnson S, Bombardier C. Evaluating changes in health status: reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. *J Clin Epidemiol*. 1997;50(1):79–93.
36. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clin Rehabil*. 1998;12(3):187–99.
37. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307–10.
38. Bland JM, Altman DG. Measurement error. *BMJ*. 1996;313(7059):744.
39. Burger H, Marincek C, Isakov E. Mobility of persons after traumatic lower limb amputation. *Disabil Rehabil*. 1997;19(7):272–77.
40. Hoffman RD, Saltzman CL, Buckwalter JA. Outcome of lower extremity malignancy survivors treated with transfemoral amputation. *Arch Phys Med Rehabil*. 2002;83(2):177–82.
41. Walker CR, Ingram RR, Hullin MG, McCreath SW. Lower limb amputation following injury: a survey of long-term functional outcome. *Injury*. 1994;25(6):387–92.
42. Gallagher P, Allen D, Maclachlan M. Phantom limb pain and residual limb pain following lower limb amputation: a descriptive analysis. *Disabil Rehabil*. 2001;23(12):522–30.
43. Young NL, Williams JI, Yoshida KK, Bombardier C, Wright JG. The context of measuring disability: does it matter whether capability or performance is measured? *J Clin Epidemiol*. 1996;49(10):1097–101.
44. Hays RD, Hahn H, Marshall G. Use of the SF-36 and other health-related quality of life measures to assess persons with disabilities. *Arch Phys Med Rehabil*. 2002;83(12 Suppl 2):4–9.

Submitted for publication November 13, 2003. Accepted in revised form April 14, 2004.

