

## Lower Limb Functional Index: Development and Clinimetric Properties

Charles Philip Gabel, Markus Melloh, Brendan Burkett, Lori A. Michener

C.P. Gabel, MSc, Faculty of Science, Health and Education, Centre for Healthy Activities, Sport and Exercise, University of the Sunshine Coast, PO Box 760, Coolool Beach, Queensland, Australia 4573. Address all correspondence to Mr Gabel at: cp.gabel@bigpond.com.

M. Melloh, MD, MPH, Department of Orthopaedic Surgery, Dunedin School of Medicine, University of Otago, Dunedin, New Zealand.

B. Burkett, PhD, Faculty of Science, Health and Education, Centre for Healthy Activities, Sport and Exercise, University of the Sunshine Coast.

L.A. Michener, PT, PhD, ATC, Clinical Orthopedic and Sports Outcomes Research Laboratory, Department of Physical Therapy, Virginia Commonwealth University-MCV Campus, Richmond, Virginia.

[Gabel CP, Melloh M, Burkett B, Michener LA. Lower Limb Functional Index: development and clinimetric properties. *Phys Ther*. 2012;92:98-110.]

© 2012 American Physical Therapy Association

Published Ahead of Print:  
November 3, 2011  
Accepted: July 26, 2011  
Submitted: June 6, 2010

**Background.** Existing lower-limb, region-specific, patient-reported outcome measures have clinimetric limitations, including limitations in psychometric characteristics (eg, lack of internal consistency, lack of responsiveness, measurement error) and the lack of reported practical and general characteristics. A new patient-reported outcome measure, the Lower Limb Functional Index (LLFI), was developed to address these limitations.

**Objective.** The purpose of this study was to overcome recognized deficiencies in existing lower-limb, region-specific, patient-reported outcome measures through: (1) development of a new lower-extremity outcome scale (ie, the LLFI) and (2) evaluation of the clinimetric properties of the LLFI using the Lower Extremity Functional Scale (LEFS) as a criterion measure.

**Design.** This was a prospective observational study.

**Methods.** The LLFI was developed in a 3-stage process of: (1) item generation, (2) item reduction with an expert panel, and (3) pilot field testing (n=18) for reliability, responsiveness, and sample size requirements for a larger study. The main study used a convenience sample (n=127) from 10 physical therapy clinics. Participants completed the LLFI and LEFS every 2 weeks for 6 weeks and then every 4 weeks until discharge. Data were used to assess the psychometric, practical, and general characteristics of the LLFI and the LEFS. The characteristics also were evaluated for overall performance using the Measurement of Outcome Measures and Bot clinimetric assessment scales.

**Results.** The LLFI and LEFS demonstrated a single-factor structure, comparable reliability (intraclass correlation coefficient [2,1]=.97), scale width, and high criterion validity (Pearson  $r=.88$ , with 95% confidence interval [CI]). Clinimetric performance was higher for the LLFI compared with the LEFS on the Measurement of Outcome Measures scale (96% and 95%, respectively) and the Bot scale (100% and 83%, respectively). The LLFI, compared with the LEFS, had improved responsiveness (standardized response mean=1.75 and 1.64, respectively), minimal detectable change with 90% CI (6.6% and 8.1%, respectively), and internal consistency ( $\alpha=.91$  and  $.95$ , respectively), as well as readability with reduced user error and completion and scoring times.

**Limitations.** Limitations of the study were that only participants recruited from outpatient physical therapy clinics were included and that no specific conditions or diagnostic subgroups were investigated.

**Conclusion.** The LLFI demonstrated sound clinimetric properties. There was lower response error, efficient completion and scoring, and improved responsiveness and overall performance compared with the LEFS. The LLFI is suitable for assessment of lower-limb function.



Post a Rapid Response to  
this article at:  
[ptjournal.apta.org](http://ptjournal.apta.org)

Patient-reported outcome measures of function are critical for assessing musculoskeletal conditions.<sup>1-3</sup> *Function* is the level of activities an individual performs to realize the needs of daily living.<sup>4</sup> Numerous lower-limb patient-reported outcome measures assess function for specific joints,<sup>5-8</sup> joint conditions,<sup>9-12</sup> or region-specific conditions.<sup>13-16</sup> However, there is limited consensus regarding which tools to use.<sup>5,8,17</sup> Consequently, a need exists for a simple, reliable, and valid tool that effectively measures patient-rated lower-limb function.<sup>18,19</sup>

Five tools<sup>20-24</sup> were designed to measure the lower limb as a single regional kinetic chain.<sup>25,26</sup> Only the Lower Extremity Functional Scale (LEFS)<sup>22</sup> and the Foot and Ankle Ability Measure (FAAM)<sup>24</sup> have detailed their clinimetric properties, but both have limitations<sup>6,17</sup> and neither have demonstrated readability, completion time, and scoring time. Clinimetric properties include psychometric, practical, and general characteristics<sup>19,26,27</sup> (eAppendix 1, available at [ptjournal.apta.org](http://ptjournal.apta.org)). The LEFS has potential excessive internal consistency<sup>22,28</sup> that may cause item redundancy.<sup>29</sup> Furthermore, sensitivity<sup>30</sup> and long-term responsiveness<sup>31</sup> are lacking, and the five-point Likert scale may increase respondent burden and scoring error.<sup>26,32</sup> The FAAM originally was proposed and developed as a region-specific, patient-reported outcome measure,<sup>24</sup> but subsequent studies used only participants with disorders below the knee.<sup>6,17,33</sup> Furthermore, these studies on the FAAM show conflicting results for internal consistency. It is either demonstrated as excessive<sup>24</sup> or not demonstrated sufficiently.<sup>6,17</sup> These aspects call into question the FAAM's ability to accurately measure lower-limb functional loss as a single kinetic unit.<sup>33</sup> Consequently, the LEFS remains the only suitable criterion

measure for lower-limb functional measurement.

The literature indicates that development of new lower-limb measurement tools should be considered.<sup>30</sup> To be widely accepted and used, any new tool should improve the clinimetric properties of advocated patient-reported outcome measures.<sup>26,27</sup> Psychometrics should be robust, assess function, and measure change over time. Practicality should improve readability, provide a user-friendly format, and minimize response errors through efficient completion and scoring processes. General characteristics should clarify that distribution is normalized without floor or ceiling effects and that constructs represent both function and quality of life. The objective of this study was to overcome recognized deficiencies in existing lower-limb, region-specific, patient-reported outcome measures through: (1) development of a new lower-extremity outcome scale, the Lower Limb Functional Index (LLFI), and (2) evaluation of the clinimetric properties of the LLFI using the LEFS as a criterion measure.

## Materials and Method

A prospective observational study of the development and clinimetric assessment of the LLFI (Appendix 1) was completed in 2 phases (Figure). Phase 1 developed the LLFI using a 3-stage process, and phase 2 validated the LLFI in patients receiving care for lower-limb musculoskeletal conditions.


### Phase 1—Development of the LLFI

Development of the LLFI followed the Kirshner and Guyatt<sup>34,35</sup> established 3-stage process.

**Stage 1—item generation.** The electronic databases PubMed, CINAHL, EMBASE, and PEDro were reviewed from 1980 to 2009 with

the key words “outcomes,” “self-report,” “function,” “disability,” “impairment,” “lower limb,” “leg,” “knee,” “hip,” “foot,” and “ankle.” This review identified 130 patient-reported outcome measures (eAppendix 2, available at [ptjournal.apta.org](http://ptjournal.apta.org)). A 4-person peer panel (occupational therapist, physical therapist, general practitioner, and occupational physician) used consensus opinion, which required consensus of a minimum of 3 panel members,<sup>34,36</sup> to review and shorten the list to 34 patient-reported outcome measures, with 873 items for lower-limb and general musculoskeletal injuries. The 873 items were further reduced to 421 items through binning and winnowing methods, which removed duplicate and non-applicable items.<sup>37-39</sup>

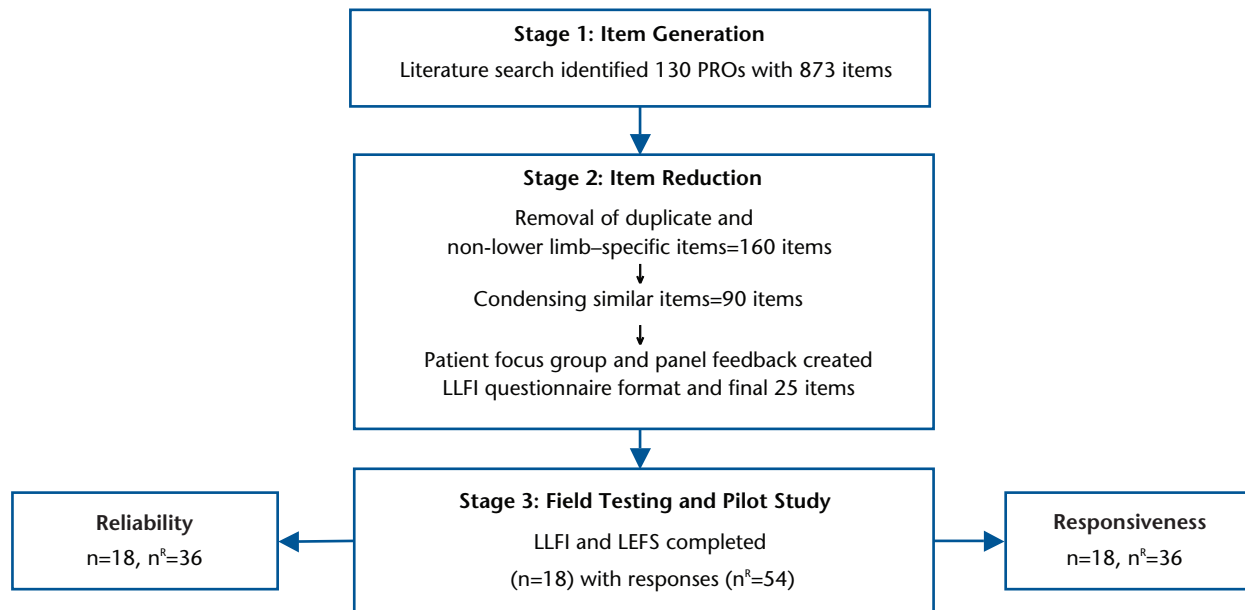
**Stage 2—item reduction.** The peer panel further reduced the 421 items in 5 substages (2a-e).<sup>40</sup> Substage 2a reduced the list to 203 items through item pooling (eg, “stairs,” “steps-up,” “steps-down,” and “slopes” became “stairs”). Substage 2b classified<sup>26,35</sup> items using the World Health Organization *International Classification of Functioning, Disability and Health* (ICF)<sup>41</sup> codes<sup>42</sup>: b=body functions, s=body structures, d=activities and participation, and e=environmental factors.<sup>43</sup> Substage 2c reduced the 203 items to 90 by combining the ICF codes to create common descriptive construct titles (eg, “stairs,” “ladders,” and “curbs” became “code


Available With This Article at [ptjournal.apta.org](http://ptjournal.apta.org)

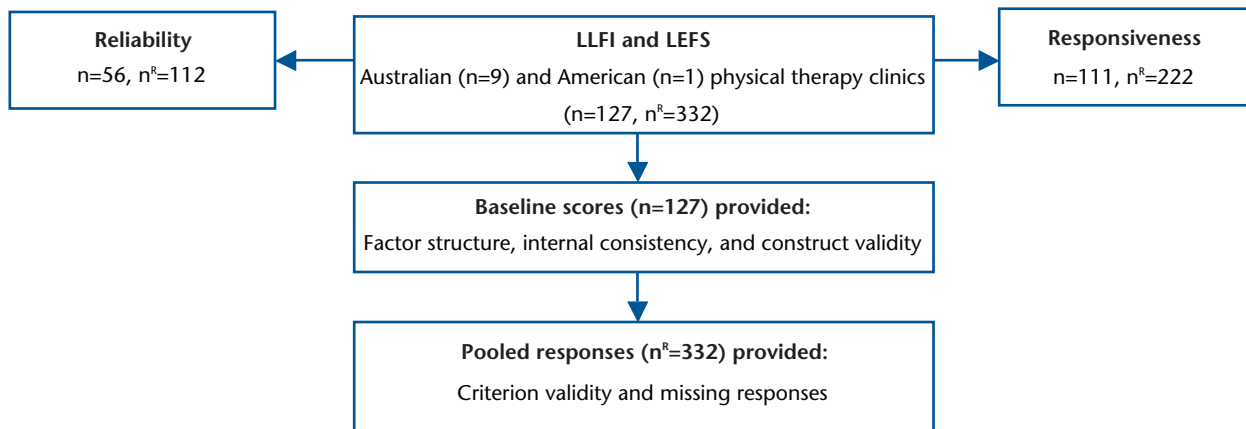
- [eAppendix 1](#): Glossary of Clinimetric Terms
- [eAppendix 2](#): List of Outcome Measures Used in Stage 1 of Development of the Lower Limb Functional Index

## Lower Limb Functional Index Development and Validation

### Phase 1: Development and Pilot Study (n=18)



### Phase 2: Validation Study (n=127)



#### Figure.

Flow chart of Lower Limb Functional Index (LLFI) development and validation. PRO=patient-reported outcome measure, LEFS=Lower Extremity Functional Scale, n=number of participants, n<sup>R</sup>=number of responses.

d4551-climbing”). Substage 2d reduced the list to 74 items by grouping and deletion (eg, “dressing,” “putting on footwear,” and “pants” were retained, but “dressing slower” and “fastening clothing” were deleted). Substage 2e further combined items via panel consensus on importance and relevance to achieve the final 25 items (15 general and 10 lower-limb specific). The ques-

tion stems then were formulated: “Due to my leg/s: I have difficulty/problems . . .” or “I stay/change/avoid/get others. . .” Final question wording required consensus of the peer panel and a focus group of 10 patients (3 with hip conditions, 3 with knee conditions, and 4 with foot and ankle conditions) for face and content validity.<sup>35</sup>

The LLFI format was based on usability for human-system interaction<sup>44</sup> and user-centered design.<sup>45</sup> Text boxing was used, which places questions within larger boxes to improve reader acceptability, and shadowing of alternate lines. The 3-item response option scale of “Yes,” “Partly,” and “No”<sup>19</sup> was selected to provide stable, equally spaced responses.<sup>46,47</sup>

**Stage 3—field testing.** A pilot investigation (n=18, total number of responses [n<sup>R</sup>]=54) used outpatients with lower-limb conditions for preliminary LLFI assessment (Tab. 1). This pilot investigation demonstrated reliability<sup>48,49</sup> (n=18, n<sup>R</sup>=36) was high (intraclass correlation coefficient [ICC (2,1)]=.97, with 95% confidence interval [CI]), responsiveness<sup>50,51</sup> (n=18, n<sup>R</sup>=36) was high (effect size=1.3, standard response mean >1.80), and no floor or ceiling effects were present. Sampling methodology was confirmed as suitable and enabled sample size calculations for the larger main study. It also identified the patient-reported outcome measures' change characteristic as heterogeneous,<sup>52</sup> with the change coefficient demonstrated by the individual patients who had true change that varied by different amounts over 2 points in time.<sup>53</sup>

**Sample size.** Minimum sample sizes for the validation study were calculated from the pilot study results, with an 80% likelihood of detecting differences and allowing for 15% attrition with *P*<.05.<sup>54,55</sup> Power calculations indicated the need for a total sample of n=120 (reliability, n=55; responsiveness, n=99; and concurrent criterion validity, n=104).<sup>56,57</sup> Exploratory factor analysis indicated a single-factor structure was likely; therefore, more than 100 participants were required.<sup>54,58</sup>

**Phase 2—Validation of the LLFI in a Cohort Population**

A prospective, cohort design was used.<sup>26</sup> Participants with lower-limb musculoskeletal conditions (n=127, n<sup>R</sup>=332) were recruited consecutively from 9 Australian and 1 American physical therapy clinics between 2003 and 2009. Inclusion criteria were medical practitioner referral and musculoskeletal lower-limb symptoms (acute, subacute, and chronic) that enabled a broad

**Table 1.**  
Participant Demographics

Variable	Phase 1: Pilot Study	Phase 2: Validation Study
No. of participants	18	127
No. of responses	54	332
Age (y), $\bar{X}\pm SD$	44.5±14.1	44.8±15.6
Sex (% female)	38.9%	39.7%
Injury: duration (wk), $\bar{X}\pm SD$ , range	5.7±10.0, 1–65	7.3±9.1, 1–256
Subregion <sup>a</sup>		
Hip/groin	11.1%	11.8%
Knee	27.7%	30.7%
Ankle	16.7%	22.0%
Foot	5.6%	12.6%
Lower leg (calf, shin)	22.2%	14.2%
Upper leg (thigh)	5.6%	5.5%
Other (whole leg, ulcers)	5.6%	9.4%
Multiple areas	5.6%	5.5%

<sup>a</sup> Subregion percentages include individuals with multiple (2 or more) affected subregions, making the total greater than 100%. A statistical difference was present in the percentage representation of the categories of: ankle, foot, lower leg, and other.

range of participants and conditions (Tabs. 1 and 2). Exclusion criteria, defined and determined by the referring medical practitioner and the participating therapist were: pregnancy; less than 18 years of age; English comprehension difficulty; and “red flag” signs indicating nonmusculoskeletal symptoms and lower-limb conditions, including peripheral arterial occlusive disease, deep vein thrombosis, septic arthritis, and cellulitis.<sup>59</sup> In total, 142 participants were referred, with 15 being excluded (8 declined to participate, 1 was pregnant, 2 had English difficulty, 2 were excluded due to age, and 2 were referred for spinal conditions). Of the 127 participants, 111 received repeated measures, with 16 assessed only at baseline. Duration of symptom status was classified as: acute at 0 to 6 weeks, subacute at >6 to 12 weeks, and chronic at >12 weeks.<sup>60</sup> Each participant's injury was classified by region and subregion (Tabs. 1 and 2) to determine proportional representation.<sup>19,35</sup>

**Procedure**

At initial evaluation, the participants completed the LEFS and LLFI and two 11-point external criterion clinical change scales: a global numeric rating scale (NRS) of perceived present overall status<sup>61,62</sup> and a patient-specific index (PSI)<sup>26</sup> that generated a list of 5 items the individual has difficulty doing. Those participants who received treatment were re-measured every 2 weeks for 6 weeks and then every 4 weeks until discharge. The LEFS is a single-page, 20-item patient-reported outcome survey questionnaire with a 5-point (0–4) Likert scale in a matrix format.<sup>22</sup> The raw score is computed by totaling the points ranging from 0 to 80 (80=no disability) and multiplying the total points by 1.25 to provide a score of 0% to 100%. Up to 2 missing responses are permitted. The LLFI is a single-page, 25-item patient-reported outcome survey questionnaire with a 3-point Likert scale of 1 point for “Yes,” 0.5 point for “Partly,” and 0 points for “No.”

## Lower Limb Functional Index Development and Validation

**Table 2.**

Validation Study Participants (n=127) by Diagnoses and Percentage Representation of Total Sample as an Indication of Generalizability<sup>a</sup>

Subregion	Percentage <sup>b</sup>	Diagnoses (Alphabetic Listing)
Hip/groin	11.8	
	0.8	A-Capsule strain
	0.8	B-Impingement, hip joint
	3.1	C-Nonspecific pain
	2.4	D-Osteoarthritis
	0.8	E-ORIF
	0.8	F-Perthes disease
	2.4	G-Soft tissue strain, groin
	1.6	H-THR/PHR
Knee	30.7	
	3.9	A-ACL, both conservative and surgical
	5.5	B-Ligaments of knee: collateral, etc
	0.8	C-Meniscal cyst
	2.4	D-Meniscal, postoperative
	1.6	E-Meniscal trauma
	5.5	F-Nonspecific knee pain
	5.5	G-Osteoarthritis
	0.8	H-Patellar fracture
	6.3	I-Patellofemoral joint pain/dysfunction
	0.8	J-Patellofemoral joint, postoperative
	0.8	K-Patellofemoral joint subluxation
	0.8	L-Pes anserinus inflammation
	0.8	M-Soft tissue injury of knee
	0.8	N-Tibiofemoral joint
0.8	O-TKR/PKR	
Ankle	22.0	
	2.4	A-Fracture of ankle, tibia/fibula
	15.0	B-Ligament, collateral sprain, grade 1-3
	0.8	C-ORIF
	2.4	D-Soft tissue injury (not affecting collateral ligament)
	0.8	E-Syndesmosis injury
	0.8	F-Talus trauma and ligament
	0.8	G-Talar dome fracture
Foot	12.6	
	0.8	A-Calcaneum trauma
	0.8	B-Cuboid
	0.8	C-Hallux
	0.8	D-Heel pad pain
	1.6	E-Jones fracture
	1.6	F-Lisfranc ligament
	0.8	G-Metatarsal fracture
	0.8	H-Metatarsophalangeal joint
	2.4	I-Nonspecific foot pain
	0.8	J-Peripheral neuropathy
	1.6	K-Plantar fascia
	0.8	L-Sesamoid bone
Lower leg (calf, shin)	14.2	
	3.9	A-Achilles tendon
	5.5	B-Calf strain
	0.8	C-Compartment syndrome
	0.8	D-Medial tibial stress syndrome
	0.8	E-Peroneal strain
	1.6	F-Tibiofibular fracture, mid shaft
	0.8	G-Tibialis anterior muscle strain
Upper leg (thigh)	5.5	
	3.1	A-Hamstring muscle strain/tear grade 1-3
	0.8	B-Iliotibial band
	1.6	C-Rectus femoris and quadriceps muscles
Other (whole leg, ulcers)	9.4	
	7.0	A-Nonspecific leg pain
	1.6	B-Neuroma
	0.8	C-Ulcers
Multiple areas	5.5	Diagnoses were included above

<sup>a</sup> ORIF=open reduction internal fixation, THR=total hip replacement, PHR=partial hip replacement, ACL=anterior cruciate ligament, TKR=total knee replacement, PKR=partial knee replacement.

<sup>b</sup> Subregion and diagnoses percentage values include individuals with multiple (2 or more) affected subregions. Consequently, totals are greater than 100%.

Points are totaled, multiplied by 4, and subtracted from 100 to provide a score of 0% to 100% (0%=maximum function). Up to 2 missing responses are permitted. The global NRS for perceived status and the PSI used anchors at 0 and 10 (0="worst possible," 10="normal/no problem").<sup>61,62</sup> The PSI individual scores are summated, doubled, and subtracted from 100 to give a maximum functional impairment score of 0% to 100%.<sup>26</sup>

### Data Analysis

The phase 2 validation (Figure) utilized participants' baseline responses (n=127) to assess distribution, internal consistency, and factor structure. The total responses (n<sup>R</sup>=332) assessed floor and ceiling effects, missing responses, and criterion validity between the LLFI and LEFS. Subgroups of the baseline responses were formed for reliability (n=56, n<sup>R</sup>=112) and responsiveness (n=111, n<sup>R</sup>=222). For responsiveness, the time period for repeated measures was based on known group differences expected to occur with natural healing and the effects of treatment. This time period, which included the immediate postoperative and postfracture period, was classified as 2 weeks for patients with acute conditions, as 4 weeks for patients with subacute conditions, and as 6 weeks for patients with chronic conditions.<sup>19,25,63</sup> Error values of the standard error of the measurement (SEM) and the minimal detectable change (MDC) were calculated using the responsiveness subgroup (Figure). The minimal clinically important difference (MCID) was calculated using a subgroup of responses based on the clinically important change as determined by a 2-point change on the global NRS.<sup>64</sup>

### General Characteristics

*Distribution* and *normality* were assessed by visual inspection of the baseline scores histogram and

the one-sample Kolmogorov-Smirnov test<sup>54</sup> cutoff at a significance level of  $P > .05$ . The presence of items that represented both function and quality of life were verified from face and content validity.

### Psychometric Characteristics

*Internal consistency* was assessed by Cronbach alpha ( $\alpha = 0-1.00$ ) from baseline scores, with a cutoff of  $\geq .95$  indicating redundancy.<sup>65</sup>

*Test-retest reliability* (subsample  $n = 56$ ) used the ICC (2,1) with 95% CI<sup>66</sup> comparing baseline scores with scores obtained 3 days later, prior to the next treatment.<sup>26,52</sup> Only participants who were unchanged, with the NRS criterion reference standard at a change level of 0 or  $\pm 1$ <sup>19</sup> between tests, were used for assessment.

*Responsiveness* was assessed from effect size (ES)<sup>67</sup> and standardized response mean (SRM).<sup>68</sup> Participants with repeated measures who fulfilled the *a priori* periods of expected change based on natural healing or intervention were selected as having a known group difference<sup>69</sup> and classified by limb subregion.

*Validity* was assessed in all 4 forms. Face and content validity were assessed through the patient focus group, panel feedback, and readability scores.<sup>27</sup> Criterion-related validity was assessed through Pearson  $r$  coefficient for concurrently completed LLFI and LEFS responses ( $n^R = 332$ ). Construct validity compared groups that changed with groups that did not change. Change was determined from the 2 external criteria over 2 time points with the NRS score at  $\geq 20\%$  change and the PSI score at  $\geq 12.5\%$  change.<sup>19</sup> Both were required to categorize participants as improved or deteriorated,<sup>70</sup> with an *a priori* requirement of statistical difference between the baseline and repeated groups' paired  $t$  tests.

*Error scores* were determined with MDC at the 90% confidence interval (MDC<sub>90</sub>) from the SEM formula using the ICC.<sup>70</sup> The MCID was calculated using the distribution-based method with the construct validity definition of change on the 11-point NRS criterion measure.<sup>62,64</sup>

*Factor analysis* was assessed from baseline LLFI and LEFS data using maximum likelihood extraction<sup>54,71</sup> to clarify one factor for a single summated score,<sup>71-73</sup> with loading coefficient absolute value suppression at  $.40$ .<sup>54,74</sup> Factor extraction had 3 *a priori* requirements: scree plot point of inflection at the second eigenvalue,<sup>75</sup> eigenvalue cutoff  $> 1.0$ ,<sup>76</sup> and  $\geq 10\%$  variance.<sup>54,74</sup> SPSS version 14.0 (SPSS Inc, Chicago, Illinois) was used, with the level of significance at  $P < .05$ .

### Practical Characteristics

*Practicality* considered 9 distinct aspects.<sup>26,27,77,78</sup> The initial 5 aspects were: (1) self-administered, (2) applicable across a variety of conditions, (3) applicable across different severity levels, (4) relevance to defined populations, and (5) single-page length. The remaining 4 areas were determined individually through the patient focus group: (1) ease of understanding and ease of completion, assessed by an 11-point global NRS anchored at 0 for "Extremely difficult" and 10 for "Extremely easy"; (2) questionnaire completion time, the average of 3 completions timed manually following 2 minutes of familiarization; (3) scoring time, the average of 3 timed scorings, following 2 practice trials, by one therapist per clinic; and (4) missing responses, a percentage of total responses. *Readability* was determined from word-processing software<sup>79,80</sup> with Flesch-Kincaid grade scales (range = 0-12, optimum score is  $< 7$ ) and Flesch reading ease (optimum score is  $> 60\%$ ).

### Assessment of Clinimetrics

*Clinimetric performance* was assessed from 2 established clinimetric scales: (1) the Measurement of Outcome Measures (MOM) scale, which evaluates 25 aspects in 4 categories (methodological, practical, distributional, and general) by means of 3 response options ("Yes," "Partial," and "No"),<sup>19,26</sup> and (2) the Bot scale, which evaluates 12 aspects through 4 response options ("Good," "Doubtful," "Poor," and "Not available").<sup>27,81</sup> The Bot scale cutoff criteria were adjusted in 2 categories: (1) "Time to administer" was reduced from 10 to 3 minutes, and (2) "Readability and comprehension" was quantified by the Flesch-Kincaid readability criteria.<sup>19,27</sup>

### Role of the Funding Source

Research support was provided by the University of the Sunshine Coast.

## Results

### Phase 1—Development of the LLFI

The LLFI final version determined from field testing in stage 3 is presented in the Appendix. The method for item reduction used in stage 2 is presented in the Figure.

### Phase 2—Validation of the LLFI in a Cohort Population

Patient demographics are reported in Tables 1 and 2 for both the pilot and validation phases. No inferential statistics were used.

### General Characteristics

Distribution and normality were demonstrated through the Kolmogorov-Smirnov test for the LLFI ( $D[127] = 0.074$ ,  $P = .087$ ) and the LEFS ( $D[114] = 0.049$ ,  $P = .200$ ). Both measures had identical baseline score ranges (0%-98%), suitable histogram shape, and no floor or ceiling effects on visual examination. The LLFI "Partly" option was used by 43.3% of participants at baseline and in 40.1% of all responses. Face

## Lower Limb Functional Index Development and Validation

**Table 3.**

Clinimetric Properties of the Lower Limb Functional Index (LLFI) and the Lower Extremity Functional Scale (LEFS)<sup>a</sup>

Measure	Test-Retest Reliability, ICC (2,1)	Internal Consistency, Cronbach Alpha	Error Score		Responsiveness			Missing Responses
			SEM	MDC <sub>90</sub>	SD <sub>100</sub>	ES	SRM	Percentage
LLFI	.97	.91	2.84	6.63	21.86	1.27	1.75	0
LEFS	.97	.95	3.48	8.13	23.22	1.23	1.64	35

<sup>a</sup> ICC=intraclass correlation coefficient, SEM=standard error of the measurement, MDC<sub>90</sub>=minimal detectable change (90% confidence interval), SD<sub>100</sub>=standard deviation at baseline (100% scale), ES=effect size, SRM=standardized response mean.

and content validity were demonstrated only through the development (pilot) phase, with both function and quality of life being represented.

### Psychometric Characteristics

The methodological characteristics of internal consistency, reliability, responsiveness, and error score are summarized in Table 3. Criterion validity was high ( $r=.88$ ). Construct validity was demonstrated with a statistically significant ( $P<.001$ ) difference between baseline and repeated measures. Mean and score differences were comparable for both patient-reported outcome measures (Tab. 4).

Factor analysis indicated a single-factor structure for the LLFI and the LEFS, with all three *a priori* criteria met: suitable scree plot, eigenvalue  $>1.0$ , and variances  $>10\%$  (Tab. 5). The correlation matrix Kaiser-Meyer-Olkin values were .87 for the LLFI and .93 for the LEFS, with a significant Bartlett test of sphericity ( $P<.001$ ).

### Practical Characteristics

Ease of understanding and ease of completion were not significantly different, with an average score of 8.6 for the LLFI and 8.2 for the LEFS. Completion time for the LLFI required a mean ( $\pm$ SD) of  $131\pm 23$  seconds, 29% less than the LEFS at  $184\pm 31$  seconds. Scoring time required  $17\pm 5$  seconds for the LLFI and  $50\pm 19$  seconds for the LEFS, which increased to  $150\pm 39$  seconds with missing responses. The LLFI combined completion and scoring was  $148\pm 28$  seconds. The LEFS combined completion and scoring was  $234\pm 50$  seconds with no missing responses, but  $334\pm 70$  seconds with missing responses. Missing responses for the LLFI affected 1 of 332 responses, and no questionnaires were invalid. For the LEFS, 35% of the questionnaires had missing responses and 10% were invalid. Readability for the LLFI had a grade level of 7.2 and a reading ease of 61%, and the LEFS had a comparable grade level of 7.8 but a less preferable reading ease of 51%. Clinimetric performance on the MOM scale was

96% for the LLFI and 82% for the LEFS. On the Bot scale, the LLFI score was 100% (12/12) and the LEFS score was 83% (10/12).

## Discussion

### Main Findings

The LLFI was demonstrated as a psychometrically sound and practical patient-reported outcome survey tool suitable for assessing lower-limb function. This prospective study used external criteria that were not retrospective or reliant on recall<sup>82,83</sup> and that allowed direct comparison between the LLFI and LEFS. The LLFI demonstrated superior or equivalent clinimetric properties, in particular the psychometric characteristics of responsiveness and error values and the improved practical characteristics of missing responses, completion and scoring times, and mildly preferred readability. These findings were supported by overall clinimetric assessment via the MOM and Bot scale scores. The sample covered a large range of conditions and symptoms, which implied the results are generalizable and representative of the broader population with lower-limb musculoskeletal conditions treated in outpatient physical therapy clinics. Furthermore, these results added to the clinimetric knowledge of the LEFS.

### General Characteristics

The 3-stage development process enabled the LLFI face and content validity to be established and supported the previous findings for the LEFS.<sup>22,84</sup> This process ensured the

**Table 4.**

Construct Validity: Baseline and Repeated Scores for the Lower Limb Functional Index (LLFI) and the Lower Extremity Functional Scale (LEFS)

Measure <sup>a</sup>	Sample Size	Baseline, $\bar{X}\pm$ SD	Repeat Test, <sup>b</sup> $\bar{X}\pm$ SD
LLFI	111	52.5% $\pm$ 21.9%	78.6% $\pm$ 21.5% <sup>c</sup>
LEFS	111	51.8% $\pm$ 23.2%	79.3% $\pm$ 20.2% <sup>c</sup>

<sup>a</sup> 100% is maximum function.

<sup>b</sup> Repeated measures were made after a period of known natural healing: after 2 weeks in participants with acute conditions, after 4 weeks in participants with subacute conditions, and after 12 weeks in participants with chronic conditions.

<sup>c</sup> Significant difference with  $P<.001$  for all *t*-statistic measures.

**Table 5.**

Factor Analysis: Variance Explained for the Lower Limb Functional Index (LLFI) and the Lower Extremity Functional Scale (LEFS)<sup>a</sup>

Factor	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
LLFI			
1	7.57	30.29	30.29
2	1.92	7.68	37.96
3	1.30	5.22	43.18
4	1.18	4.73	47.91
5	1.11	4.44	52.35
6	1.08	4.31	56.66
7	1.01	4.03	60.68
8–25	0.95–0.21	39.32	100.00
LEFS			
1	10.68	53.42	53.42
2	2.04	10.19	63.61
3	1.21	6.05	69.67
4–20	0.83–0.06	30.33	100.00

<sup>a</sup> A forced one-factor solution was used with maximum likelihood extraction, varimax rotation, and a suppression of 0.40.

**Table 6.**

Baseline Mean ( $\pm$ SD) Scores by Subregion for the Lower Limb Functional Index (LLFI) and the Lower Extremity Functional Scale (LEFS)

Subregion	LLFI	LEFS
Hip/groin	47.8 $\pm$ 24.6	48.6 $\pm$ 29.4
Upper leg (thigh)	35.4 $\pm$ 22.5	27.5 $\pm$ 19.4
Knee	49.4 $\pm$ 22.1	51.2 $\pm$ 24.9
Lower leg (calf, Achilles tendon, shin)	47.2 $\pm$ 26.3	47.5 $\pm$ 20.3
Ankle	48.9 $\pm$ 23.0	48.8 $\pm$ 26.9
Foot	36.3 $\pm$ 27.2	37.3 $\pm$ 22.4
Other (whole leg, ulcers, nerve palsy)	57.2 $\pm$ 24.2	47.7 $\pm$ 19.3
Multiple areas	33.0 $\pm$ 14.8	24.5 $\pm$ 5.4
Total average	48.1 $\pm$ 24.0 n=127	52.3 $\pm$ 24.7 n=127

LLFI was acceptable and satisfactory for both clinicians and patients. The pilot study facilitated the main validation study, as it enabled estimates for sample size and assessment of the change characteristic, which reduced the potential for errors.<sup>52</sup> The baseline distribution of LEFS and LLFI scores was similar (Tab. 6). The LLFI “Partly”<sup>19</sup> response was accepted by patients and used 43% of the time. Although the selection

of this response afforded fewer response options compared with the 5-point LEFS format, it provided stable, equally spaced responses with 2 advantages. First, 3 options reduce the psychological dilemma of selecting a response without increased cognitive demand, as the required options of “for,” “against,” and “intermediate” are met.<sup>46,47</sup> Second, it allows results to be analyzed with more powerful parametric statistics,

as the requirement is met for the lowest level of interval data needed with individual scores that are summed to provide a total score.<sup>85,86</sup>

### Psychometric Characteristics

The LLFI psychometric characteristics were established and preferred to those of the LEFS. The levels of internal consistency, responsiveness, error score, and factor structure all favored the LLFI. The LEFS and LLFI were equal with respect to reliability without a preference for one scale over another, with a level similar to previously published LEFS values (ICC = .88–.94).<sup>22,30,87,88</sup> Also, the findings for construct validity were similar for both patient-reported outcome measures where both baseline and repeated measures were comparable.

Internal consistency values favored the LLFI ( $\alpha$  = .91) compared with the LEFS ( $\alpha$  = .95). This result for the LEFS’s internal consistency is similar to the findings of previous studies<sup>22,28,31</sup> and was sufficiently high to indicate potential item redundancy.<sup>29,89</sup> The responsiveness values were consistently greater for the LLFI than for the LEFS. This result may simply be a consequence of the slightly lower baseline standard deviation for the LLFI.<sup>67,68</sup> As an observational study, other influences may have included: lower baseline severity (as change rates vary between patients with acute and chronic conditions), the variation in follow-up duration for patients with acute and subacute conditions and those with chronic conditions (as an instrument is less responsive over shorter follow-up periods), and interventions provided were at the treating therapist’s discretion. For the LEFS, the responsiveness values (SEM and ES) also were marginally lower than previously reported.<sup>31,87</sup> This finding may be attributed to our use of all known group participants anticipated to improve through natural



healing or treatment intervention, as opposed to only participants categorized as “responders.” This may have potentially reduced error indexes, but does not distinguish between the MDC<sup>64</sup> and minimal important change.<sup>90</sup>

The LLFI error scores (SEM=2.84%, MDC<sub>90</sub>=6.63%) were preferred to those of the LEFS (SEM=3.48%, MDC<sub>90</sub>=8.13%) in the current study. Interestingly, the LEFS’s MDC value in this study was lower than the MDC values reported in 4 prior studies (11.3%–12.4%, 9.0–9.9 scale points).<sup>22,87,88,91</sup> These differences probably were due to the higher reliability in the current study and, compared with the 2 recent studies,<sup>87,88</sup> to the fact that our patients would be expected to show greater improvement through natural healing at a faster rate due to age and injury mechanism.

Factor analysis results indicated a single-factor structure; thus, the LLFI and LEFS scores can be summed for a single score. The LEFS factor structure was determined using maximum likelihood extraction<sup>19,71</sup> and is reported for the first time. However, there were some concerns. Specifically, the cross-loading of multiple LEFS items indicated a tendency for a multifactor structure, which supported the internal consistency results and suggests the LEFS may have item redundancy.<sup>29,89</sup> Both the LLFI and LEFS had additional factors that accounted for substantial variance. This finding suggests item reduction may be appropriate for both patient-reported outcome scales and warrants further research.

### Practical Characteristics

There were practical advantages of the LLFI over the LEFS, including reduced user burden as shown by fewer missing responses, improved readability, and shorter completion and scoring times.

The clinimetric assessment on both the MOM and Bot scales demonstrated the LLFI was preferred. The LLFI’s higher levels of responsiveness and sensitivity would enable greater efficiency for researchers and clinicians in determining whether selected intervention strategies were effective. Moreover, a smaller change score and shorter time period would be required to evaluate an intervention’s outcome.

### Study Limitations and Strengths

Limitations of the study were that only participants from physical therapy outpatient clinics were included and that specific conditions or diagnostic subgroups were not investigated. The results cannot be generalized to inpatient or community settings or to other body regions. The study’s strengths are the provision of a single region-specific, lower-limb, patient-reported outcome scale that improved sensitivity and reduced clinician burden and missing responses compared with a recognized criterion standard. The sample was from multiple centers and included patients with conditions affecting each subregion of the lower limb, with varied degrees of severity and duration, who represented both the general and work-injured populations. These attributes ensure the LLFI fulfills the recommendations stipulated by previous researchers<sup>6,8,30</sup> for any new patient-reported outcome measure to ensure it is applicable to outpatients with lower-limb disorders.

### Implications for Further Research

The high correlation indicating criterion validity between the LLFI and LEFS implied the LLFI could be generalized to populations in which the LEFS had been validated.<sup>31,92</sup> However, this generalization must be made with caution until further investigations of the LLFI in other general and diagnostically specific populations are made. Independent

validation is essential, and further concurrent investigation is needed. Furthermore, investigations are warranted to assess potential item reduction to shorten the LLFI, further reducing respondent and clinician burden.<sup>32</sup>

## Conclusions

The LLFI is a practical patient-reported outcome measure to assess functional status in patients with lower-limb conditions. Compared with the LEFS, the LLFI demonstrated preferred clinical utility and improved clinimetric performance due to superior psychometric and practical characteristics. These findings indicate the LLFI is a viable patient-reported outcome measure for the evaluation of lower-limb status and impairment in clinical and research settings.

Mr Gabel, Dr Burkett, and Dr Michener provided concept/idea/research design. All authors provided writing and data analysis. Mr Gabel and Dr Burkett provided data collection and project management. Mr Gabel provided participants, facilities/equipment, and clerical support. Dr Melloh provided institutional liaisons. Mr Gabel, Dr Melloh, and Dr Michener provided consultation (including review of manuscript before submission). The authors thank all participating patients, general practitioners, and therapists for their time and effort.

The study was approved by the Human Research Ethics Committee of the University of the Sunshine Coast.

This research, in part, was presented at the Scientific Meeting and Annual Conference of the American Physical Therapy Association; June 8–11, 2005; Boston, Massachusetts, and at the Fourteenth Biennial Australian Physiotherapy Association-Musculoskeletal Physiotherapy Australia Conference; November 24–26, 2005; Brisbane, Australia.

Research support was provided by the University of the Sunshine Coast.

DOI: 10.2522/ptj.20100199

## References

- 1 Jette AM, Tao W, Norweg A, Haley S. Interpreting rehabilitation outcome measurements. *J Rehabil Med.* 2007;39:585–590.

- 2 Sackett D, Sharon E, Straus W, et al. *Evidence-Based Medicine: How to Practice and Teach EBM*. 2nd ed. New York, NY: Churchill Livingstone; 2000.
- 3 Tranfield DR, Denyer D, Smart P. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal Management*. 2003;14:207-222.
- 4 Wang TJ. Concept analysis of functional status. *Int J Nurs Stud*. 2004;41:457-462.
- 5 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;25:521-525.
- 6 Eecheute C, Vaes P, Van Aerschot L, et al. The clinimetric qualities of patient-assessed instruments for measuring chronic ankle instability: a systematic review. *BMC Musculoskelet Disord*. 2007; 8:6.
- 7 Riddle DL, Stratford PW, Bowman DH. Findings of extensive variation in the types of outcome measures used in hip and knee replacement clinical trials: a systematic review. *Arthritis Rheum*. 2008;59:876-883.
- 8 Naal FD, Impellizzeri FM, Rippstein PF. Which are the most frequently used outcome instruments in studies on total ankle arthroplasty? *Clin Orthop Relat Res*. 2010; 468:815-826.
- 9 Eng JJ, Pierrynowski MR. Evaluation of soft foot orthotics in the treatment of patellofemoral pain syndrome [erratum in: *Phys Ther*. 1993;73:330]. *Phys Ther*. 1993; 73:62-68.
- 10 Flandry F, Hunt JP, Terry GC, Hughston JC. Analysis of subjective knee complaints using visual analog scales. *Am J Sports Med*. 1991;19:112-118.
- 11 Hawker G, Melfi C, Paul J, et al. Comparison of a generic (SF-36) and a disease-specific (WOMAC) (Western Ontario and McMaster Universities Osteoarthritis Index) instrument in the measurement of outcomes after knee replacement surgery. *J Rheumatol*. 1995;22:1193-1196.
- 12 Mohtadi N. Development and validation of the Quality Of Life Outcome Measure (Questionnaire) for chronic anterior cruciate ligament deficiency. *Am J Sports Med*. 1998;26:350-359.
- 13 Barber FA, Aziz-Jacobo J, Oro FB. Anterior cruciate ligament reconstruction using patellar tendon allograft: an age-dependent outcome evaluation. *Arthroscopy*. 2010;26:488-493.
- 14 Kujala UM, Jaakkola LH, Koskinen SK, et al. Scoring of patellofemoral disorders. *Arthroscopy*. 1993;9:159-163.
- 15 Robinson JM, Cook JL, Purdam C, et al. The VISA-A questionnaire: a valid and reliable index of the clinical severity of Achilles tendinopathy. *Br J Sports Med*. 2001; 35:335-341.
- 16 White DK, Keysor JJ, Lavalley MP, et al. Clinically important improvement in function is common in people with or at high risk of knee OA: the MOST study. *J Rheumatol*. 2010;37:1244-1251.
- 17 Martin RL, Irrgang JJ. A survey of self-reported outcome instruments for the foot and ankle. *J Orthop Sports Phys Ther*. 2007;37:72-84.
- 18 Australian Physiotherapy Association. Policy on outcome measures and treatment justification. Available at: <http://www.tac.vic.gov.au/upload/TAC+Physio+Chart.pdf> and <http://www.physiotherapy.asn.au/quality-practice/outcome-measures/>.
- 19 Gabel CP, Michener LA, Melloh M, Burkett B. Modification of the Upper Limb Functional Index to a three-point response improves clinimetric properties. *J Hand Ther*. 2010;23:41-52.
- 20 Oberg U, Oberg B, Oberg T. Validity and reliability of a new assessment of lower-extremity dysfunction. *Phys Ther*. 1994; 74:861-871.
- 21 Finch E, Walsh M, Thomas SG, Woodhouse LJ. Functional ability perceived by individuals following total knee arthroplasty compared to age-matched individuals without knee disability. *J Orthop Sports Phys Ther*. 1998;27:255-263.
- 22 Binkley JM, Stratford PW, Lott SA, Riddle DL. The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. *Phys Ther*. 1999;79:371-383.
- 23 Johanson NA, Liang MH, Daltroy L, et al. American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments: reliability, validity, and sensitivity to change. *J Bone Joint Surg Am*. 2004;86:902-909.
- 24 Martin RL, Irrgang JJ, Burdett RG, et al. Evidence of validity for the Foot and Ankle Ability Measure (FAAM). *Foot Ankle Int*. 2005;26:968-983.
- 25 Beaton DE, Katz JN, Fossel AH, et al. Measuring the whole or the parts: validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. *J Hand Ther*. 2001; 14:128-146.
- 26 Gabel CP, Michener LA, Burkett B, Neller A. The Upper Limb Functional Index: development and determination of reliability, validity and responsiveness. *J Hand Ther*. 2006;19:328-349.
- 27 Bot SD, Terwee CB, van der Windt DA, et al. Clinimetric evaluation of shoulder disability questionnaires: a systematic review of the literature. *Ann Rheum Dis*. 2004;63:335-341.
- 28 Cacchio A, De Blasis E, Necozone S, et al. The Italian version of the Lower Extremity Functional Scale was reliable, valid, and responsive. *J Clin Epidemiol*. 2010;63: 550-557.
- 29 Boyle GJ. Does item homogeneity indicate internal consistency or item redundancy in psychometric scales? *Pers Individ Dif*. 1991;12:291-294.
- 30 Watson CJ, Propps M, Ratner J, et al. Reliability and responsiveness of the Lower Extremity Functional Scale and the Anterior Knee Pain Scale in patients with anterior knee pain. *J Orthop Sports Phys Ther*. 2005;35:136-146.
- 31 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;89:580-588.
- 32 Gabel CP, Burkett B, Yelland M. Balancing fidelity and practicality in short version musculoskeletal outcome measures. *Phys Ther Rev*. 2009;14:221-225.
- 33 Carcia CR, Martin RL, Drouin JM. Validity of the foot and ankle ability measure in athletes with chronic ankle instability. *J Athl Train*. 2008;43:179-183.
- 34 Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chronic Dis*. 1985;38:27-36.
- 35 Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to their Development and Use*. 4th ed. New York, NY: Oxford University Press; 2008.
- 36 Kopec JA, Sayre EC, Davis AM, et al. Assessment of health-related quality of life in arthritis: conceptualization and development of five item banks using item response theory. *Health Qual Life Outcomes*. 2006;4:33.
- 37 Patient-Reported Outcome Measurement Information Systems (PROMIS). Version 1.0 Item Banks; 2010. Available at: <http://www.nihpromis.org/science/ItemClassification>.
- 38 Reeve BB, Hays RD, Bjorner JB, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care*. 2007;45(5 suppl 1):S22-S31.
- 39 Kopec JA. Measuring functional outcomes in persons with back pain: a review of back-specific questionnaires. *Spine (Phila Pa 1976)*. 2000;25:3110-3114.
- 40 DeWalt DA, Rothrock N, Yount S, et al. Evaluation of item candidates: the PROMIS qualitative item review. *Med Care*. 2007; 45(5 suppl 1):S12-S21.
- 41 World Health Organization. *International Classification of Functioning, Disability and Health (ICF)*; 2001. Available at: [www.who.int/icidh](http://www.who.int/icidh). Accessed July 31, 2009.
- 42 World Health Organization. The ICF Browser. Available at: <http://apps.who.int/classifications/icfbrowser/>. Accessed January 12, 2010.
- 43 Escorpizo R, Stucki G, Cieza A, et al. Creating an interface between the *International Classification of Functioning, Disability and Health* and physical therapist practice. *Phys Ther*. 2010;90:1053-1063.
- 44 International Organization for Standardization. ISO/TR 16982:2002; Ergonomics of human-system interaction: usability methods supporting human-centred design. Available at: [http://www.iso.org/iso/catalogue\\_detail?csnumber=31176:44](http://www.iso.org/iso/catalogue_detail?csnumber=31176:44). Accessed November 29, 2010.
- 45 Holm I. *Ideas and Beliefs in Architecture and Industrial design: How Attitudes, Orientations, and Underlying Assumptions Shape the Built Environment* [doctoral thesis]. Oslo, Norway: Oslo School of Architecture and Design; 2006.
- 46 Krosnick JA. *The Handbook of Questionnaire Design*. New York, NY: Oxford University Press; 1991.

## Lower Limb Functional Index Development and Validation

- 47 Albarracín D, Johnson BT, Zanna MP. *The Handbook of Attitudes*. Hillsdale, NJ: Lawrence Erlbaum Associates; 2005.
- 48 George K, Batterham A, Sullivan I. Validity in clinical research: a review of basic concepts and definitions. *Phys Ther Sport*. 2000;1:19-27.
- 49 Guyatt GH, Osoba D, Wu AW, et al. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc*. 2002;77:371-383.
- 50 Terwee CB, Dekker FW, Wiersinga WM, et al. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. *Qual Life Res*. 2003;12:349-362.
- 51 Eurich DT, Johnson JA, Reid KJ, Spertus JA. Assessing responsiveness of generic and specific health related quality of life measures in heart failure. *Health Qual Life Outcomes*. 2006;4:89.
- 52 Stratford PW, Riddle DL. Assessing sensitivity to change: choosing the appropriate change coefficient. *Health Qual Life Outcomes*. 2005;3:23.
- 53 Lowe A, Wessel J, Battié MC. Test-retest reliability, construct validity, and responsiveness of a functional pain scale for tennis elbow. *Physiother Can*. 2003;55:114-122.
- 54 Field A. *Discovering Statistics Using SPSS*. 2nd ed. London, United Kingdom: Sage Publications Ltd; 2005.
- 55 Dawson B, Trapp R. *Basic and Clinical Biostatistics*. 2nd ed. Sydney, Australia: McGraw Hill Publishers; 2001.
- 56 Meng X, Rosenthal R, Rubin DB. Comparing correlated correlation coefficients. *Psychol Bull*. 1992;111:172-175.
- 57 Anglim J. Using R to test for significant differences between two nonindependent correlations. Available at: <http://jeromyanglim.blogspot.com/2008/11/using-r-to-test-for-significance-test.html>. Accessed December 4, 2010.
- 58 Guadagnoli E, Velicer WF. Relation of sample size to the stability of component patterns. *Psychol Bull*. 1988;103:265-275.
- 59 Samanta J, Kendall J, Samanta A. 10-minute consultation: chronic low back pain. *BMJ*. 2003;326:535.
- 60 van Tulder M, Becker A, Bekkering T, et al. Chapter 3: European guidelines for the management of acute nonspecific low back pain in primary care. *Eur Spine J*. 2006;15(suppl 2):S169-S191.
- 61 Bowling A. Just one question: if one question works, why ask several? *J Epidemiol Community Health*. 2005;59:342-345.
- 62 Ostelo RW, Deyo RA, Stratford PW, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila PA 1976)*. 2008;33:90-94.
- 63 Hogg-Johnson S, Cole DC. Early prognostic factors for duration on temporary total benefits in the first year among workers with compensated occupational soft tissue injuries. *Occup Environ Med*. 2003;60:244-253.
- 64 Copay AG, Subach BR, Glassman SD, et al. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J*. 2007;7:541-546.
- 65 Nunnally JC, Bernstein IH. *Psychometric Theory*. New York, NY: McGraw-Hill; 1994.
- 66 Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86:420-428.
- 67 Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
- 68 Liang MH, Fossel AH, Larson MG. Comparison of five health status instruments for orthopaedic evaluation. *Med Care*. 1990;28:632-642.
- 69 Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol*. 2000;53:459-468.
- 70 Jacobson NS, Follette WC, Revensdorf D. Psychotherapy outcomes research: methods for reporting variability and evaluating clinical significance. *Behav Ther*. 1984;15:336-352.
- 71 Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. *Psychol Methods*. 1999;4:272-299.
- 72 Costello AB, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment Research Evaluation*. 2005;10:1-9.
- 73 Doward LC, McKenna SP. Defining patient-reported outcomes. *Value Health*. 2004;7(suppl 1):S4-S8.
- 74 Stevens JP. *Applied Multivariate Statistics for the Social Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1992.
- 75 Cattell RB. The scree test for the number of factors. *Multivariate Behav Res*. 1966;1:245-276.
- 76 Kaiser HF. The application of electronic computers to factor analysis. *Educ Psychol Meas*. 1960;20:141-151.
- 77 Liang MH, Jette AM. Measuring functional ability in chronic arthritis: a critical review. *Arthritis Rheum*. 1981;24:80-86.
- 78 Michener LA, Leggin BG. A review of self-report scales for the assessment of functional limitation and disability of the shoulder. *J Hand Ther*. 2001;14:68-76.
- 79 Doak CC, Doak LG, Root JH. *Teaching Patients With Low Literacy Skills*. 2nd ed. Philadelphia, PA: JB Lippincott Co; 1996.
- 80 Paasche-Orlow MK, Taylor HA, Brancati FL. Readability standards for informed-consent forms as compared with actual readability. *N Engl J Med*. 2003;348:721-726.
- 81 Lohr KN, Aaronson NK, Alonso J, et al. Evaluating quality of life and health status instruments: development of scientific review criteria. *Clin Ther*. 1996;18:979-992.
- 82 Norman GR, Stratford PW, Regehr G. Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. *J Clin Epidemiol*. 1997;50:869-879.
- 83 Schmitt J, Di Fabio RP. The validity of prospective and retrospective global change criterion measures. *Arch Phys Med Rehabil*. 2005;86:2270-2276.
- 84 Garratt AM, Brealey S, Gillespie WJ; DAM-ASK Trial Team. Patient-assessed health instruments for the knee: a structured review. *Rheumatology (Oxford)*. 2004;43:1414-1423.
- 85 Newcombe RG. Confidence intervals for the mean of a variable taking the values 0, 1 and 2. *Stat Med*. 2003;22:2737-2750.
- 86 Jacoby J, Matell MS. Three point Likert scales are good enough. *J Mark Res*. 1971;8:495-500.
- 87 Yeung TS, Wessel J, Stratford PW, Macdermid J. Reliability, validity, and responsiveness of the lower extremity functional scale for inpatients of an orthopaedic rehabilitation ward. *J Orthop Sports Phys Ther*. 2009;39:468-477.
- 88 Pua YH, Cowan SM, Wrigley TV, Bennell KL. The Lower Extremity Functional Scale could be an alternative to the Western Ontario and McMaster Universities Osteoarthritis Index physical function scale. *J Clin Epidemiol*. 2009;62:1103-1111.
- 89 McDonald RP. The dimensionality of tests and items. *Br J Math Stat Psychol*. 1981;34:110-117.
- 90 de Vet HC, Terwee CB, Ostelo RW, et al. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health Qual Life Outcomes*. 2006;4:54.
- 91 Alcock GK, Stratford PW. Validation of the Lower Extremity Functional Scale on athletic subjects with ankle sprains. *Physiother Can*. 2002;54:233-240.
- 92 Stratford PW, Kennedy DM, Hanna SE. Condition-specific Western Ontario McMaster Osteoarthritis Index was not superior to region-specific Lower Extremity Functional Scale at detecting change. *J Clin Epidemiol*. 2004;57:1025-1032.

**Appendix.**

The Lower Limb Functional Index (LLFI)<sup>a</sup>

---

**LOWER LIMB FUNCTIONAL INDEX**

**DATE:** \_\_\_\_\_

**NAME:** \_\_\_\_\_ **INJURY** \_\_\_\_\_  **LEFT LEG**  **RIGHT LEG**

**PLEASE COMPLETE:** Your leg/s may make it difficult to do some things you normally do. This list contains sentences people use to describe themselves with such problems. Think of yourself over the last few days.

**If an item describes you, mark the line. If not, leave it blank.**

**If an item partly describes you - Use a Half (1/2) Mark.**

**DUE TO MY LEG/S:**

- \_\_\_ 1. I stay at home most of the time.
- \_\_\_ 2. I change position frequently for comfort.
- \_\_\_ 3. I avoid heavy jobs (eg, cleaning, lifting more than 5 kg or 10 lb, gardening).
- \_\_\_ 4. I rest more often.
- \_\_\_ 5. I get others to do things for me.
- \_\_\_ 6. I have the pain/problem almost all the time.
- \_\_\_ 7. I have difficulty lifting and carrying (eg, shopping bags up to 5 kg or 10 lb).
- \_\_\_ 8. My appetite is now different.
- \_\_\_ 9. My walking or normal recreation or sporting activity is affected.
- \_\_\_ 10. I have difficulty with normal home or family duties and chores.
- \_\_\_ 11. I sleep less well.
- \_\_\_ 12. I need assistance with personal care (eg, washing, hygiene).
- \_\_\_ 13. My regular daily activities (work, social contacts) are affected.
- \_\_\_ 14. I am more irritable and/or bad tempered.
- \_\_\_ 15. I feel weaker and/or stiffer.
- \_\_\_ 16. My transport independence is affected (driving, public transport).
- \_\_\_ 17. I have difficulty or need help with dressing (eg, trousers/pants/shoes and socks).
- \_\_\_ 18. I have difficulty changing directions, twisting or turning.

(Continued)

---

## Lower Limb Functional Index Development and Validation

---

### Appendix.

Continued

---

- \_\_\_ 19. I am unable to move as fast as I would wish.
- \_\_\_ 20. I have difficulty with prolonged or extended standing.
- \_\_\_ 21. I have difficulty bending, squatting, and/or reaching down.
- \_\_\_ 22. I have difficulty with long or extended walks.
- \_\_\_ 23. I have difficulty with steps and stairs.
- \_\_\_ 24. I have difficulty with sitting for prolonged or extended times.
- \_\_\_ 25. I have problems with my balance on uneven surfaces and/or with unaccustomed footwear.

**LLFI SCORE: To score the upper part, add the marked boxes:**

**TOTAL (LLFI points above) × 4 = \_\_\_/100 points.**      **FINAL TOTAL (100 – TOTAL) =  %**

**MDC (90% CI):** 6.67% or 1.67 LLFI points. Change less than this may be due to error.

---

<sup>a</sup> MDC (90%CI)=minimal detectable change (90% confidence interval). The LLFI may not be used or reproduced without written permission of the authors.