Reliability and validity of the Finnish version of the Prosthesis Evaluation Questionnaire

Abstract

BACKGROUND Thus far there have been no specific patient-reported intruments in Finnish for health-related quality of life assesment after major lower extremity amputation and successful prosthesis fitting.

METHODS The Prosthesis Evaluation Questionnaire was translated and cross-culturally adapted to Finnish. Participants completed a questionnaire package including the Finnish version of the Prosthesis Evaluation Questionnaire and the 15D health-related quality of life instrument. Scales (n = 10) were tested for internal consistency, floor-ceiling effect, and reproducibility for which participants completed the Prothesis Evaluation Questionnaire twice within a 2-week interval. Validity was tested by estimating the correlation between the 15D index and the scales. The authors included 122 participants who had completed the questionnaire on two separate occasions in the final analysis.

RESULTS Mean scale scores of the 10 scales varied from 52 to 83. Cronbach alphas ranged from 0.67 to 0.96. The total score showed no floor-ceiling effect. Reproducibility of the scales was good (intraclass correlation coefficient, 0.78-0.87; coefficient of repeatability, 19-36). Significant correlations were observed between the 15D index and the scales for Ambulation, Social Burden, Usefulness and Well-being.

CONCLUSIONS This study provided evidence of the reliability and validity of the Finnish version of the Prothesis Evaluation Questionnaire in assessing the health-related quality of life among major lower extremity amputated patients who have been fitted with prosthesis.

Key Words: Rehabilitation; Prosthesis; Validity; Reliability; Amputation; Psychometrics

Introduction

Assessing rehabilitation effectiveness with high quality patient-reported outcome instruments makes it possible to obtain an amputee-centered experience in a relevant way [1]. Generic instruments that are designed to obtain information from a broad variety of health parameters may not be specific enough to measure the specific problems encountered by amputees. Thus, the Prosthesis Profile of the Amputee questionnaire was introduced in 1994 to provide a tool for lower-extremity amputee-specific assessment [2]. Qualitative studies may provide deep insight into patients' biopsychosocial perspectives that otherwise would be hard to obtain. However, quantitative data obtained from patient-reported instruments can provide accurate and reliable outcomes that can be statistically analyzed for the assessment of effectiveness of different methods of surgical approaches or rehabilitation.

The rehabilitation assessment further evolved towards emphasizing the impact of rehabilitation on the health-related quality of life (HRQoL), when Legro et al. developed and validated the English version of the Prosthesis Evaluation Questionnaire (PEQ) [3]. The PEQ is an amputee-specific quality of life instrument that can be used to assess the HRQoL of lower-extremity prosthesis users. It has been further psychometrically investigated and validated after being translated into several other languages [4-8]. Furthermore, the PEQ has been used in a great variety of studies [9].

There has hitherto been no validated lower-extremity amputee-specific patient-reported

outcome instrument in Finnish. The authors aimed, therefore, to transculturally adapt the English PEQ into a Finnish version, which was then tested for reliability and validity among patients who have undergone major lower extremity amputation and have rehabilitated to prosthesis users.

Methods

Ethical considerations and participants

The Ethics Committee of the Helsinki University Hospital approved the study. The authors included in the study patients, who had undergone major lower extremity amputation, were at least 18 years old, had full ability to understand written Finnish and had rehabilitated to prosthesis users in the Helsinki and Uusimaa Hopital District or the Central Finland Health Care District, Finland. The participants provided their written consent according to the Helsinki Declaration. The authors approached by mail 597 consecutive patients who had undergone major lower extremity amputation and had successful prosthesis fitting.

Translation and adaptation

The authors contacted the developer of the PEQ to obtain permission to use the English language questionnaire. The translation and adaptation process adhered to the International Society for Pharmacoeconomics and Outcomes Research guidelines [10].

Two native Finnish-speaking translators who were professionals in the field of rehabilitation and fluent in English produced a forward-translation independently of each other. Differences encountered between the two forward translations were discussed by the steering group who then synthesized one forward-translation. A back-

translation was produced by an English language-expert who was fluent in Finnish and familiar with the Finnish culture and translating patient-reported outcome instruments but unfamiliar with the current instrument. A back-translation panel consisting of all three translators reviewed the translation drafts and compared them to the original English version and provided a written report. In addition to this a language expert of the Finnish Medical Society Duodecim was consulted when translation problems were encountered. A multidisciplinary committee reviewed each part of the translation processes separately.

The pre-final version underwent pre-testing together with cognitive debriefing among 14 Finnish patients who had undergone transtibial amputation and who were transtibial prosthesis users. The cognitive debriefing followed the European Organisation for Research and Treatment of Cancer (EORTC) guidelines [11] to identify any offensive content, problems with understandability, cultural relevance, difficulties in anwering or in interpretation of the questions and whether the participants would ask any question differently. In the last phase, the multidisciplinary committee reviewed the pre-testing outcomes and interview reports. The final version was introduced and was then proofread by the language expert of the Finnish Medical Society Duodecim (Supplementary file).

Instruments

Prosthesis Evaluation Questionnaire. The lower extremity amputee-specific PEQ is a valid, comprehensive instrument comprising 82 items with seven different main themes. The items refer to the preceding four weeks. The PEQ also contains items with

checkboxes for assessing topics such as Satisfaction, Pain, Transfers, Prosthesis care and Self-efficacy. These items are scored individually.

The PEQ can be separated into 10 validated scales: Ambulation, Appearance, Frustration, Perceived Response, Social Burden, Utility, Residual Limb Health, Sounds, Transfers and Well-being [3]. The items are completed on a visual analogue scale (0-100 mm; from worst to best). The total scores for each scale are calculated through the arithmetic mean of all items of the scale.

15D instrument. The 15D is a valid, generic HRQoL instrument containing 15 dimensions: moving, seeing, hearing, breathing, sleeping, eating, speech, excretion, usual activities, mental function, discomfort and symptoms, depression, distress, vitality and sexual activity [12]. Respondents choose one of the five levels in each dimension that best describes their current state of health (1-5; best to worst possible score). The 15D produces both a HRQoL profile and a single index score that represents the overall HRQoL. The single index score ranges from 0 to 1, with 0 equivalent to being dead and 1 being in the best imaginable HRQoL state. Reproducibility and the minimum important change of the 15D have been reported to be 0.90 and 0.015, respectively [13,14].

Sociodemographic and clinical questionnaire. The authors obtained information on participants' age, sex, cause for amputation, comorbidities, amputation level (disarticulation amputation was considered as above-knee amputation), time since amputation, and beginning of the prosthesis use. In addition, a visual analogue scale on a 0 to 100 mm scale (0-100 mm; best to worst) was used for measuring participants'

self-reported general health and pain during the preceding week. The NRS is another instruments as it is a segmented numeric version of the visual analog scale (VAS) in which a respondent selects a whole number (0–10 integers) that best reflects the intensity of their pain. The visual analog scale, which the authors used, is a widely accepted measure and validated for pain assessment [15].

PEQ validation course and reproducibility setting

In addition to the pre-information form, the authors included the following intruments in the first questionnaire package: the Finnish PEQ, the 15D and the general health and pain visual analogue scale questions. Participants returned the completed questionnaires together with the signed informed consent. Potential participants who did not return the first questionnaire set within a week received a reminder letter. After the participants had completed the first questionnaire, the authors mailed them the PEQ instrument a second time along with a survey. The purpose of both was to ascertain whether the patients' health status had changed between completing the first round of questionnaires. The authors included participants who had completed the PEQ twice in the final analyses.

Statistics

The authors present the data as means with standard deviations (SD), medians with interquartile ranges (IQR), 95% confidence intervals (95% CI), or as counts with percentages or ranges. The scale completion rate is provided to illustrate the percentage of missing items in the analysis. Predefined hypotheses were placed based on the existing literature or general presumptions [table 1].

A one-way random-effects model with absolute agreement was used to measure relative reliability or intraclass correlation coefficient. The intraclass correlation coefficient value was classified according to Cicchetti et al. as poor (< 0.40), fair (0.40-0.59), good (0.60-0.74) or excellent (0.75-1.00) [16].

The internal consistency was estimated by calculating Cronbach's alpha [17] with bootstrapped 95% CIs.

The coefficient of repeatability expressed the expected maximum size of 95% of the absolute differences between paired observations. The 95% CI was obtained by bias corrected and accelerated bootstrapping (5000 replications).

The Pearson method served to calculate the correlation coefficients. Statistical significance in the correlation coefficient was set at $p<0.05$ and calculated using Sidak-adjusted probabilities. Bias-corrected bootstrapping was used to obtain the confidence intervals for the mean changes between the two measurements and reproducibility.

The authors used linear regression analyses to identify the appropriate predictors of the 15D age- and gender-standardized regression coefficients Beta (β). The β-value is a measure of how strongly each predictor variable influences the criterion (dependent) variable. The β was measured in units of standard deviation. Cohen's standard for β-values above 0.10, 0.30 and 0.50 represent small, moderate and large relationships, respectively.

Results

Of the 167 participants (response rate, 28%), who returned the questionnaires together with their signed written consent, a total of 122 patients (73%) had completed both the first and the second questionnaires and were included in the study. The participants' ages ranged from 19 to 93 [table 2]. The most common indication for primary major lower-extremity amputation was vascular disease (29.5%). Thirty-six percent (n = 44) of participants reported having no comorbidities [table 2]. The time from amputation to completion of the outcome measures varied from four months to 69 years. Fifty percent of the participants had undergone amputation less than five years earlier.

Translation and adaptation

Minor linguistic differences were noted between the two forward translations. A back-translation panel review revealed no major problems between the back-translation and the original English version. The multidisciplinary committee required that "rate the weight of your prosthesis" in item 1C be changed to "evaluate the weight of your prosthesis" in order to improve the Finnish. Item 1N required amending "prosthesis cover" to "cosmetic surface" which is preferred in Finnish. In item 1Q the word "stump" was added for clarification. In the Finnish language, the word "stump" is well accepted to describe the distal end of an amputated limb. Translation of the words: " "shooting", "searing", "stabbing", "sharp", "ache" in the "Group 2" of the PEQ instrument required the help of the language expert to find suitable matches in Finnish. The pre-testing and participants' cognitive debriefing gave no reason for changes.

Reliability

Floor-ceiling effect. The PEQ showed no floor-effect (0 score) on the total score. Nine of the scales had no floor effect. Altogerher 1% had the lowest score in Ampulation scale. A ceiling effect of one to five percent was found in five of the scales [table 3]. The highest ceiling effect was strongest in the Perceived Responses scale (5%).

Internal consistency. Cronbach's alpha for the 10 scales revealed an internal consistency ranging from 0.67 (Appearance) to 0.96 (Ambulation) [table 3].

***Reproducibility*.** The mean value (SD) of the PEQ subscales at measurement one was 65.1 (23.7) (table 4). The mean change between the two measurement times ranged from 0.0 to 2.1 in the separate scales. All scales had good reproducibility [table 4]. The coefficient of repeatability ranged from 19 for Usefulness to 36 for the Frustration scales [table 4].

Validity

Convergent validity. Pearson correlation coefficients between the PEQ scales and age were low (range, -0.28 to 0.15) [table 5]. The correlation of the PEQ scale scores with time since prosthetization was also poor. Strong correlation was found between general pain or general health and Usefulness, Ambulation, Transfers, Perceived responses, Social Burden and Well-Being scales.

Strong correlation was found between the15D index and the scales of Ambulation Social burden, Transfers, Usefulness and Well-being [Figure 1].

Discussion

The authors successfully produced a Finnish PEQ instrument and evaluated its psychometric properties. To the authors' knowledge this study has the largest study population to assess the psychometrics of the PEQ. The psychometric analyses showed evidence of good reproducibility and validity for the Finnish PEQ. The Finnish version of the PEQ instrument can now be used to assess the effectiveness of different amputation techniques, stump reconstruction methods, and rehabilitation after successful prosthesis fitting.

Translation and adaptation

The translation and cross-cultural adaptation process adhered rigorously to the International Society for Pharmacoeconomics and Outcomes Research guidelines [10]. All the discrepancies and changes made during the translation phases were meticulously recorded in written reports. One previous translation report addressed the linguistic or cultural problems encountered during the translation process [5]. The authors found that adjustments were required to adjust for linguistic differences between the Finnish version of the PEQ and the original English version.

In the Arabic translation of the PEQ, the authors found the word "phantom" could be interpreted as a "ghost sensation" among the Saudi people [5]. The word "phantom" does not have a negative connotation in Finnish nor is it linked to ghosts. The identification of items in the Arabic version was changed to match the group number rather than the page number as in the original English version [3,5]. The Finnish version also uses the group numbers to identify the items. The new numbering of items should be taken into consideration when using the Finnish PEQ.

Reliability

A floor-ceiling effect of less than 15% is considered acceptable [18]. Reliability testing for the PEQ by Legro et al. found a floor effect of 22% in the scales of Frustration and a ceiling effect of 25% in the Transfers scale in a similar study population to that of the present study [3]. No explanation for this was provided by the Legro group. It could be hypothesized however that the ceiling effect was a consequence of the answers of those participants who had been amputated 9 to 28 years before assessment took place as they received the highest scores in the Transfers scale [3]. Other validation studies of PEQ did not report floor-ceiling values [5-8]. In the present study, five percent of participants received the maximum score in the Perceived Responses scale. Not a single participant reported the maximum scores in the Transfers scale. The PEQ scales of the Finnish version seemed to have no floor or ceiling effect based on the present study's findings. Thus, the present analysis provided evidence that it is somewhat unlikely that the PEQ would yield inaccurate maximum scores.

The internal consistency of the original English PEQ varies between 0.67 and 0.89 in the 10 scales [3]. Cronhach's alphas between 0.67 and 0.96 were noted in the present study. According to the literature, Cronbach's alpha of 0.8 or more is considered sufficient [19]. In the present study four of the 10 subscales were slightly lower than the proposed benchmark, but these values can be considered acceptable. Benavent et al. found poor internal consistency in the scales of Appearance and Residual Limb (0.37 and 0.15, respectively) [8]. Cronbach's alpha varied in the remaining scales between 0.55 and 0.93 in that study [8]. Other studies have reported the internal consistency of Appearance and Residual Limb Health of 0.73-0.77 and 0.77-0.80, respectively [3,6,7]. The results of the present study were similar to those the previous studies [3,6,7] as the

internal consistency of the Appearance scale was 0.79 and that of the Residual Limb was 0.67. Internal consistency of the other eight scales were also mainly in concordance with previously published literature [3,6-8]. The internal consistency in the present study was below 0.9 in all scales, indicating that there was no item repetition [20]. The authors assessed reproducibility after a mean interval of two weeks. The participants' health was stable in the interim period. The optimal interim time between the two assessments has previously been placed at two weeks in assessment of the reproducibility in situations where there is no acute change in the participants' health [21]. According to the classification by Cicchetti et al. [14], all scales used in the Finnish PEQ had excellent intraclass correlation coefficient values (0.78-0.87). Conrad and colleagues reported intraclass correlation coefficient values that ranged from good (0.65, Well-being) to excellent (0.92, Ambulation) between the scales in the Brazilian Portuguese version of the PEQ [6]. However, the Conrad group reported on a smaller study population that consisted only of 65 patients who had undergone major lower-extremity amputation [6]. The authors also calculated the coefficient of repeatability for the PEQ scales. The coefficient of repeatability can be used to obtain the value for absolute reliability, the expected maximum size of 95% of the absolute differences between paired observations. The present study reflects the good reproducibility of the PEQ instrument scales. The authors found that the coefficient of repeatability ranged from 19 to 36 between the different scales in the present study. The alternative of calculating the coefficient of repeatability values may be more accurate compared to the standard error of measurement as it takes into account both random and systematic errors [22].

Validity

Age has previously been reported to correlate with Residual Limb Health and Frustration scale [3]. In the study by Legro et al., scores were higher in patients who were younger than 40 years old [3]. The present study found low negative correlation between age and Usefulness and Ambulation. The negative value indicates that as the age of the patient increases, the worse the score gets. Locomotor activity might be decreased in older individuals, which could explain the correlation. Interestingly, time from prosthesis fitting to assessment had no correlation with the PEQ score, which supports the findings reported by Legro and colleagues [3]. Both general health and general pain correlated strongly with the scales of Usefulness, Ambulation, Transfers, Perceived Responses, Social Burden and Well-being. Previous psychometric studies of the PEQ have not assessed scale correlations with separate measurements of general health or general pain [3-8]. However, the Usefulness scale correlated well with General Health summary score in the study by Benavent et al.[8], which also supports the findings of the present study.

The authors found a notable relationship between the scales of Ambulation, Social burden, Usefulness and Well-being and the 15D HRQoL index in the construct validity analysis [Figure 1]. The evidence suggests that PEQ has good criteria validity when it comes to assessing HRQoL. Previously there has been no validated prosthesis-related quality of life instrument in Finnish. Evidence of validity of the PEQ presented here supports its use to assess the HRQoL of patients who have undergone major lower extremity amputation and have been fitted with prosthesis. Legro et al. found strong correlation with Ambulation and the SF-36 summary score of Physical Function (r= 0.61) [3]. Further, Benavent and others [8] found that there was strong correlation

between the Ambulation scale and the SF-36 summary scores of General Health (*r*= 0.71), Vitality (*r*= 0.73), Social Function (*r*= 0.78 and Mental Health (*r*= 0.67). A strong correlation (r= 0.73) between the PEQ Social Burden scale and the SF-36 Social Function summary score was also found. The authors used the 15D HRQoL instrument in the present study as it is widely accepted in health care internationally and especially in Finland. The 15D can be linked to the ICF-classification [23]. Its properties have proven superior to several other widely used HRQoL patient-reported instruments [13,14, 24,25,26].

Clinical applications

Amputation has a significant impact on patients' lives. Optimally, rehabilitation allows the patients to return to their previous daily activities and social affairs. However, prosthesis fitting and rehabilitation cause notable cost to society. There is a need for assessment tools in measuring the need of treatment and rehabilitation as well as their effectiveness. Several different techniques (e.g. in flap design) are used for major lower extremity amputation. Furthermore, the amputation stump may not always have a sufficient amount of healthy soft tissue for local flap stump coverage and microvascular reconstruction or bone-lengthening techniques are thus needed in selected cases. These surgical techniques may have an impact on how the prosthesis fits. Inadequate rehabilitation methods may lead to poor results and abandonment of the prosthesis. The effectiveness of different surgical methods and rehabilitation processes and their impact on health-related quality of life can be assessed using the PEQ instrument in patients who have been fitted with prosthesis. However, the PEQ is a comprehensive questionnaire that has a large amount of items (N = 82). It gives extensive information

about the patient and prosthesis use. The 10 validated scales might be better in clinical practice as they can be used as a patient profile.

Strengths and limitations

The study recruited a heterogeneous population of patients who had undergone major lower extremity amputation. Some may consider this approach as a weakness. However, a heterogeneous study population allows a better generalization to be made about the outcomes of this study. One limitation was the low response rate that, nonetheless, can be considered acceptable for a psychometric study. Previous studies have shown that ischaemia is the major cause or major lower extremity amputation [27]. However, no epidemiological studies have been conducted to provide information of the amputation etiology of patients who are fitted with prosthesis. Using several reference outcomes would have brought even deeper knowledge of the convergent validity of the Finnish PEQ. However, the authors did not have another validated amputee-specific instrument in Finnish to compare. Francihignoni et al. analyzed the PEQ Ambulation scale using item response theory [4]. A single item was omitted and a 5-point answer scale established [4]. A Rasch analysis could have provided even more insight into the construct validity of the Finnish PEQ in the present study. Further studies should therefore aim to assess the construct of the PEQ scales using *inter alia* Rasch analysis and the responsiveness with a longitudinal study design.

Conclusions

The authors conclude that the PEQ instrument was succesfully translated and cross-culturally adapted into the Finnish language version. Psychometric testing of the Finnish version of the PEQ showed evidence of its reliability and validity in assessing

prosthesis-related quality of life in patients who have undergone major lower extremity amputation and who have rehabilitated to prosthesis users. The Finnish PEQ is a suitable patient-reported outcome instrument for clinical use and in scientific studies for assessing the efficacy and outcomes of different amputation techniques, stump reconstruction methods, and rehabilitation in patients who have been fitted with prosthesis.

Declaration of Interests The authors report no conflicts of interest.

References

1. Hawkins AT, Henry AJ, Crandell DM, Nguyen LL. A systematic review of functional and quality of life assessment after major lower extremity amputation. Ann Vasc Surg. 2014;28(3):763-80.

2. Gauthier-Gagnon C, Grisé MC. Prosthetic profile of the amputee questionnaire: validity and reliability. Arch Phys Med Rehabil 1994;75(12):1309-14.

3. 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):931-8.

4. Franchignoni F, Giordano A, Ferriero G, Orlandini D, Amoresano A, Perucca L. Measuring mobility in people with lower limb amputation: Rasch analysis of the mobility section of the prosthesis evaluation questionnaire. J Rehabil Med. 2007;39(2):138-44.

5. Day SJ, Buis A. Cross cultural equivalence testing of the Prosthetic Evaluation Questionnaire (PEQ) for an Arabic speaking population. Prosthet Orthot Int. 2012;36(2):173-80.

399 6. Conrad C, Chamlian TR, Ogasowara MS, Pinto MA, Masiero D. Translation into
400 Brazilian Portuguese, cultural adaptation and validation of the Prosthesis Evaluation
401 Questionnaire. Jornal Vascular Brasileiro. 2015;14(2):110-114.
402 7. Safer VB, Yavuzer G, Demir SO, Yanikoglu I, Guneri FD. The prosthesis evaluation
403 questionnaire: reliability and cross-validation of the Turkish version. J Phys Ther Sci.
404 2015;27(6):1677-80.
405 8. Benavent JV, Igual C, Mora E, Antonio R, Tenias JM. Cross-cultural validation of
406 the Prosthesis Evaluation Questionnaire in vascular amputees fitted with prostheses in
407 Spain. Prosthet Orthot Int. 2015;40(6):713-9.
408 9. Boone DA, Coleman KL. Use of the Prosthesis Evaluation Questionnaire. JPO.
409 2006:18(6):P68-P79.
410 10. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A et al.
411 Principles of good practice for the translation and cultural adaptation process for
412 patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for
413 Translation and Cultural Adaptation. Value Health 2005;8:94-104.
414 11. Dewolf L, Koller M, Velikova G et al. Bottomley A. EORTC Quality of life group
415 translation procedure. 3rd ed., 2009.
416 12. Sintonen H. The 15D instrument of health-related quality of life: properties and
417 applications. Ann Med 2001;33:328–36.
418 13. Stavem K. Reliability, validity and responsiveness of two multiattribute utility
419 measures in patients with chronic obstructive pulmonary disease. Qual Life Res.
420 1999;8:45–54.
421 14. Alanne S, Roine RP, Räsänen P, Vainiola T, Sintonen H. Estimating the minimun
422 important change in the 15D scores. Qual Life Res. 2015;24(3):599-606.

15. Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. Pain 1983;17:45–56.

16. Cicchetti D, Bronen R, Spencer S, Haut S, Berg A, Oliver P, Tyrer P. Rating scales, scales of measurement, issues of reliability: resolving some critical issues for clinicians and researchers. J Nerv Ment Dis 2006;194(8):557-64.

17. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16(3):297-334.

18. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res. 1995;4:293–307.

19. Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. J Pers Assess 2003;80:99-103.

20. Streiner DL. Health measurement scales. A practical guide to their development and use 3rd ed. New York: Oxford Medical Publications; 2003. 7p.

21. Beurskens AJ, de Vet HC, Koke AJ, van der Heijden GJ, Knipschild PG. Measuring the functional status of patients with low back pain. Assessment of the quality of four disease-specific questionnaires. Spine. 1995;20:1017–28.

22. Vaz S, Falkmer T, Passmore AE, Parsons R, Andreou P. The case for using the repeatability coefficient when calculating test-retest reliability. PLoS One. 2013;8(9):e73990.

23. Neljän elämänlaatumittarin sisältöanalyysi ICF-luokituksen viitekehyksessä. Journal of Social Medicine. 2009;46(3):196-207.

24. Hawthorne G, Richardson J, Day NA. A comparison of the assessment of quality of life (AQoL) with four other generic utility instruments. Ann Med 2001;33: 358–370.

25. Moock J, Kohlmann T. Comparing preference-based quality-of-life measures: results from rehabilitation patients with musculoskeletal, cardiovascular, or

448 psychosomatic disorders. Qual Life Res 2008;17:485–495.
449 26. Richardson J, Iezzi A, Khan MA, Chen G, Maxwell A. Measuring the Sensitivity
450 and Construct Validity of 6 Utility Instruments in 7 Disease Areas. Med Decis Making.
451 2016;36:147-59.
452 27. Eskelinen E, Lepäntalo M, Hietala EM, Sell H, Kauppila L, Mäenpää I, Pitkänen J,
453 Salminen-Peltola P, Leutola S, Eskelinen A, Kivioja A, Tukiainen E, Lukinmaa A,
454 Brasken P, Railo M. Lower limb amputations in Southern Finland in 2000 and trends up
455 to 2001. Eur J Vasc Endovasc Surg. 2004;27(2):193-200.

456

457

458 Table 1. Predefined hypotheses and their confirmation of rejection.

	Statistical Method	Rejected/ Confirmed
Reliability		
The floor and ceiling values are ≤ 15%	Max or min scores in %	0/10*
Internal consistency 0.80-0.90	Cronbach's alpha	4/6*
Criterion validity		
Moderate correlation between time of amputation and beginning of prosthesis use	Pearson	10/0*
Moderate correlation with general pain	Pearson	3/7*
Moderate correlation with general health	Pearson	4/6*
Convergent validity		
Large correlation between the 15D and	Standardized regression	
Ambulation	coefficients β.	Confirmed
Well-being		Confirmed

459 *Presents the number of confirmed and rejected hypotheses for all the 10 scales. β, beta.

Table 2. Participants' sociodemographic and clinical characteristics.

Characteristics	N = 122
Men, n (%)	76 (62.3)
Age, years, mean (SD; range)	63.7 (13.9;19-93)
Time since amputation, years, median (IQR)	4.6 (6.0)
Level of amputation, n (%)	
Transtibial	81 (66.4)
Transfemoral	41 (33.6)
Bilateral amputation, n (%)	11 (9.0)
Indication for amputation, n (%)	
Vascular disease	36 (29.5)
Trauma	25 (20.5)
Infection	17 (13.9)
Cancer	14 (11.5)
Other	30 (24.6)
Patient-reported comorbidities, n (%)	
Diabetes	44 (36.0)
Vascular disease	41 (33.6)
Hypertension	39 (32.0)
Heart disease	29 (23.8)
Neurological disease	10 (8.2)
Respiratory disease	5 (4.1)
Other	52 (42.6)
General Health, VAS, mm, mean (SD)	35.71 (23.7)
General Pain, VAS, mm, mean (SD)	34.0 (25.9)
15D, mean score (SD)	0.820 (0.125)

IQR, interquadral range; SD, standard deviation;

VAS, visual analogue scale

Table 3. Mean scores, floor and ceiling effects and the internal consistency of each of the scales at first administration.

	Items	Response Rate (%)	Mean Score (SD)	Score Range	Floor Effect (%)	Ceiling Effect (%)	Internal Consistency (95% CI)*
Prosthesis function							
Usefulness	8	100	64 (19)	7-95	0	0	0.87 (0.83 to 0.92)
Residual Limb Health	6	100	60 (22)	10-98	0	0	0.79 (0.70 to 0.89)
Appearence	5	100	62 (21)	4-99	0	0	0.67 (0.52 to 0.82)
Sounds	2	98	66 (27)	5-100	0	2	0.82 (0.71 to 0.92)
Mobility							
Ambulation	8	100	52 (28)	0-96	1	0	0.96 (0.95 to 0.97)
Transfers	5	100	66 (25)	1-99	0	0	0.81 (0.75 to 0.88)
Psychosocial experience							
Perceived Responses	5	100	83 (17)	14-100	0	5	0.69 (0.55 to 0.83)
Frustration	2	96	65 (30)	2-100	0	3	0.85 (0.76 to 0.93)
Social Burden	3	98	67 (25)	3-100	0	2	0.75 (0.65 to 0.84)
Well-being							
Well-being	2	99	66 (23)	3-100	0	1	0.80 (0.68 to 0.91)

*Expresses the expected maximum size of 95% of the absolute differences between paired observations. 95% CI obtained by bias corrected and accelerated bootstrapping.

Table 4. The change between the two measurements and reproducibility of each separate PEQ scales.

	Change From First to Second Measurement	Reproducibility	
	Mean (95% CI)	ICC (95% CI)*	CR (95% CI)**
Prosthesis function			
Usefulness	0.6 (-1.2 to 2.3)	0.87 (0.82 to 0.91)	19 (17 to 23)
Residual Limb Health	2.1 (0.4 to 4.7)	0.80 (0.73 to 0.86)	28 (24 to 31)
Appearence	0.9 (-1.1 to 3.0)	0.85 (0.79 to 0.89)	22 (19 to 27)
Sounds	1.7 (-1.6 to 4.9)	0.80 (0.72 to 0.86)	34 (28 to 40)
Mobility			
Ambulation	1.9 (-0.7 to 4.5)	0.87 (0.82 to 0.91)	28 (23 to 34)
Transfers	1.5 (-1.0 to 4.0)	0.83 (0.77 to 0.88)	27 (22 to 35)
Psychosocial experience			
Perceived Responses	0.0 (-1.9 to 2.0)	0.78 (0.70 to 0.84)	21 (16 to 26)
Frustration	0.4 (-3.0 to 3.9)	0.81 (0.73 to 0.86)	36 (30 to 43)
Sosial Burden	2.0 (-0.9 to 4.8)	0.79 (0.71 to 0.85)	31 (26 to 35)
Well-being			
Well-being	0.4 (-2.6 to 3.0)	0.79 (0.71 to 0.85)	28 (24 to 32)

ICC, intraclass correlation coefficient; CR, coefficient of repeatability. *Obtained by one-way random-effects model with absolute agreement. **Expresses the expected maximum size of 95% of the absolute differences between paired observations. 95% CI obtained by bias corrected and accelerated bootstrapping.

Table 5. PEQ correlation with age, time between prosthesis and the assessment, and general pain and health on visual analogue scale.

PEG Scale	Age	Time Since Amputation	General Pain	General Health
Prosthesis function				-
Usefulness	-0.28*	0.05	-0.39***	-0.40***
Residual Limb Health	0.23	0.00	-0.30**	-0.25
Appearence	0.15	-0.05	-0.23	-0.17
Sounds	0.24	-0.18	-0.11	-0.18
Mobility				
Ambulation	-0.27*	0.18	-0.44***	-0.48***
Transfers	-0.19	0.14	-0.40***	-0.40***
Psychosocial experience				
Perceived responses	-0.04	0.17	-0.45***	-0.42***
Frustration	0.04	0.10	-0.10	-0.05
Sosial Burden	-0.23	0.19	-0.40***	-0.38***
Well-being				
Well-being	-0.12	0.11	-0.48***	-0.43***

*p<0.05; **p<0.001; p<0.0001; statistical significance calculated using Sidak-adjusted probabilities.

Figure 1. Predictors of the 15D age- and gender-standardized regression coefficients β. Values 0.10, 0.30 and 0.50 represent small, moderate and large correlations, respectively. The box plot indicates mean values and the whiskers represent standard deviations.