

Cross-Cultural Adaptation and Validation of the Lower Extremity Measure Into German

Geriatric Orthopaedic Surgery
& Rehabilitation
2015, Vol. 6(4) 282-288
© The Author(s) 2015
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/2151458515604715
gos.sagepub.com



Sabine Goldhahn, MD¹, Kurt Käch, MD², Hans Curd Frei, MD³,
Mark Rudin, MD², Melanie Leimbacher, MD⁴,
and Andreas Platz, MD⁴

Abstract

Introduction: The Lower Extremity Measure (LEM) was developed to provide a specific instrument to detect changes in physical function in patients with hip fracture. Of 29 questions, 3 have a valid “not applicable” answer option. The goal of this study was to validate the LEM in German and to determine the added value to the physical functioning (pf) subscale of the Short Form 36 (SF-36). **Materials and Methods:** The LEM was translated according to published guidelines and administered to patients with hip fracture (31 A1-A3 and 31 B1-B3) shortly after surgery (baseline), at 3 months (3M), and for reliability testing at 3 months plus 1 week (3M+). The reproducibility, internal consistency, floor and ceiling effects, construct validity, and responsiveness of the German LEM were assessed. **Results:** A total of 106 patients completed the LEM and SF-36 (mean age 75.5; 67% women) at baseline (mean of 4.9 days after operation), and 88 completed both questionnaires at both the 3M and 3M+ assessments. At each assessment time point, between 6% and 23% of the patients answered 7 questions as “not applicable.” Reproducibility and internal consistency were high (intraclass correlation coefficient = 0.93; Cronbach’s α = .96). No floor effect (0%) and a minor ceiling effect (7.87%) were found for the total LEM score. The strongest correlation was found between the LEM and the SF-36 subscale pf (Spearman ρ = .93). Responsiveness was similar for the SF-36 pf subscale and the LEM when using effect size (SF-36 pf 0.71 vs LEM 0.72) and better for the LEM when using standardized response mean (SF-36 pf 0.65 vs LEM 0.76). **Discussion:** The German LEM is a reliable, valid, and responsive measure for the self-assessment of patients after hip fracture surgery. As a number of questions are not applicable to elderly patients, the added value of this lengthy questionnaire in these often frail, sometimes cognitively impaired patients is still open for debate.

Keywords

fragility fractures, geriatric trauma, hip fracture, patient-reported outcome, validation

Introduction

Evaluation of hip fracture surgery has traditionally focused on clinical or surgeon-defined measures, such as the Harris Hip Score or the Charnley score.^{1,2} Although there are a number of patient-reported outcome measures (PROMs) and generic quality-of-life instruments validated for hip osteoarthritis and other hip-related disorders,³⁻⁶ there are no validated functional PROMs specific to hip fractures.⁷ The Lower Extremity Functional Scale, a self-reported questionnaire that has been validated in the context of traumatic injuries of the lower extremities in general, has not been specifically designed for hip fractures and includes questions regarding vigorous activities, such as running and hopping, which are typically not applicable to elderly patients.^{8,9} The Musculoskeletal Function Assessment (MFA) or its short version, the SMFA, has been used in patients with a wide spectrum of musculoskeletal problems including fractures but is also not hip fracture specific.¹⁰ The Lower Extremity Measure (LEM) is a PROM that was developed based on the Toronto Extremity Salvage Score.¹¹

Emphasis was put on designing a short, simple questionnaire considering the advanced age of most patients with hip fracture. The LEM, which is available in French and in English, was shown to be a reliable, valid, and responsive tool to evaluate function in patients with a hip fracture.¹² As no such validated tool is available in German, the goal of the present study was to validate the LEM in German to quantify its psychometric properties and thereby to determine the added value to

¹ AO Clinical Investigation, AO Foundation, Davos/Switzerland

² Department of Trauma Surgery, Cantonal Hospital Winterthur, Winterthur/Switzerland

³ Department of Surgery/Orthopaedics, Hospital Davos, Davos/Switzerland

⁴ Department of General, Hand, and Trauma Surgery, City Hospital Triemli, Zurich/Switzerland

Corresponding Author:

Sabine Goldhahn, AO Clinical Investigation, AO Foundation, Clavadelerstrasse 8, 7270 Davos/Switzerland.

Email: sabine.goldhahn@aofoundation.org

Table 1. Floor and Ceiling Effects, Internal Consistency, and Reliability of Individual Lower Extremity Measure items.

Item ^a	3 Months								Reliability (retest at 3 months + 1 week)	
	Min	Min %	Max	Max%	Mean	SD	ρ^b	α^c	ICC	95%CI
1. Getting out of bed is ...	2	0.00	5	69.41	4.59	0.71	.65	.962	0.72	0.60-0.81
2. Getting in/out bathtub is ...	1	20.90	5	40.30	3.54	1.56	.79	.963	0.89	0.83-0.93
3. Getting on/off toilet is ...	1	1.16	5	74.42	4.64	0.73	.50	.962	0.76	0.66-0.84
4. Showering is ...	1	3.45	5	66.67	4.41	1.02	.69	.961	0.88	0.82-0.92
5. Putting on a pair of pants is ...	1	3.41	5	50.00	4.26	0.96	.67	.962	0.74	0.62-0.82
6. Putting on socks/stockings is ...	1	5.81	5	31.40	3.84	1.15	.70	.962	0.78	0.68-0.85
7. Putting shoes is ...	1	3.49	5	41.86	4.07	1.04	.66	.962	0.78	0.69-0.85
8. Rising from a chair is ...	2	0.00	5	68.97	4.56	0.73	.59	.962	0.82	0.74-0.88
9. Standing upright is ...	1	2.33	5	67.44	4.47	0.93	.63	.962	0.76	0.65-0.83
10. Kneeling is ...	1	28.41	5	22.73	2.92	1.56	.77	.961	0.90	0.85-0.93
11. Getting up from kneeling is ...	1	31.76	5	15.29	2.74	1.49	.79	.962	0.88	0.82-0.92
12. Bending to pick something up off the floor is ...	1	5.68	5	43.18	3.88	1.26	.75	.961	0.84	0.77-0.89
13. Sitting is ...	3	0.00	5	86.05	4.83	0.47	.40	.963	0.66	0.52-0.76
14. Walking within the house is ...	1	2.30	5	67.82	4.45	0.97	.67	.961	0.88	0.82-0.92
15. Walking downstairs is ...	1	5.75	5	40.23	3.93	1.18	.77	.961	0.86	0.79-0.90
16. Walking upstairs is ...	1	5.75	5	41.38	3.98	1.16	.79	.960	0.88	0.82-0.92
17. Walking outside is ...	1	6.82	5	45.45	4.02	1.19	.73	.960	0.82	0.74-0.88
18. Walking up/down ramps is ...	1	6.98	5	29.07	3.73	1.20	.79	.960	0.89	0.83-0.92
19. Getting in/out car is ...	1	2.33	5	41.86	4.06	1.02	.71	.962	0.79	0.70-0.86
20. Using public transportation is ...	1	14.67	5	42.67	3.76	1.42	.86	.961	0.94	0.91-0.96
21. Preparing light meals is ...	1	3.75	5	70.00	4.44	1.03	.71	.961	0.84	0.76-0.89
22. Tidying, dusting, washing dishes is ...	1	6.10	5	65.85	4.32	1.16	.76	.961	0.87	0.81-0.91
23. Doing laundry, vacuuming (heavy housework) is ...	1	12.66	5	34.18	3.58	1.38	.89	.961	0.87	0.81-0.91
24. Gardening/yard work is ...	1	36.84	5	19.30	2.63	1.57	.89	.963	0.91	0.87-0.94
25. Food shopping is ...	1	13.10	5	46.43	3.89	1.40	.81	.961	0.85	0.79-0.90
26. Socializing with friends is ...	1	5.88	5	70.59	4.39	1.17	.71	.961	0.92	0.89-0.95
27. Doing the usual number of hours for your normal daily activities is ...	1	9.20	5	40.23	3.87	1.28	.80	.960	0.85	0.78-0.90
28. Completing your usual daily activities is ...	1	6.82	5	46.59	4.02	1.22	.84	.960	0.87	0.81-0.92
29. Participating in usual leisure activities is ...	1	18.52	5	38.27	3.48	1.54	.75	.962	0.79	0.70-0.86

Abbreviations: ICC, intraclass correlation coefficient; 95% CI, 95% confidence interval for ICC; SD, standard deviation.

^aEach item is scored based on a Likert-type 1-5 point scale: 1 = impossible to do; 2 = extremely difficult; 3 = moderately difficult; 4 = a little bit difficult; 5 = not at all difficult.

^bItem-total correlation: Spearman rank correlation between each item and the total score.

^cCronbach's α if item removed.

the physical functioning subscale of the Short Form 36 (SF-36) when assessed in patients with hip fractures.

Materials and Methods

Lower Extremity Measure and Cross-Cultural Adaptation

The English LEM consists of 29 items on activities of daily living rated from 1 (impossible) to 5 (not at all difficult), including a "not applicable" option. A total score from 0 to 100 is calculated, with higher scores indicating higher levels of function.¹² Three questions (Qs) have a valid "not applicable" option: Q4 (Showering), Q20 (Using public transportation), and Q24 (Gardening/yard work). If the task is not applicable, it does not contribute to the total score. We performed the cross-cultural adaptation of the LEM into German with forward and backward translations, pretesting, and agreement meetings according to the established guidelines.¹³⁻¹⁵ The

German LEM (Untere Extremitäten-Fragebogen LEM-D) is shown in Appendix A.

Validation study

The validation study was performed in 2 clinics after the approval of the applicable ethics committee. A total of 106 patients with a mean age of 75.5 ± 13.3 years (67% women) having a fresh hip fracture (AO classification type 31 A1-A3 and 31 B1-B3) gave their written informed consent. Only patients who were able to read and understand German language and to understand the patient information and sign the consent form themselves were included. Proxy answers by relatives were not permitted. The patients were surgically treated with an intramedullary nail (58%), a partial or total hip prosthesis (30% and 1%, respectively), or hip screws (11%). The German versions of the LEM and the SF-36 were completed by the patients at a mean of 4.9 ± 3.3 days after operation

(baseline), at 3 months (3M), and at 3 months plus 1 week (3M+) for reliability testing. The time point for the 3M follow-up visit was chosen based on the standard follow-up schedule for patients with hip fractures at the selected clinics. Additional follow-up visits were not scheduled to avoid additional burden to these elderly patients. The mean test–retest interval was 8.0 ± 4.0 days. The questionnaires were self-administered after explanation by a trained nurse or doctor. Because 2 patients died and 16 withdrew their informed consent, 88 (83%) patients were available for both the 3M and 3M+ follow-up assessments.

Reproducibility and Internal Consistency

Reproducibility (test–retest reliability) was assessed using the intraclass correlation coefficient (ICC) between the German LEMs completed at the 3M and 3M+ follow-up visits. At these time points, patients have already achieved a stable clinical condition that is necessary to assess reliability between repeated measurements without the influence of factors that might be related to symptom changes in the early rehabilitation or acute postoperative phase. An ICC $\geq .75$ was deemed to indicate reliability.¹⁶ We expected values of $>.85$.¹² Cronbach's α (CA) was used to determine internal consistency. In addition, item–total correlations were calculated. We expected CA values of $>.8$.¹⁷

Floor and Ceiling Effects

Floor and ceiling effects of the LEM were assessed by calculating the percentage of answers scored 1 (for floor effect, worst clinical result) or 5 (for ceiling effect, best clinical result) at the 3M assessment. A proportion of more than 15% was regarded as effect.¹⁸

Construct Validity

Spearman rank correlation coefficients were used to examine construct validity of the LEM relative to the SF-36 v2 at the 3M assessment. The SF-36 is a widely used instrument to assess quality of life that has previously been validated in German and demonstrated its good psychometric properties.¹⁹⁻²¹ As scales of similar content should demonstrate a convergent validity, we hypothesized a strong ($|r| \geq .60$) correlation between the LEM and the physically dominated subscales (physical functioning [pf] and role-physical [rp]) and the Physical Component Summary (PCS) of SF-36.¹² In contrast, we hypothesized a divergent validity with low or only moderate correlations ($|r| < .60$) between the LEM and the SF-36 role-emotional (re) subscale, mental health (mh) subscale, and Mental Component Summary (MCS).

Responsiveness

Effect size (ES) and standardized response mean (SRM) were used to assess responsiveness between the baseline and 3M assessments when compared to the SF-36 pf subscale. We expected ES and SRM values of >0.9 .¹²

Table 2. Construct Validity Between the Lower Extremity Measure and the SF-36.

	n ^a	Mean score (SD)	R ^b	P value ^c
SF-36				
Physical Functioning	79	37.02 (12.00)	.93	<.001
Role-Physical	78	35.25 (10.92)	.73	<.001
Bodily Pain	80	42.86 (11.29)	.65	<.001
General Health	80	48.85 (8.32)	.67	<.001
Vitality	80	49.80 (10.74)	.63	<.001
Social Functioning	80	45.74 (13.00)	.66	<.001
Role-Emotional	79	39.25 (14.48)	.48	<.001
Mental Health	80	49.97 (9.49)	.54	<.001
Physical Component Summary	77	39.37 (10.20)	.83	<.001
Mental Component Summary	77	49.52 (11.52)	.46	<.001

Abbreviations: SD, standard deviation; SF-36, 36-Item Short Form Health Survey.

^aNumber of patients contributing to the calculation of the Spearman rank correlation.

^bSpearman rank correlation between SF-36 subscale/summary score and Lower Extremity Measure (LEM) total score.

^cTwo-sided tests for each SF-36 subscale/summary score and LEM total score are independent.

Statistical Analysis

The LEM and SF-36 scores were reported as mean \pm standard deviation values. A *P* value of .05 was defined as level of significance. The “not applicable” response was treated as missing for all questions except for the 3 valid options of the original English LEM. For patients with more than 3 missing values, the total score was not calculated. For patients with 3 or less missing values, the mean of the respective item of the overall study population was used. The statistical analyses were performed using STATA 13 (StataCorp LP, College Station, Texas).

Results

Cross-Cultural Adaptation and Basic LEM Data

During the translation procedure, Q21 (Preparing light meals) required particular discussion because the German expression of “light meal” means more a healthy than a simple meal. Therefore, the German expression for “simple” (=“einfach”) was used.

More than 3 missing items were found for 17% (18 of 106) of the patients at baseline, 10% (9 of 89) of the patients at the 3M assessment, and 8% (7 of 88) of the patients at the 3M+ assessment. The Q24 (Gardening/yard work) was answered as “not applicable” by 35% of the surveyed patients. At each assessment time point, between 6% and 23% of the patients chose “not applicable” for the following questions: Q2 (Getting in/out bathtub), Q21 (Preparing light meals), Q22 (Tidying, dusting, washing dishes), Q23 (Doing laundry, vacuuming), Q25 (Food shopping), and/or Q29 (Participating in usual leisure activities).

Table 3. Responsiveness of the SF-36 Physical Functioning Subscale and the Lower Extremity Measure (LEM) at 3 Months Following Surgery for Hip Fracture.

Measure	n ^a	Post-op		3 months		Change from post-op to 3 months		ES ^b	SRM ^c
		Mean	SD	Mean	SD	Mean	SD		
SF-36 Physical Functioning	64	26.7	32.8	49.8	28.2	23.1	35.5	0.71	0.65
LEM	64	49.9	32.1	73.1	21.9	23.2	30.7	0.72	0.76

Abbreviations: SD, standard deviation; ES, Effect size; post-op, postoperative; SRM, Standardized response mean; SF-36, 36-Item Short Form Health Survey.

^aNumber of patients contributing to the analysis.

^bEffect size: Calculated as the difference between the mean post-operative score and the mean 3 months score, divided by the standard deviation of the post-operative score.

^cStandardized response mean: Calculated as the mean of the patient-level change from post-op to 3 months divided by the standard deviation of this change.

Reproducibility and Internal Consistency

Test-retest reliability was excellent (ICC = 0.93 [confidence interval = 0.89-0.95]). The CA was .96 for the total score and if single items were removed (Table 1).

Floor and ceiling effects

No floor effect (0%) and only a minor ceiling effect (7.87%) were found for the total score of the LEM. Large ceiling effects were observed for all single items except Q11 (Getting up from kneeling). A floor effect was found for 5 single items (Table 1).

Construct Validity

The LEM showed strong correlations with the SF-36 subscale pf, subscale rp and PCS, and only moderate correlations with the SF-36 re subscale, mh subscale, and MCS (Table 2). All hypotheses were confirmed.

Responsiveness

Responsiveness (from baseline to 3M) was similar for the SF-36 pf subscale and the LEM when using ES (SF-36 pf 0.71 vs LEM 0.72) and better for the LEM when using SRM (SF-36 pf 0.65 vs LEM 0.76; Table 3).

Discussion

The LEM has been validated in German, showing good reliability, validity, and responsiveness in patients with a hip fracture. The mean score of the German LEM at 3 months lies between the scores for 6 weeks and 6 months of the English version and is therefore comparable. The test-retest reliability of the German version is excellent as is that of the original English version.¹² As in the English version, no floor effect was found. While none of the patients showed a ceiling effect in the English version, we observed a minor effect for the total score at 3 months. The slightly younger age of our patient population may have contributed to this finding.

The CA values of .96 for the German LEM total score and all items show very high internal consistency. However, this indicates a possible redundancy of items, that is, a too narrow questionnaire with too many questions repeatedly asking the

same content in a different way.^{22,23} Although CA values have not been reported for the original, English version and, therefore, they cannot be compared, such high CA values suggest that the questionnaire should be shortened.

Construct validity of the German LEM in relation to the SF-36 was similar to that of the English version, with the strongest correlation observed between the LEM and the pf subscale of the SF-36.¹² Only moderate correlations were found for the SF-36 subscales and summary score that measure a different construct, that is, re, mh, and the MCS.

Although the LEM contains a number of hip-specific questions, greater sensitivity to change compared to the SF-36 pf subscale was only observed when using SRM. The ES values between the baseline (postoperative) and 3M assessments were similar for the LEM and the SF-36 pf subscale. In comparison with the English version, the German LEM demonstrates a smaller effect.¹² The reason might be that the assessment time points are different between the 2 versions. This would also explain the lower responsiveness of the SF-36 in our study compared to data reported for the English version. As in the original validation, our results support the usability of the SF-36 pf subscale in patients with hip fractures.¹²

The assessment of patient's ability to understand and correctly answer to questions was left to the investigator during the informed consent procedure, and no specific test to assess cognitive impairment was used. Although this is a limitation of our study, it was skipped to limit patient's burden in the postoperative phase. Other limitations of the study are its small sample size and the fact that the patients answered more than the 3 valid questions as "not applicable." Although this is not allowed in the original questionnaire, it shows the needs and the varying characteristics of an elderly patient population. Patients of older age are often dependent on their relatives, live in nursing homes, or receive community-based support. Consequently, they need a suitable questionnaire that allows an adapted assessment of lower extremity function with clear rules to calculate a total score with missing or "not applicable" answers. The answers to the single items can show the clinician which individual tasks the patient is able to perform and to what degree. But in its current version, the total score of the German LEM shows a high redundancy of items and does not result in better responsiveness than the pf components of the SF-36.

Based on its psychometric properties, the German LEM can be used as a self-assessment outcome measure in German-speaking patients with hip fractures and to provide answers regarding specific activities in the rehabilitation phase. However, the added value of this 29-item questionnaire relative to the SF-36 in these often frail, sometimes physically and cognitively impaired patients should be better scrutinized. Further research is needed to develop a comprehensive outcome instrument for patients with hip fractures that acknowledges their individual limitations by allowing a “not applicable” option for all task questions and that provides clear calculation rules considering the influence of its use on the total score.

Appendix A

Untere Extremitäten-Fragebogen LEM-D

Liebe/r Patient/in,

Die folgenden Fragen beziehen sich auf Ihre Fähigkeit, Tätigkeiten auszuführen, die im täglichen Leben üblich sind. Bitte wählen Sie die Antwort, die Ihre Schwierigkeit mit der

genannten Aktivität am besten beschreibt. Manche werden sehr einfach für Sie sein, andere sehr schwierig oder gar unmöglich.

Beispiel. Knien ist:

1. unmöglich
2. sehr schwierig
3. mittelmässig schwierig
4. etwas schwierig
5. überhaupt nicht schwierig

Bemerkungen. Sie sollten mit “unmöglich” antworten, wenn die Tätigkeit etwas ist, das Sie üblicherweise täglich oder wöchentlich tun oder getan haben, nun aber wegen Ihrer körperlichen Einschränkungen wie Schwäche, Steifheit oder Schmerzen nicht mehr machen können.

Falls Sie für bestimmte Tätigkeiten Hilfe bekommen (zum Beispiel bei der Hausarbeit), geben Sie bitte an, wie schwierig es für Sie wäre, wenn Sie keine Hilfe hätten.

	unmöglich	sehr schwierig	mittelmässig schwierig	etwas schwierig	überhaupt nicht schwierig	Diese Tätigkeit betrifft mich nicht
1. Aus dem Bett aufstehen ist						
2. In die oder aus der Badewanne steigen ist						
3. Sich auf die Toilette setzen oder von der Toilette aufstehen ist						
4. Duschen ist						
5. Hosen anziehen ist						
6. Socken oder Strümpfe anziehen ist						
7. Schuhe anziehen ist						
8. Von einem Stuhl aufstehen ist						
9. Aufrecht stehen ist						
10. Knien ist						
11. Vom Knien aufstehen ist						
12. Sich bücken, um etwas aufzuheben, ist						
13. Sitzen ist						
14. Innerhalb des Hauses gehen ist						
15. Treppen heruntergehen ist						
16. Treppen hinaufgehen ist						
17. Ausserhalb des Hauses gehen ist						
18. Eine Steigung hinauf- oder heruntergehen ist						

(continued)

(continued)

	unmöglich	sehr schwierig	mittelmässig schwierig	etwas schwierig	überhaupt nicht schwierig	Diese Tätigkeit betrifft mich nicht
19. Ins oder aus dem Auto steigen ist						
20. Öffentliche Verkehrsmittel benutzen ist						
21. Einfache Mahlzeiten zubereiten ist						
22. Aufräumen, staubwischen, Geschirrspülen ist						
23. Wäschewaschen, staubsaugen (schwere Hausarbeit) ist						
24. Gartenarbeit oder Hofarbeit ist						
25. Lebensmittel einkaufen ist						
26. Mit Freunden treffen ist						
27. Ihre üblichen Alltagsaktivitäten im gewohnten Zeitrahmen erledigen ist						
28. Ihre gewohnten Alltagsaktivitäten erledigen ist						
29. An den gewohnten Freizeitaktivitäten teilnehmen ist						

Vielen Dank für die Beantwortung dieses Fragebogens!

Acknowledgments

We are grateful to Philip Perry (†) for the primary statistical analyses and Irene Santi for pursuing his work. We also want to thank Jan Ljungqvist, Elise Kaegi, and Susann Drerup for the support in the data collection and data checks. The presented study was funded via the AO TK Trauma Network.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Charnley J. The long-term results of low-friction arthroplasty of the hip performed as a primary intervention. *J Bone Joint Surg Br.* 1972;54(1):61-76.
- Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. *J Bone Joint Surg Am.* 1969;51(4):737-755.
- Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg Br.* 1996;78(2):185-190.
- Brazier J, Jones N, Kind P. Testing the validity of the Euroqol and comparing it with the SF-36 health survey questionnaire. *Qual Life Res.* 1993;2(3):169-180.
- Nilsdotter AK, Lohmander LS, Klässbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS) - validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord.* 2003;4:1-8.
- Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to anti-rheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol.* 1988;15(12):1833-1840.
- Hutchings L, Fox R, Chesser T. Proximal femoral fractures in the elderly: how are we measuring outcome? *Injury.* 2011;42(11):1205-1213.
- 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(4):371-383.
- 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(11):1759-1763.
- Swionkowski MF, Engelberg R, Martin DP, Agel J. Short musculoskeletal function assessment questionnaire: validity, reliability, and responsiveness. *J Bone Joint Surg Am.* 1999;81(9):1245-1260.
- Davis AM, Wright JG, Williams JI, Bombardier C, Griffin A, Bell RS. Development of a measure of physical function for patients with bone and soft tissue sarcoma. *Qual Life Res.* 1996;5(5):508-516.

12. Jaglal S, Lakhani Z, Schatzker J. Reliability, validity, and responsiveness of the lower extremity measure for patients with a hip fracture. *J Bone Joint Surg Am.* 2000;82-A(7):955-962.
13. Beaton D, Bombardier C, Guillemin F, Bosi Ferraz M. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine.* 2000;25(24):3186-3191.
14. Beaton D, Bombardier C, Guillemin F, Bosi Ferraz MB. *Recommendations for the Cross Cultural Adaptation of Health Status Measures.* Rosemont, IL: American Academy of Orthopaedic Surgeons revised; 2002:1-34.
15. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol.* 1993; 46(12):1417-1432.
16. Portney LG, Watkins MP. *Foundations of Clinical Research - Applications to Practice.* 3 rd ed: Upper Saddle River, New Jersey: Pearson Education, INc.; 2009:82.
17. Bland JM, Altman DG. Cronbach's alpha. *BMJ.* 1997;314(7080): 572.
18. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42.
19. Bullinger M. German translation and psychometric testing of the SF-36 Health Survey: preliminary results from the IQOLA Project. International Quality of Life Assessment. *Soc Sci Med.* 1995;41(10):1359-1366.
20. Bullinger M. [Assessment of health related quality of life with the SF-36 Health Survey]. *Rehabilitation (Stuttg).* 1996;35(3): XVII-XXVII.
21. Bullinger M, Alonso J, Apolone G, et al. Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. International Quality of Life Assessment. *J Clin Epidemiol.* 1998;51(11):913-923.
22. Boyle GJ. Does item homogeneity indicate internal consistency or item redundancy in psychometric scales? *Personality Individual Differences.* 1991;12(3):291-294.
23. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ.* 2011;2:53-55.