Reliability and validity of the Finnish version of the Lower Extremity Functional Scale (LEFS)

Jussi P. Repo,[1] Erkki J. Tukiainen,[1] Risto P. Roine,[2,3] Outi Ilves,[4] Salme Järvenpää,[5] Arja Häkkinen[4,5]

[1] Department of Plastic Surgery, University of Helsinki and Helsinki University Hospital, HUS, Finland

[2] Group Administration, University of Helsinki and Helsinki University Hospital, HUS, Finland

[3] Department of Health and Social Management, Research Centre for Comparative Effectiveness and Patient Safety, University of Eastern Finland, Kuopio, Finland

[4] Department of Health Sciences University of Jyväskylä, Jyväskylä, Finland

[5] Department of Physical Medicine, Central Finland Health Care District, Jyväskylä, Finland

Address for correspondence: Jussi P. Repo, Department of Plastic Surgery, Helsinki University Hospital, P.O. Box 266, 00029 HUS, Finland (e-mail: mrjussirepo@gmail.com) Tel. +358 44 359 3100, Fax: +358 9 47178580

Keywords: Lower Extremity Functional Scale, LEFS, reliability, validity, Finnish

Abstract

Purpose The present study aimed to assess the psychometric properties of the Finnish version of the Lower Extremity Functional Scale (LEFS) among foot and ankle patients.

Methods The LEFS was translated and cross-culturally adapted to Finnish. We assessed the test-retest reliability, internal consistency, floor-ceiling effect, construct validity and criterion validity in patients who underwent surgery due to musculoskeletal pathology of the foot and ankle (*N* = 166).

Results The test-retest reliability was high (ICC = 0.93, 95% CI: 0.91-0.95). The standard error of measurement was 4.1 points. The Finnish LEFS showed high internal consistency (Cronbach's α = 0.96). A slight ceiling effect occurred as 17% achieved the maximum score. The LEFS correlation was strong with the 15D Mobility dimension (r = 0.74) and overall HRQoL (r = 0.66), pain during foot and ankle activity (r = -0.69) and stiffness (r = -0.62). LEFS correlated moderately with foot and ankle pain at rest (r = -0.50) and with physical activity (r = 0.46).

Conclusions The Finnish version of the LEFS showed reliability and validity comparable to those of the original version. This study indicates that the Finnish version of the LEFS serves both clinical and scientific purposes in assessing lower-limb function.

Introduction

The literature describes around 140 different rating scales for assessing foot and ankle function [1]. In evidence-based medicine and rehabilitation, the use of patient-reported outcome measures (PROMs) has gained popularity in assessing the benefits and disadvantages of various treatment options. However, opinions continue to vary regarding which rating system to use [1-3]. A valid PROM for assessing foot and ankle function in the Finnish-speaking population is in high demand.

The English-language Lower Extremity Functional Scale (LEFS) [4] comprises 20 function-related items assessing the impairment of the lower-extremity musculoskeletal system in everyday activities. Many have lauded its superior validity over many of the more frequently used foot and ankle rating scales [1, 2]. Furthermore, previous research has validated the LEFS in several other languages [5-11].

Until now, the Finnish LEFS remains unvalidated. The present study aimed to translate and cross-culturally adapt the LEFS into the Finnish language and to assess the psychometric properties of the Finnish LEFS in foot and ankle patients.

Methods

Translation and cross-cultural adaptation

The developer of the Lower Extremity Functional Scale (LEFS) (Supplementary material) granted us permission to translate and cross-culturally adapt the questionnaire into Finnish [4]. The translation and cross-cultural adaptation followed the International Society of Pharmacoeconomics and Outcome Research (ISPOR) guidelines [12]. Two separate translators with Finnish as their first language produced the forward-translations of the LEFS from English to Finnish. The different metric systems involved in the forward-translation process revealed cultural differences in items 11 (“Walking two blocks“) and 12 (“Walking two miles“). The key in-country person, the other translator and the project manager addressed the cultural differences encountered and reconciled them in the forward-translations, resulting in the first Finnish version and a written report.

A back-translation from Finnish into English by a native English-speaking translator fluent in Finnish served to detect any flaws in content. The translator had no medical background and no knowledge of the original LEFS, but was familiar with Finnish culture. A review of the back-translation conducted by the project manager and the key in-country person served to identify any discrepancies between the back-translation and the original LEFS.

An expert committee consisting of the project manager, a specialist in internal medicine, two surgeons and a general physician compared the forward and backward translations to each other and to the original questionnaire as well as reviewed the translation reports. This phase produced a pre-final version of the Finnish LEFS and a written protocol report.

We conducted a pilot-study of the pre-final version on 20 patients who underwent lower-extremity surgery and had Finnish as their first language. Each patient was cognitively debriefed in a manner adhering to the European Organisation for Research and Treatment of Cancer (EORTC) guidelines [13] to identify any offensive content, difficulties in anwering or in understanding the questions, and whether they should ask any questions differently.

In the final phase, the project manager and the key in-country person reviewed the pilot-testing and cognitive debriefing results, and finalized the questionnaire. A Finnish language expert of the Finnish medical association (Duodecim) then proofread this version, after which the project manager later introduced the final version (Supplementary material) in a final report.

Participants for psychometrics

The Ethics Committee approved the study. We recruited for the study consecutive patients who underwent surgery at the authors' institution. We verified the clinical data from the patient records, including diagnostic and procedure codes. We excluded from the study patients under age 18. Inclusion criteria were a surgically treated foot and ankle pathology, the patient's full ability to understand written Finnish and completion of all questionnaires. Altogether 769 patients were approached by mail cross-sectionally. The participants in the study provided their written consent in accordance with the principles of the Declaration of Helsinki.

Questionnaires

The Lower Extremity Functional Scale (LEFS)

The LEFS is a 20-item PROM developed to assess the musculoskeletal function of the lower-limb [4]. The LEFS is scored on a five-point scale (0 to 4 = worst to best); the total value ranges from 0 to 80 points with higher scores indicating a higher functional level. The minimum detectable change and the minimum clinically important difference are both estimated at around nine points [4, 9, 14].

The psychometric properties of the LEFS in foot and ankle patients have proven superior to many other, more widely used, foot and ankle rating scales [1]. A computerized adaptive testing of the LEFS has revealed a high ability to identify impairment in the ankle, foot, hip or knee [3]. The scale has proven reliable, responsive and valid in assessing foot and ankle function [4, 16, 17]. The LEFS also has high content and construct validity [4].

15D health-related quality of life (HRQoL) instrument

The 15D instrument is a valid, reliable, comprehensive, generic HRQoL instrument [18]. A total score of 15 contains 15 dimensions: moving, seeing, hearing, breathing, sleeping, eating, speech, excretion, usual activities, mental function, discomfort and symptoms, depression, distress, vitality and sexual activity. Of these, we selected four dimensions that assess related function and activities (Mobility, Usual activities, Discomfort and symptoms, and Vitality) in order to assess the validity of the criteria due to the similarity of their content to the LEFS items.

Respondents choose one of the five levels in each dimension that best describes their current state of health (1 = the best; 5 = the worst). The 15D produces both a HRQoL profile and a single index score representing overall HRQoL. The 15D compares favorably with other popular, generic HRQoL instruments in their most important properties [18-21].

Physical activity

The FIT index of Kasari [22], which measures current leisure time physical activity, consists of three questions: 1) Frequency (< once per month, a few times per month, 1 or 2 times per week, 3 to 5 times per week, ≥ 6 times per week), 2) Intensity (light aerobic excercise, low to moderate, moderate, moderately high, and high intensity excercise), and 3) Time spent per workout (< 10 minutes, 10-20 minutes, 20-30 minutes and > 30 minutes). The score range is 1 to 100; the points < 36, 37-63 and 64 indicate low, moderate or high physical activity, respectively.

Overall health, pain and stiffness

Patients reported their general state of health during the previous week on a Visual Analogue

Scale (VAS) ranging from 0 to 100 mm (0 = the best imaginable state; 100 = worst possible state). The Visual Analogue Scale has previously been validated for pain assessment [23]. The VAS also served as a single-item measure to capture subjective feelings concerning foot and ankle pain during activity and at rest, as well as foot and ankle stiffness.

In addition, we obtained information on patient age and sex, weight, height, smoking habits, occupation, and employment status (employed, unemployed, on sick leave, retired, other). Clinical data consisted of information on surgery and time since most recent surgery.

LEFS validation course

We sent the questionnaires to the patients by mail. In addition to the sociodemographic data, the first questionnaire package included the Finnish LEFS, 15D, FIT index, overal health, pain and stiffness questionnaires. Participants returned the questionnaires together with their signed informed consent. Those who failed to return the first questionnaire package within a week received a reminder.

Test-retest reliability

We mailed a second questionnaire package to the patients two weeks after they had completed the first questionnaire. This package included the LEFS questionnaire and a survey assessing whether the patients' health status had changed between the two test periods. The time between the first and the second questionnaires was on average 2.5 weeks. A period of two weeks between the assessments has previously been claimed to be optimal in the test-retest evaluation [24].

Statistics

The data appear as means with standard deviation (SD) or 95% confidence intervals (95% CI), percentages and ranges. A two-way mixed model with absolute agreement served to measure relative reliability (Intraclass Correlation Coefficience, ICC). We calculated absolute reliability and the standard error of measurement (SEM) as the root mean square error term of the analysis of variance. The degrees of freedom associated with the estimated residual variance and percentage points from the corresponding chi-square distributions served to obtain confidence intervals for the SEM [25]. We estimated internal consistency by calculating Cronbach's alpha with bootstrapped 95% CIs. The floor and ceiling values were assessed by dividing the amount of participants receiving the maximum or minimum scores, respectively, by the total amount of participants.

Maximum likelihood factor analysis with oblimin rotation served to study construct validity for the LEFS items matrix of polychoric correlations. In the item analysis, the Spearman correlation coefficient served to estimate the discriminatory power (i.e., corrected item-total correlation). The Spearman method served to calculate the correlation coefficients with the following criteria for the correlation values according to Andresen: weak <0.30, moderate 0.30- 0.59, strong ≥0.60 [26]. We used bias-corrected bootstrapping to obtain the confidence intervals for the mean changes and correlation coefficients, and the bootstrapped type t-test for independent samples to test for differences between groups (surgical intervention in either foot or ankle).

Predefined hypotheses are presented in table 1. IBM SPSS Statistics, version 23.0, and R version 3.0.1 served to conduct the statistical analyses.

Results

We included a total of 166 patients for the study (Figure 1). Participants' mean age was 55 (SD 16) years, and 53% were women. The participants' detailed characteristics appear in table 1. Of the 166 participants, the indication for primary surgery on the foot or ankle was trauma in 157 participants, infection in six, en bloc tumor resection in two, and a stress fracture and destruction of the tibiotalar joint in one.

Reliability

Floor-ceiling effect

The LEFS questionnaires showed no floor-effect (0 score). When completing the LEFS for the first time, 29 patients (17%) scored the maximum total value, and 15% reached the maximum number of points in both assessments. In one group, 8% scored the maximum number of points in the first assessment completed less than one year after surgery.

Test-retest reliability

The mean value of the LEFS score at measurement one was 66.2. The mean change by measurement two was -0.5 points. The test-retest reliability with ICC was 0.93 (95% CI, 0.91 to 0.95). The value of more than 0.90 confirmed our predefined hypothesis. The standard error of the measurement (SEM) was 4.1. When we grouped the participants according to the time since surgery, the ICC was smaller and the SEM higher in the group with a shorter time since surgery (table 2).

Internal consistency and item analysis

Cronbach's alpha for the LEFS total score revealed a high internal consistency of 0.96 (95% CI, 0.94 to 0.97) confirming our predefined hypothesis. In patients completing the questionnaires within one year after surgery, the ICC for LEFS was 0.96 (95% CI, 0.93 to 0.98), whereas in those completing the questionnaires more than one year after surgery, the corresponding figure was 0.95 (95% CI, 0.92 to 0.96). Item analysis for each of the LEFS items appears in figure 2. All items showed at least moderate item correlation. Items 16 to 19 had the highest corrected item correlation. In these items, the mean value (item difficulty) was close to two. The lowest corrected item correlation was in item 20, followed by items 8, 5, and 4. In these items, almost all participants attained the highest score. The median (IQR) of all items together was 4.0 (3.4).

Construct validity

Factor analysis performed for the construct validity showed that the LEFS scale loaded on two factors (table 3). The first factor included questions about daily activities. The other factor loaded items 16 to 19. These items were distinctly related to function. The latter factor explained 83% of the total variance.

Criterion validity

Spearman's correlation coefficients between the LEFS total score and the reference outcomes appear in table 4. Eight out of nine predefined hypotheses were met (table 1). Correlation with the 15D total index was strong (r = 0.66). Strong correlation with the 15D Mobility (r = 0.74) and Usual activities (r = 60) dimensions was also found. However, in the dimensions of Discomfort and symptoms and Vitality the correlation was moderate. Foot and ankle pain at rest (r = -0.50) or during activity (r = -0.69) as well as stiffness (r = 0.62) showed a moderate to strong correlation. LEFS correlated poorly with age (r = -0.25) and BMI (r = -0.24). However, correlation with physical activity was moderate (r = 0.46).

The location of the surgical intervention did not affect the LEFS score (p = 0.57). Ankle surgery patients scored a mean of 65.9 (SD 15.8) points, and foot patients, 67.6 (SD 13.5) points.

Discussion

In the present study, we assessed the cross-cultural adaptation and the psychometric properties of the LEFS questionnaire translated into the Finnish language. We demonstrated that the Finnish version of the LEFS instrument showed test-retest reliability, internal consistency, construct and criterion validity comparable to that of the original version published in English.

The mean score of the LEFS was 66. In a LEFS study of foot and ankle function by Lin et al. [16], the mean score value ranged from 29 to 72, depending on the location of the surgery and the time between surgery and the assessment. The overall score was between 29 and 34 points in the acute phase four to six weeks after fracture and 66 to 72 points 24 to 26 weeks after ankle fracture. The results indicate that the LEFS is a useful tool for monitoring limitations of activity in people with foot and ankle fractures starting from the early postoperative phase.

Linguistic and cultural adaptation

The translation and cross-cultural adaptation followed best practice guidelines [12]. We encountered cultural differences in the adaptation process.

Because Finland does not use British imperial units of measurement, the harmonization committee consented to changing the British Imperial system to the metric system in item 12 by changing “Walking a mile” to “Walking 2 km”. Although, strictly speaking, a mile is equivalent to 1.6 km, we considered an even number of 2 km more suitable. In previously published literature, authors converted the length to a value between 1 and 1.6 km [5, 6, 8].

In addition, because the Finnish language does not use city blocks to describe distances, item 11 was modified to “Walking 200 meters”. Hoogeboom and colleagues used “250 meters”, Cacchio “10 minutes” and Hou et al. “500 m” to describe the distance of two city blocks [5, 6, 8].

Reliability of the Finnish version of the LEFS

A floor-ceiling effect of 15% or less is sufficient for acceptable psychometrics [27]. Previous reports showed no floor-ceiling effect has [5, 7, 11, 16]. Our study detected a ceiling-effect in 17% of our participants, including over 60% of those who had undergone surgery more than a year earlier. The ceiling-effect was lower (23% vs 8%), however, in those who had undergone surgery during the preceding 12 months. The ceiling-effect may be due to the fact that most patients had underwent the surgical procedure more than a year ago. This may result in higher scores. On the other hand, assessing test-retest reliability with several-day intervals in acute patients would be impossible because the healing process would likely skew the results. However, the results of this study suggest that the use of the LEFS is best applied during the acute phase.

Previous researchers have shown the internal consistency of the LEFS to be high (0.94-0.99) [4-6, 9, 10, 11]. The total LEFS score of the present study had a Cronbach’s alpha of 0.96 confirming our predefined hypothesis. Further, a Cronbach’s alpha of 0.8 or more is considered sufficient [28]. On the other hand, Cronbach's alpha values over 0.9 may indicate item repetition [29]. Users should note this when using the LEFS.

In the present study, the test-retest reliability was high (ICC 0.93). The results indicate that the symptoms did not change during the 2.5 weeks between the assessments. Previous studies have shown test-retest reliability values of 0.86-0.98 [4, 5, 7]. The high ICC value in this study may be due to the fact that 62% of the participants had undergone surgery earlier than during the preceding year. In the long-term, the recovery process is slower or healing achieves maximal recovery, which affects the test-retest results, because the impairment status may no longer show any improvement. On the other hand, the ICC value alone may not accurately reflect the acceptability of the measurement. In addition to the ICC values, calculating the SEM values is also important, as the SEM value is the standard error of the measurement and conveys the variation in the actual units of measurements. SEM estimates the precision and uncertainty of how the study sample represents the underlying population. The smaller the SEM value, the higher the agreement. The mean SEM value of 4.1 in our study is in accordance with the SEM values of 3.5 to 4.0 reported previously [9, 13].

Validity of the Finnish version of the LEFS

The construct validity analysis revealed two main dimensions. Items from 16 to 19 included extensive foot and ankle movement and loaded on a different factor. Recent studies have shown similar results with LEFS loading on two factors [7, 11]. Hoogeboom and colleagues found in their study that the LEFS loaded on three factors, however, one factor represented less than 10% of the total variance and was omitted resulting in two distinct factors [5]. The item correlation in the present study also showed that these four items scored differently. In clinical practice items loading on factor 2 may have more specificity in assessing the foot and ankle function.

Criterion validity demonstrates the ability of the instrument to measure what it is supposed to measure compared with at least one validated instrument. The LEFS and the 15D total scores showed strong correlation, indicating that these measurement procedures may partly measure the same construct. The HRQoL comparison of the present study may not be directly proportional and comparable to those of the previous studies. This is due to the fact that previous authors have used the SF-36 instrument to measure the HRQOL [7, 8, 11, 17]. Previous studies have found strong correlation (0.77-0.83) with the SF-36 Physical Component [7, 8, 11, 17]. Nevertheless, the results of the present study seem to concur with the published literature [5, 7, 8, 11]: the 15D dimensions of Mobility and Usual activities showed strong correlation with the LEFS.

The LEFS score showed strong and negative correlation with or foot and ankle pain during activity and ankle stiffness, respectively, probably because the LEFS uses highly generic items to assess lower-extremity function. Correlation with the foot and ankle pain at rest was moderate indicating that the LEFS does not measure foot and ankle musculoskeletal pathology at rest.

Finally, age and BMI correlated negatively and poorly with the LEFS, indicating that slightly higher disability may accompany with higher age or BMI. A study by Lentz et al. also showed poor correlation with age (r = -0.25) and BMI (r = -0.24) in 85 patients with foot and ankle disability [30]. Further, a moderate correlation between the total LEFS score and physical activity could be due to LEFS items focusing more on function-related problems and not reflecting the patient's physical activity.

In the present study, we noted no significant difference in the location of surgery (foot vs. ankle) in patients who had undergone surgery.

To the authors' knowledge, the present Finnish LEFS validation study has a representative study population of foot and ankle patients who have undegone surgery, a large proportion (95%) of whom underwent injury-related surgery. Thus, clinicians and reserchers should exercise caution when generalizing the suitability of the LEFS to other lower-limb problems.

One limitation of this study was that most of the participating patients had a rather long period since surgery, so we calculated the validity and reliability of the Finnish LEFS based on material whose mean score was likely skewed towards the maximum score, which may have affected our ICC and SEM values.

The outcomes of our study confirmed the reliability and validity of the Finnish version of the LEFS, and that it serves well in assessing lower-limb musculoskeletal function. Although the LEFS can serve in both clinical work and research, its properties seem more suitable when the follow-up to detect changes in foot and ankle function already begins in the acute phase.

Declarations of Interest

The authors report no declarations of interest.

References

[1] Hunt KJ, Hurwit D. Use of patient-reported outcome measures in foot and ankle research. J Bone Joint Surg Am 2013;95:e118(1-9).

[2] Button G, Pinney S. A meta-analysis of outcome rating scales in foot and ankle surgery: is there a valid, reliable, and responsive system? Foot Ankle Int 2004;25:521-5.

[3] Pinsker E, Daniels TR, Inrig T, Warmington K, Beaton DE. The ability of outcome questionnaires to capture patient concerns following ankle reconstruction. Foot Ankle Int

2013;34:65-74.

[4] Binkley JM, Stratford PW, Lott SA, Riddle DL. The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. North American Orthopaedic Rehabilitation Research Network. Phys Ther 1999;79:371-83.

[5] Hoogeboom TJ, de Bie RA, den Broeder AA, van den Ende CH. The Dutch Lower Extremity Functional Scale was highly reliable, valid and responsive in individuals with hip/knee osteoarthritis: a validation study. BMC Musculoskelet Disord 2012;13:117.

[6] Cacchio A, De Blasis E, Necozione S, Rosa F, Riddle DL, di Orio F, De Blasis D, Santilli V. The Italian version of the lower extremity functional scale was reliable, valid, and responsive. J Clin Epidemiol 2010;63:550-7.

[7] Naal FD, Impellizzeri FM, Torka S, Wellauer V, Leunig M, von Eisenhart-Rothe R. The German Lower Extremity Functional Scale (LEFS) is reliable, valid and responsive in patients undergoing hip or knee replacement. Qual Life Res 2015;24:405-10.

[8] Hou WH, Yeh TS, Liang HW. Reliability and validity of the Taiwan Chinese version of the Lower Extremity Functional Scale. J Formos Med Assoc 2014;113:313-20.

[9] Alnahdi AH, Alrashid GI, Alkhaldi HA, Aldali AZ. Cross-cultural adaptation, validity and reliability of the Arabic version of the Lower Extremity Functional Scale. Disabil Rehabil. 2015:1-8. [Epub ahead of print]

[10] Cruz-Diaz D, Lomas-Vega R, Osuna-Pérez MC, Hita-Contreras F, Fernández ÁD, Martínez-Amat A. The Spanish lower extremity functional scale: a reliable, valid and responsive questionnaire to assess musculoskeletal disorders in the lower extremity. Disabil Rehabil 2014;36:2005-11.

[11] Negahban H, Hessam M, Tabatabaei S, Salehi R, Sohani SM, Mehravar M. Reliability and validity of the Persian lower extremity functional scale (LEFS) in a heterogeneous sample

of outpatients with lower limb musculoskeletal disorders. Disability and rehabilitation. 2014;36(1):10-5.

[12] Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, Erikson P. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. Value Health 2005;8:94-104.

[13] Dewolf L, Koller M, Velikova G, Johnson C, Scott N, Bottomley A. EORTC Quality of life group translation procedure (3rd ed.); 2009.

[14] Alcock AG, Stratford PW. Validation of the Lower Extremity Functional Scale on athletic subjects with ankle sprains. Psysiother Can 54:233-40.

[15] Hart D, Mioduski JE, Stratford PW. Simulated computerized adaptive tests for measuring functional status were efficient with good discriminant validity in patients with hip, knee, or foot/ankle impairments. J Clin Epidemiol 2005;58:629-38.

[16] Lin CW, Moseley AM, Refshauge KM, Bundy AC. The lower extremity functional scale has good clinimetric properties in people with ankle fracture. Phys Ther 2009;89:580-8.

[17] Pan SL, Liang HW, Hou WH, Yeh TS. Responsiveness of SF-36 and Lower Extremity Functional Scale for assessing outcomes in traumatic injuries of lower extremities. Injury 2014;45:1759-63.

[18] Sintonen H. The 15D instrument of health-related quality of life: properties and applications. Ann Med 2001;33:328–36.

[19] 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;33:358–70.

[20] Stavem K. Reliability, validity and responsiveness of two multiattribute utility measures

in patients with chronic obstructive pulmonary disease. Qual Life Res 1999;8:45–54.

[21] Moock J, Kohlmann T. Comparing preference-based quality-of-life measures: results from rehabilitation patients with musculoskeletal, cardiovascular, or psychosomatic disorders. Qual Life Res 2008;17:485–95.

[22] Kasari DS. The effects of exercise on serum lipid levels in college women. Unpublished Master's Thesis. University of Montana, Missoula; 1976:46.

[23] 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.

[24] Beurskens AJ, de Vet HC, Koke AJ, et al. Measuring the functional status of patients with low back pain. Assessment of the quality of four disease-specific questionnaires. *Spine* 1995;20:1017–28.

[25] Milliken GA, Johnson DE. Analysis of Messy Data. 1, Designed Experiments (2nd ed.). Boca Raton: CRC Press; 2009.

[26] Andresen EM. Criteria for assessing the tools of disability outcomes research. Arch Phys Med Rehabil 2000;81:15-20.

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

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

[29] Streiner DL. Health measurement scales. A practical guide to their development and use (3rd ed.). New York: Oxford Medical Publications; 2003.

[30] Lentz TA, Sutton Z, Greenberg S, Bishop MD. Pain-related fear contributes to self-reported disability in patients with foot and ankle pathology. Arch Phys Med Rehabil

2010;91:557-61.

Figure 1. Flow chart for patient recruitment and inclusion.

Figure 2. Item analysis for each of the LEFS items. Numbers indicate the corresponding items in the LEFS questionnaire.