

<https://helda.helsinki.fi>

Lower extremity functional scale -mittarin suomenkielisen version 155 Rasch-osioanalyysin tulokset nilkka- ja jalkateräpotilailla

Repo, Jussi

2017

Repo , J , Tukiainen , E , Roine , R P , Sampo , M , Henrik , S & Häkkinen , A 2017 , ' Lower extremity functional scale -mittarin suomenkielisen version 155 Rasch-osioanalyysin tulokset nilkka- ja jalkateräpotilailla ' , Suomen ortopedia ja traumatologia , Vuosikerta. 40 , Nro 3 , Sivut 155-159 . < http://www.soy.fi/files/sot_3_2017_web.pdf >

<http://hdl.handle.net/10138/231134>

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

Lower extremity functional scale -mittarin suomenkielisen version Rasch-osioanalyysin tulokset nilkka- ja jalkateräpotilailla

Jussi Repo^{1,2}, Erkki Tukiainen¹, Risto P. Roine^{2,3}, Mika Sampo⁴, Henrik Sandelin⁵, Arja Häkkinen⁶

1. HYKS Plastiikkakirurgian klinikka ja Helsingin yliopisto
2. HUS yhtymähallinto
3. Itä-Suomen yliopisto, Sosiaali- ja terveysjohtamisen laitos
4. HYKS Syöpäkeskus ja Helsingin yliopisto
5. HYKS Ortopedian ja Traumatologian klinikka ja Helsingin yliopisto
6. Jyväskylän yliopisto, Terveystieteiden laitos ja KSKS Fysiatrian poliklinikka

The Lower extremity functional scale (LEFS) is a widely used patient-reported outcome instrument. This study further investigated the measurement properties of the Finnish version of the LEFS. Altogether 182 patients who had undergone foot and ankle surgery for various reasons and had completed the LEFS were included to analyze construct validity, model and individual item fit, and reliability of the instrument using the Rasch measurement theory. The original scale was refined to meet the pertinent psychometric criteria. The new 15-item LEFS scale with four response categories proved to fit the Rasch model better than the original 20-item LEFS. Further studies are still needed to estimate minimal important change and responsiveness of the new version.

Johdanto

Nilkan ja jalkaterän sairaudet ja niiden hoito voivat vaikuttaa toimintakykyyn. Potilaan itse kuvailemaa häirtää tai hyötyä kartoittavat kyselyt antavat arvokasta tietoa hoidon vaikuttavuudesta. Potilaan subjektiivisen toimintakyvyn arviointiin nilkan ja jalkaterän osalta on kehitetty lukuisia mittareita (1,2).

Lower extremity functional scale (LEFS) –mittari täyttää erinomaisesti oleelliset mittareille asetetut psykometriset kriteerit (3,4,5). Sen mittausominaisuudet ovat paremmat verrattuna useisiin muihin spesifeihin nilkan ja jalkaterän mittareihin (1,6). LEFS-mittari on aiemmin validoitu useille eri kielille (7-15) sekä nilkka että jalkateräpotilaille (3, 15-17). Mittari on luotettava, muutosherkkä ja pätevä (3,5). Näitä ominaisuuksia pidetään kaikkein tärkeimpinä potilaan itse täyttämässä kyselyssä (18). LEFS-mittarilla on todettu olevan hyvä sisältövaliditeetti (laajuus,

jolla mittarin toimialue vastaa mittarin kysymyksiä) ja vahva rakennevaliditeetti (taso, jolla mittari mittaa sitä, mitä sen pitäisi mitata) (3,19). Useissa viimeaikaisissa tutkimuksissa on kuitenkin kyseenalaistettu LEFS-mittarin rakennevaliditeettia. LEFS on joko latautunut kahdelle faktorille faktorianalysissä (9,15,20) tai analyysit eivät ole tukeneet mittarin yksiuolotteisuutta (21,22).

Tämän tutkimuksen tavoitteena oli selvittää LEFS-mittarin suomenkielisen version rakenteen luotettavuus ja pätevyys nilkka- ja jalkateräpotilaille. Erityisesti mielenkiinto kohdistui mittarin rakennevaliditeettiin.

Menetelmät

LEFS-mittari koostuu 20 kysymyksestä, joissa on 5 vastausluokkaa (0= Erittäin suuria vaikeuksia tai ei onnistu lainkaan, 1= Melko suuria vaikeuksia, 2 = Kohtalaisia

vaikeuksia, 3 = Lieviä vaikeuksia, 4 = Ei vaikeuksia). LEFS-mittarin ominaisuuksia tutkittiin Raschin mittausteoriolla (23) käyttäen Rumm2030-ohjelmistoa. Analyysiin sisällytettiin potilaita, joille oli suoritettu nilkan ja/tai jalkaterän leikkaus. Tutkimuksesta pois suljettiin potilaat, joilla patologia sijaitti varpaissa. Raschin (osioanalyysin) mittausteorian periaatteita, analyysia ja tulosten tulkintaan käytettäviä kriteereitä on kuvailtu kattavasti toisaalla (24-31). Kuvailimme tärkeimmät tässä nyt raportoitavassa tutkimuksessa käytetyt menetelmät alla.

1. Kysymysten vastauskategorioiden raja-arvoja tutkittiin havainnoimaan vastausvaihtoehtojen porrastuksen järjestelmällisyyttä. Raja-arvoa kuvaa kohta, jossa potilaan vastauksella on 50 % todennäköisyys päätyä jompaankumpaan kahdesta vastauskategorista. Polytomisen mallin (31) raja-arvojen oletettiin olevan säännönmukaisia.

2. Kysymysten sopivuutta arvioitiin selvittämällä logaritmien jäännösarvoja (potilaan ja kysymyksen interaktiota), khii toiseen (χ^2) -arvoja ja kysymykselle ominaisia kuvaajia. Sopivuuden jäännösarvoja välillä -2.5 ja +2.5 pidetään yleisesti hyväksyttävänä (29). χ^2 -arvojen oletettiin olevan ei-merkitseviä Bonferroinin korjauksen jälkeen.

3. Kysymysten riippumattomuutta selvitettiin kahden kysymyksen välisellä jäännöskorrelaatiolla. Vähintään 0.30 käännöskorrelaatiota pidettiin raja-arvona. Tätä korkeammat arvot tarkoittavat heikompa luotettavuutta, virheellisen tarkkuuden todennäköisyyttä ja testauksen huonompaa laatua.

4. Mittarin kohdentamista tutkittiin yksilön ja kysymyksen raja-arvojen sijainneilla. Tällä pyrittiin määrittämään, vastaako kysymysten jakauma otoksessa mitattua rakenteen etäisyyttä.

5. Yksilöiden erottelevuusindeksi (Person separation index; PSI) (32) ja mittarin sisäinen yhteneväisyys (Cronbachin alfa) (33) mitattiin luotettavuuden arvioimiseksi.

6. Hypoteesina oli, että LEFS-mittari olisi yksiulotteinen. Yksiulotteisuuden kriteerinä pidettiin enintään 5 % merkitsevien t-testien prosentuaalista osuutta \pm 1,96 alueen ulkopuolella (34).

Tulokset

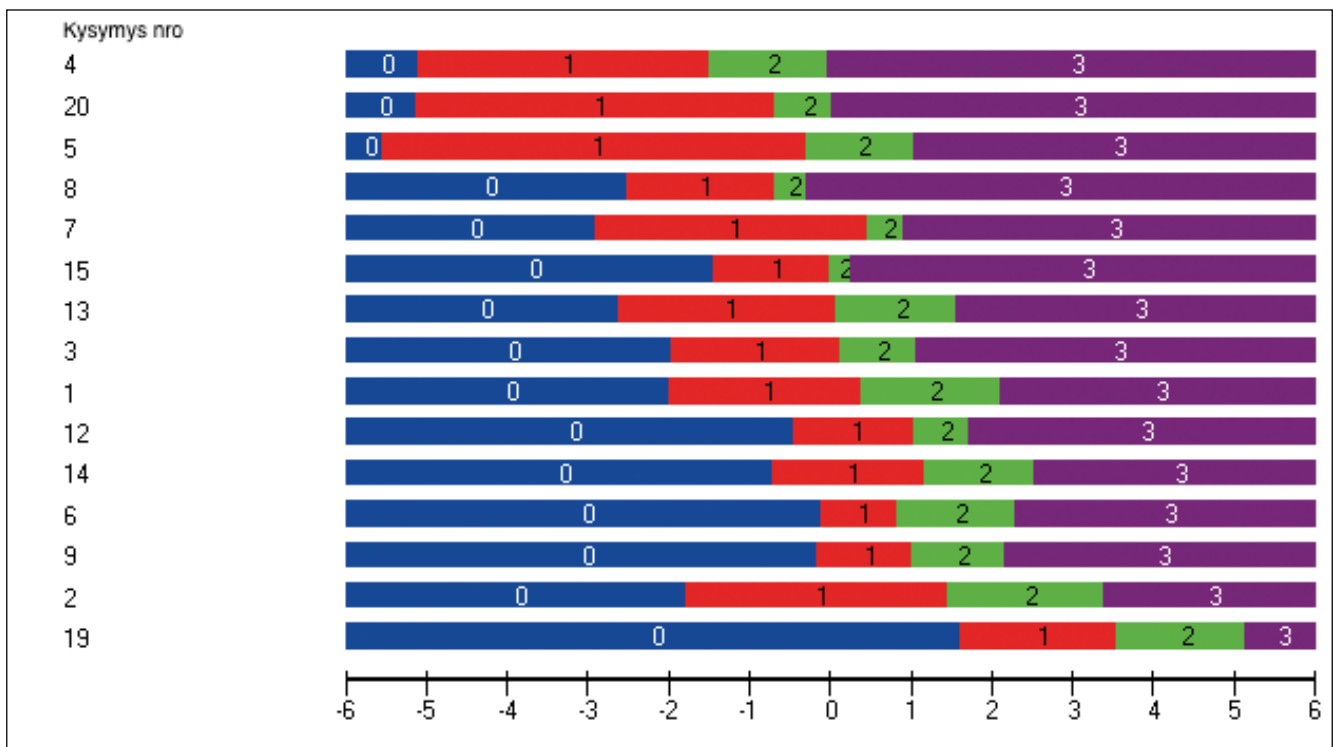
Analyysiin sisällytettiin 182 nilkka- ja jalkateräpotilasta (96 naista, keski-ikä 55 vuotta), jotka oli leikattu trauman (n = 173), infektion (n = 6), kasvaimen (n = 2) tai nivelrikon takia (n = 1). Murtumat olivat joko suljettuja (n = 144) tai avomurtumia (n = 18). Leikkauskohta oli joko nilkka (n = 152), jalkaterän takaosa (calcaneus, talus) (n = 16), jalkaterän keskiosa (naviculare, cuboideum tai cuneiformit) (n = 9) tai jalkaterän etuosa (metatarsaaliluut) (n = 5). Viidellä potilaalla leikkaus tehtiin useampaan (2-3) eri anatomiseen alueeseen.

Kysymysten vastausluokat järjestäytyivät puutteellisesti alkuperäisen mittarin 13/20 kysymyksessä. Kun vastausluokat 1 ("Melko suuria vaikeuksia") ja 2 ("Kohtalaisia vaikeuksia") yhdistettiin, todettiin vastausluokkien epäjärjestys kahden kysymyksen kohdalla (kysymys 10 "Siirtyä autoon ja autosta pois?" ja kysymys 11 "Kävellä 200 metriä?"). Koska kysymysten 10 ja 11 poistamisen jälkeen mittari ei vielä ollu yksiulotteinen, luovuttiin myös kolmesta muusta kysymyksestä (16 "Juosta tasaisella?", 17 "Juosta epätasaisella?" ja 19 "Hyppiä kipeällä jalalla?"), koska niistä luopuminen vahvisti mittarin yksiulotteisuutta.

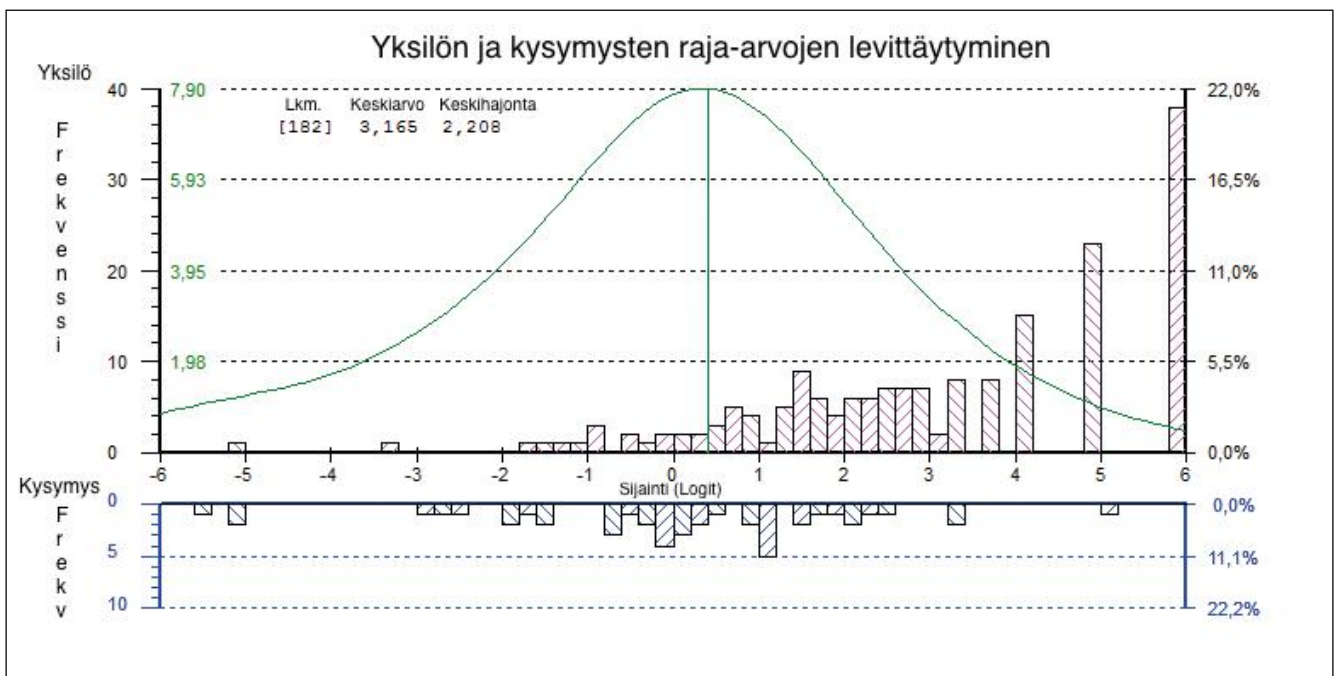
Muokatun version kaikkien 15 kysymyksen vastausluokat olivat järjestäytyneitä (kuva 1). Kysymyksissä ei ollut sopivuuden ylijäämää (kaikki sopivuuden arvot välillä -2,5 ja 2,5). Kysymysten sopivuuden χ^2 -arvot eivät olleet merkitseviä. Jäännöskorrelaatiota oli lopullisessa versiossa kysymysten 6 ("Kyykistyä?") ja 13 ("Istua yhden tunnin ajan?") välillä sekä kysymysten 12 ("Seistä yhden tunnin ajan?") ja 14 ("Tehdä nopeita käännöksiä") välillä. Yksilöiden ja kysymysten raja-arvojen sijainnit jakautuivat tasaisesti (kuva 2). PSI oli 0,85 ja sisäinen yhteneväisyys 0.95. Uusi muokattu 15 kysymyksen mittari, joka sisälsi neljä vastausluokkaa, osoittautui yksiulotteiseksi (merkitsevien t-testien osuus, 1,4 %).

Pohdinta

Kliinisten tulosten arvioinnilla pyritään selvittämään, onko hoidosta hyötyä potilaalle. Hoidon kliinistä tulosta voidaan selvittää esimerkiksi arvioimalla toimintakykyä, oireita, tai terveyteen liittyvää elämänlaatua. Sairauden ja sen hoidon vaikutusta potilaan hyvinvointiin, oireisiin ja toimintakykyyn voidaan käyttää arvioimaan suoria tai epäsuoria hyötyjä. Potilaan kokemaa terveyshyötyä tai -haittaa pystytään



Kuva 1. Uudistetun 4-luokkaisen LEFS-mittarin kysymysten vastausluokkien raja-arvot. x-akselilla mittarin rakenne (toimintakyky). y-akselilla kysymysten vastausluokat: 0, Erittäin suuria vaikeuksia tai ei onnistu lainkaan (sininen); 1, Melko suuria vaikeuksia/kohtalaisia vaikeuksia (punainen); 2, Lieviä vaikeuksia (vihreä); 3, Ei vaikeuksia (violetti).



Kuva 2. Kuvassa esitetään yksilö-kysymys raja-arvojen jakautumista. x-akseli kuvaa kysymysten raja-arvojen paikkaa tai vaikeusastetta (alempi puolisko) ja yksilöiden ilmoittamaa paikkaa tai toimintakyvyn tasoa (ylempi puolisko). x-akseli kuvaa mittarin rakennetta (toimintakykyä), jossa vasemmalle sijoittuvat potilaat, joilla on huonompi toimintakyky. y-akselilla nähdään kysymysten raja-arvojen (alempi puolisko) ja yksilöiden (ylempi puolisko) frekvenssi. Käyrän mukaisesti mittari soveltuu parhaiten potilaille, jotka sijoittuvat -1 ja 1.5 logitin välille. Ryhmitys on asetettu 0,20 välipituuksille tehden 60 ryhmää.

luotettavimmin arvioimaan potilaan itse täyttämällä kyselyillä. Tutkimuksen arviointityökaluksi tai kliiniseen arvioon valittavasta mittarista tulee tietää, soveltuuko se kyseiseen tarkoitukseen ja mittaako se oikeaa asiaa. LEFS-mittari koostuu eri vastausluokkiin jaotelluista toimintakykyä kartoittavista kysymyksistä. Sen antamat tulokset mahdollistavat johtopäätökset toimintakyvystä nilkka- ja/tai jalkaterävampapotilailla. Mittarin rakennevaliditeetista on käyty viimeaikoina kriittistä keskustelua. Tämä tutkimus osoitti, että uusi LEFS-mittarin muokattu versio mittasi luotettavasti ja pätevästi nilkka- ja jalkateräpotilaiden toimintakyvyn tasoa. Sen rakennevaliditeetti oli vahvempi kuin alkuperäisessä versiossa.

Raschin mittausteorialla pystytään arvioimaan testin tai mittarin ominaisuuksia osittain paremmin kuin klassisella testiteorialla tai osiovasteteorialla. Mittareiden ominaisuuksien kartoittamisessa on kuitenkin edelleen näille kaikille kolmelle teorialle oma paikkansa. Niitä voidaan käyttää myös yhdessä tuottamaan lisätietoa mittarin luotettavuudesta ja validiteetista. Klassinen testiteoria on kolmesta mainitusta teoriasta yleisimmin kliinisten mittareiden psykometristen ominaisuuksien testauksessa käytetty. Raschin osioanalyysi perustuu hieman monimutkaisempiin matemaattisiin ja teoreettisiin perusteisiin. Osana psykometrisia tutkimuksia, mittareiden kehittämisessä ja olemassa olevien mittareiden parantamisessa Raschin mittausteoria on erittäin käytännöllinen (26,29,35).

Alnahdin sekä Bravinin ja työtovereiden tutkimukset ovat aiemmin osoittaneet, että alkuperäinen LEFS ei sopinut Raschin malliin heterogeenisessä alaraajapotilasaineistossa, joilla oli lonkan, reiden, polven, säären, nilkan tai jalkaterän vaivaa (21,22). Molemmissa tutkimuksissa LEFS-mittarista poistettiin 5 erillistä kysymystä. On mahdollista että eri kysymykset toimivat eri tavalla riippuen potilasotoksesta. Alnahdi yhdisti vain osassa kysymyksistä vastauskategoriat (21). Sen sijaan Bravini työtovereineen yhdisti vastausluokat 0 ("Erittäin suuria vaikeuksia tai ei onnistu lainkaan") ja 1 ("Melko suuria vaikeuksia") (22). Poistamalla kysymyksiä ja yhdistelemällä vastausluokkia LEFS-mittari saatiin sopimaan Raschin malliin ja sen validiteetti vahvistui (21,22). Myös tässä nyt raportoitavassa tutkimuksessa LEFS-mittarin rakennevaliditeetti parani ja teki mittarin soveliaammaksi nilkan ja jalkaterän toimintakyvyn arviointiin. Uuden version kysymykset ja vastauskategoriat antavat kyseisellä potilasryhmällä tarkempia mittaustuloksia eikä kysymyksissä ole turhaa toistoa.

Tutkimuksessa käytettiin Raschin mittausteoriaa selvittämään suhdetta datan ja teorian välillä. Raschin mallilla saatiin selvitettyä mittarin ominaisuuksia paremmin kuin esim. klassisella testiteorialla tai osiovasteteorialla. Suomenkielisen LEFS-mittarin ominaisuuksia on aiemmin testattu klassisella testiteorialla (15). Tämä nyt raportoitava tutkimus toi oleellista lisätietoa LEFS-mittarin mittaussominaisuuksista nilkka- ja jalkateräpotilailla. Tutkimuksen heikkouksina voidaan pitää poikkileikkausasetelmaa ja pientä otoskoko, joka ei mahdollistanut erottelevaa kysymyksen toiminnan (differential item functioning, DIF) analyysiä. DIF-analyysi sukupuolen tai iän suhteen olisi antanut lisätietoa mittarin rakennevaliditeetista. Tämä tutkimus ei myöskään arvioinut uuden, muokatun mittarin toistettavuutta ja validiteettia suhteessa muihin mittareihin (konvergentti validiteetti), muutosherkkyyttä tai pienintä merkittävää eroa.

Yhteenvedona voidaan todeta, että uusi muokattu LEFS-mittarin suomenkielinen versio on pätevämpi kuin alkuperäinen mittari arvioimaan nilkan ja jalkaterän toimintakykyä. Uuden mittarin ominaisuudet tarvitsevat kuitenkin lisätutkimuksia määrittämään mittarin konvergentti validiteettia, muutosherkkyyttä ja pienintä merkittävää eroa.

Viitteet

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. Jia Y, Huang H, Gagnier JJ. A systematic review of measurement properties of patient-reported outcome measures for use in patients with foot or ankle diseases. *Qual Life Res*. 2017;26:1969-2010
3. 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;79:371-383.
4. Mokkink LB, Terwee CB, Patrick DL et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010 May;19:539-549.
5. Mehta SP, Fulton A, Quach C, Thistle M, Toledo C, Evans NA. Measurement Properties of the Lower Extremity Functional Scale: A Systematic Review. *J Orthop Sports Phys Ther*. 2016;46:200-216.
6. 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.
7. Hoogboom 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. 11
 8. 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.
 9. Naal FD, Impellizzeri FM, Torka S et al. The German Lower Extremity Functional Scale (LEFS) is reliable, valid and responsive in patients undergoing hip or knee replacement. *Qual Life Res* 2015;24:405-410.
 10. 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;113:313-320.
 11. Alnahdi AH, Alrashid GI, Alkhalidi HA, Aldali AZ. Cross-cultural adaptation, validity and reliability of the Arabic version of the Lower Extremity Functional Scale. *Disabil Rehabil.* 2016;38(9):897-904.
 12. Cruz-Díaz D, Lomas-Vega R, Osuna-Pérez MC et al. The Spanish lower extremity functional scale: a reliable, valid and responsive questionnaire to assess musculoskeletal disorders in the lower extremity. *Disabil Rehabil* 2014;36:2005-2011.
 13. Pereira LM, Dias JM, Mazuquin BF, Castanhas LG, Menacho MO, Cardoso JR. Translation, cross-cultural adaptation and analysis of the psychometric properties of the lower extremity functional scale (LEFS): LEFS- BRAZIL. *Braz J Phys Ther.* 2013;17:272-280.
 14. Dell'Era S, Dimaro M, Gamboa A, Spath MB, Salzberg S, Hernández D. [Cross-cultural adaptation and Argentine validation of the Lower Extremity Functional Scale Questionnaire]. *Medicina (B Aires).* 2016;76:279-285.
 15. Repo JP, Tukiainen EJ, Roine RP, Ilves O, Järvenpää S, Häkkinen A. Reliability and validity of the Finnish version of the Lower Extremity Functional Scale (LEFS). *Disabil Rehabil.* 2017;39:1228-1234.
 16. 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;45:1759-1763.
 17. 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.
 18. Hobart JC, Cano SJ, Zajicek JP, Thompson AJ. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations. *Lancet Neurol.* 2007;6:1094-1105.
 19. Yeung TS, Wessel J, Stratford P, 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.
 20. Negahban H, Hessam M, Tabatabaei S et al. Reliability and validity of the Persian lower extremity functional scale (LEFS) in a heterogeneous sample of outpatients with lower limb musculoskeletal disorders. *Disabil Rehabil.* 2014;36:10-15.
 21. Alnahdi AH. Rasch validation of the Arabic version of the lower extremity functional scale. *Disabil Rehabil.* 2016:1-7. doi: 10.1080/09638288.2016.1254285.
 22. Bravini E, Giordano A, Sartorio F, Ferriero G, Vercelli S. Rasch analysis of the Italian Lower Extremity Functional Scale: insights on dimensionality and suggestions for an improved 15-item version. *Clin Rehabil.* 2017;31:532-543.
 23. Rasch G. Probabilistic models for some intelligence and attainment tests. Copenhagen: Danish Institution for Educational Research; 1960.
 24. Andrich D. Rasch models for measurement. Newbury Park CA: Sage, 1988.
 25. Tesio L. Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation research. *J Rehabil Med* 2003;35:105-115.
 26. Tennant A, Conaghan PG. The Rasch Measurement Model in Rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum* 2007;57:1358-1362.
 27. Bond TG, Fox CM. Applying the Rasch model: fundamental measurement in the human sciences, 2nd edn. Mahwah: Lawrence Erlbaum Associates; 2007.
 28. Hagquist C, Bruce M, Gustavsson JP. Using the Rasch model in nursing research: an introduction and illustrative example. *Int J Nurs Stud.* 2009;3:380-393.
 29. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. *Health Technol Assess.* 2009; 13(12). doi:10.3310/hta13120
 30. Ehlan AH, Kucukdeveci AA, Tennant A. The Rasch Measurement Model. In: Franco Franchignoni (Ed). *Research Issues in Physical & Rehabilitation Medicine.* Pavia: Mageri Foundation, p. 89-102, 2010.
 31. Andrich D. Rating scales and Rasch measurement. *Expert Rev. Pharmacoecon Outcomes Res.* 2011;11:571-575
 32. Wright BD, Stone MH. Reliability in Rasch measurement. In *Measurement Essentials*, 2nd edition. 1999. Wide Range, INC. Wilmington, Delaware.
 33. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika.* 1951;16:297-334.
 34. Smith EV. Detecting and evaluation the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement.* 2002;3:205-231.
 35. McClimans L, Browne J, Cano Stefan. Clinical outcome measurement: Model, theory, psychometrics and practice. *Studies in History and Philosophy of Science.* 2017:1-7. Vierailtu 18.9.2017: <http://dx.doi.org/10.1016/j.shpsa.2017.06.004>