
**This is an electronic reprint of the original article.
This reprint *may differ* from the original in pagination and typographic detail.**

Author(s): Kyrölä, Kati; Repo, Jussi; Ylinen, Jari; Häkkinen, Arja

Title: DEPS-testin luotettavuus ja pätevyys aikuisilla selkäpotilailla

Year: 2017

Version:

Please cite the original version:

Kyrölä, K., Repo, J., Ylinen, J., & Häkkinen, A. (2017). DEPS-testin luotettavuus ja pätevyys aikuisilla selkäpotilailla. *Suomen ortopedia ja traumatologia*, 40(3), 227-231. http://www.soy.fi/files/sot_3_2017_web.pdf

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

DEPS-testin luotettavuus ja pätevyys aikuisilla selkäpotilailla

Kati Kyrölä¹, Jussi Repo¹, Jari Ylinen², Arja Häkkinen^{2,3}

1. Keski-Suomen keskussairaala, ortopedian ja traumatologian klinikka

2. Keski-Suomen keskussairaala, fysiatrian klinikka

3. Jyväskylän yliopisto, terveystieteiden laitos

The Depression Scale (DEPS) is a patient-reported outcome instrument developed to recognize depressive symptoms. This study measured the psychometric properties of the DEPS among patients with degenerative spinal disease. Altogether 637 patients were included in the analysis. Main research question was whether the DEPS would meet pertinent psychometric criteria for reliability and validity. The results of analyses supported the DEPS instrument as a reliable and valid test to measure depressive symptoms among adult patients with spinal disease fulfilling the psychometric criteria.

Tausta

Kliinisten tulosten arviointi auttaa selkäsairauksien hoidon kehittämisessä mittaamalla hoidon tehokkuutta ja laatua sekä mahdollistaa tulosten vertailun tieteelliseen näyttöön. Yksi kliinisten tulosten arvioinnin tärkeimmistä osista on potilaan itsensä antamaan arvioon perustuvat mittarit, joilla pystytään arvioimaan potilaan näkökulmaa hoidosta ja arvioida tuloksia joilla on merkitystä potilaalle itselleen. Potilaan itse arvioimaa tilannetta kartoittavien kyselyiden kehittäminen ja niiden mittauseräominaisuuksien arviointimenetelmät ovat kehittyneet ja tulleet laajemman joukon arkikäyttöön. Vaativampaan matematiikkaan perustuvat teoriat mahdollistavat tarkemman selvityksen mittareiden luonteesta. Raschin mittausteoriaa (1) on pidetty ylivoimaisena verrattuna klassiseen psykometriseen teoriaan ja sillä on myös etunsa verrattuna osiovasteteoriaan, vaikka molemmat perustuvat samantyyppiseen latentin ominaisuuden teoriaan. Hyvin toimivat kliiniset vaikuttavuusmittarit mahdollistavat kustannustehokkaiden hoitomuotojen valinnan ja auttavat välttämään turhia

tai liiallisia hoitoja, joista potilas eikä terveyspalvelujen maksaja loppujen lopuksi hyödy.

Depression Scale (DEPS) testi on kehitetty parantamaan yleisten masennusoireiden tunnistamista perusterveydenhuollossa. Se on osoitettu toimivaksi ja validiksi masennuksen seulontatyökaluksi koko väestössä ja testi kykenee tunnistamaan korkean riskin potilaat (2). DEPS on testattu työikäisillä ja toimii parhaiten 18-64-vuotiailla ja pitkälle koulutetuilla tutkittavilla, joilla on yli 9 vuoden koulutustausta (3). Verrattuna muihin depressiomittareihin, DEPS kattaa hyvin olennaiset masennuksen oireet ja on käytössä olevista mittareista yksinkertaisin (4). DEPS-testiä käytetään yleisesti myös tukirankavaivojen liitännäissairauksien seulontaan. Lanneselän spinaalisten oireiden takia leikatuilla potilailla vähäinkin masennusoireilu johtaa pitkäaikaiseen lisääntyneeseen kipuun ja toimintakyvyn laskuun (5). Sen sijaan lanneselän luudutuspotilaat hyötyvät toimintakykykymittareiden perusteella leikkauksesta masennuksesta huolimatta, mutta saavutettu perustaso jää matalammaksi, ja masennuksen seulontaa ja hoitoa suositetaan lanneselän luudutuspotilaille (6). Tutkimuksen tarkoituksena oli testata,

onko DEPS-testi validi ja luotettava degeneratiivista selkäsairautta potevilla potilailla ja soveltuuko mittari tämän populaation masennusoireiden mittaamiseen kun mitattava potilasjoukko ei kokonaan täytä testin optimialuetta iän ja koulutustaustan suhteen (7).

Menetelmät

Potilaat rekrytoitiin 12 kuukauden ajalta vuosina 2012-2013 Keski-Suomen 255 000 asukkaan sairaanhoitopiiriin selkäpoliklinikalta. Kokonaisuudessaan 874 aikuispotilasta lähetettiin pitkittyneen degeneratiivisen selkäsairauden vuoksi konsultaatioon erikoissairaanhoidon fysiatrian ja ortopedian selkäpoliklinikalle sairaanhoitopiiriin hoidonpöytäkirjan mukaisesti. Potilaat edustavat koko sairaanhoitopiiriin väestöpohjaa. Yhteensä 643 potilasta antoi kirjallisen suostumuksen tutkimukseen. Poissulkukriteerit olivat alle 18 vuoden ikä, puuttuvat tai huonolaatuiset röntgenkuvat, akuutit murtumat ja neuromuskulaariset tai pahanlaatuiset sairaudet selkäoireen aiheuttajana. Potilaat täyttivät muun muassa yleistietokyselyn, Oswestry Disability Index (ODI) (8), DEPS-testin, SF-36- (9) ja kipujanakyselyt (10).

DEPS-testi on 10 kysymyksen masennusoireiden seulontaan tarkoitettu kysely, jossa on neljä vastauskategoriaa antaen pisteitä välillä 0-3 (4). DEPS-testi kartoittaa potilaan unettomuutta, surullisuutta, ponnistelun tarvetta, tarmottomuutta, yksinäisyyttä, toivottomuutta, kyvyttömyyttä nauttia asioista, ilottomuutta ja alakuloisuutta. Potilaalla todetaan olevan 50 % todennäköisyydellä masennus, mikäli kokonaispistemäärä ylittää 12 pistettä. Pisteiden noustessa tämä todennäköisyys nousee (Taulukko 1, DEPS-testi).

Raschin mittausteoriaa (1) sovellettiin otokseen käyttäen Rumm2030-ohjelmistoa tutkimaan rakenneliditeettia, mallin sopivuutta ja yksittäisten kysymysten sopivuutta, kysymysten erottelukykyä (Differential Item Functioning, DIF) sukupuolten välillä ja testin luotettavuutta. Raschin mallissa latentin piirteen oletetaan olevan yksiulotteinen. Raschin malli kuuluu niin kutsutun latentin piirteen teorian kattotermin alle. Tämä teoria pohjautuu yksilöiden suoritustasojen suhteesta testin yhteen kysymykseen ja kysymyksen yleiseen kykyyn mitata mitä sen pitäisi mitata. Karkeasti kuvailtuna latentin piirteen teoria koostuu testien suunnittelusta, analyysistä ja pisteytyksestä. Tässä nyt raportoitavassa tutkimuksessa polytominen malli valittiin analyysin perustaksi. Analyysit tehtiin käyttäen tilastolaskennallisia ja ku-

vailevia testejä. Mallin kuvailu, analyysin vaiheet ja psykometriset vaatimukset on kuvailtu tarkemmin toisaalla (1,11-17).

1. Kysymysten vastauskategorioiden raja-arvoja tutkittiin havainnoimaan kysymysten raja-arvoja vastausvaihtoehtojen välillä. Raja-arvo kuvaa kohtaa, jossa vastauksella on 50 % todennäköisyys päätyä jompaan kumpaan vastauskategoriaan. Tässä tutkimuksessa hypoteesiksi asetettiin, että polytomisen mallin raja-arvot olisivat säännönmukaiset.

2. Kysymysten sopivuutta arvioitiin selvittämällä logaritmien jäännösarvoja (potilas-kysymysinteraktiota), khii toiseen (χ^2) -arvoja ja kysymykselle ominaisia käyriä. Sopivuuden jäännösarvoja välillä -2.5 ja +2.5 pidettiin hyväksyttävänä. χ^2 -arvojen hypotettiin olevan ei-merkittäviä Bonferronin korjauksen jälkeen.

3. Oletusta kysymysten riippumattomuudesta mitattiin tutkimalla jäännöskorrelaatiot jokaisen kysymysparin välillä. Arvoa 0.30 tai siitä suurempaa käytettiin tunnistamaan jäännöskorrelaatiot. Korkeammat arvot tulkittiin tarkoittavan heikompaa luotettavuutta ja todennäköisyyttä näyttää virheellistä tarkkuutta ja testauksen laatua.

4. Testin kohdentamista tutkittiin yksilön ja kysymyksen raja-arvojen sijainneilla määrittämään, vastaako kysymysten jakauma otoksessa mitattua rakenteen etäisyyttä. Sukupuolien merkitsevää eroa arvioitiin yksisuuntaisella varianssianalyysillä.

5. Person separation index (PSI) (18) ja Chronbachin alfa (19) (sisäinen yhtenevyys) määritettiin luotettavuuden arvioimiseksi. PSI mittaa, kuinka hyvin mittari erottelee vastaajia. Se arvioi mittarin ominaisuuksia suhteessa tiettyyn potilasotokseen, jota sillä hetkellä mitataan. Luotettavuutta arvioidaan välillä 0 (kaikki virheellistä) ja 1 (ei virheitä). Tutkimushypoteesiksi asetettiin, että arvot olisivat vähintään 0.80.

6. Kysymysten vakauden selvittämiseksi analysoitiin DIF sukupuolelle. DIF todettiin tunnistetuksi, jos χ^2 -arvot olivat merkitseviä Bonferronin korjauksen jälkeen.

7. Lisäksi DEPS-testiä ja SF-36:n kahdeksan osaluueen korrelaatiota verrattiin. SF-36:n kahdeksan

Taulukko 1. DEPS Depression Scale testin kysymykset ja pisteytys (2).

1	Kärsin unettomuudesta	Ei lainkaan 0 p Jonkin verran 1 p Melko paljon 2 p Erittäin paljon 3 p
2	Tunsin itseni surumieliseksi	
3	Minusta tuntui, että kaikki vaati ponnistusta	
4	Tunsin itseni tarmottomaksi	
5	Tunsin itseni yksinäiseksi	
6	Tulevaisuus näytti toivottomalta	
7	En nauttinut elämästäni	
8	Tunsin itseni arvottomaksi	
9	Tunsin, että kaikki ilo on hävinnyt elämästäni	
10	Minusta tuntui, ettei alakuloisuuteni hellittänyt edes perheeni ja ystäväni avulla	

osa-aluetta ovat "Yleinen terveys", "Fyysinen toimintakyky", "Sosiaalinen toimintakyky", "Psykykinen hyvinvointi", "Tarmokkuus", "Kipu", "Fyysinen rooli-toiminta" ja "Psykykinen rooli-toiminta". Korrelaation laskemiseen käytettiin Pearsonin korrelaatiokerrointa.

Tulokset

Kokonaisuudessaan 637 tutkimukseen sopivaa potilasta täytti kyselyt. Keski-Suomen keskussairaalan selkäpoliklinikan aikuispotilaista yhden vuoden otoksen perusteella 30% oli yli 64-vuotiaita ja alle 9 vuoden koulutustausta oli tutkimushetkellä vuonna 2013 noin 15%:lla. Potilaat olivat keskimäärin (SD) 54,8 (15,2) vuotiaita, vaihteluväli 18-88 vuotta. Potilaiden korkeampi ikä korreloi koulutusvuosien vähäisyyteen ($p < 0,01$), korkeamman DEPS-pistemäärään ($p < 0,05$), kehon massaindeksiin ($p < 0,05$), VAS-alaraajakipuun ($p < 0,05$) ja ODI pistemäärään ($p < 0,01$).

DEPS-pisteiden keskiarvo oli 9,2 (6,6; vaihteluväli 0-29). Yli 64-vuotiaiden DEPS-pisteet olivat 9,3 (5,7) ja alle 9 vuotta koulutettujen 10,3 (6,5). Testin minimi- ja maksimipisteet saivat 2,4 % ja 0,8 % potilaista.

Kysymyksissä oli säännönmukaiset raja-arvot (Kuva 1). Kaiken kaikkiaan 7/10 kysymyksessä todettiin hyvä kysymyksen sopivuus. (Yhteensä 15 vastaajalla osa kysymyksistä oli puutteellisesti vastattuja, joten sopivuuden loppuanalyysiin sisällytettiin 622 potilasta.) Sen sijaan kolme kysymystä ylitti asetetut (-2,5 ja +2,5) sopivuuden jäännösarvot ($p < 0,001$): 1 "Kärsin unettomuudesta" (10,1),

9 "Tunsin, että kaikki ilo oli hävinnyt elämästäni" (-4,8), 10 "Minusta tuntui, ettei alakuloisuuteni hellittänyt edes perheeni tai ystäväni avulla" (-3,5). Kahdessa näistä kysymyksestä (1 ja 9) myös χ^2 -arvo oli merkitsevä ($p < 0,0001$). Hyvin pieni jäännöskorrelaatio (-0,325) todettiin kysymysten 1 ("Kärsin unettomuudesta") ja 9 ("Tunsin, että kaikki ilo oli hävinnyt elämästäni") välillä.

DEPS-testin yksiuoloitteisuus vahvistettiin, sillä merkitsevien t-testien osuus oli 4,5 % (hyväksyttävä merkitsevien t-testien osuus on $\leq 5,0$ %). Kysymysten raja-arvojen jakautuminen ja yksilön (vastaajan) arviot olivat yhdenmukaiset paitsi jatkumon ääripäissä havaittujen lievien puutteiden osalta (kuva 2). Sukupuolella ei ollut merkitsevää eroa raja-arvojen jakautumisessa ($p = 0,10$).

PSI oli 0,88. Cronbachin alfa oli 0,92 viitaten DEPS-testin hyvään sisäiseen yhtenevyyteen.

Yhdenmukainen DIF ($p < 0,001$) löydettiin naisia kohden kysymyksessä 3 ("Minusta tuntui että kaikki vaati ponnistusta") ja miehiä kohden kysymyksessä 7 ("En nauttinut elämästäni").

DEPS-testin korrelaatio oli vahva ($p < 0,001$) kaikkien SF-36:n osa-alueiden kanssa. Vahvin korrelaatio ($r = 0,79$) löydettiin psykikistä hyvinvointia kartoittavasta osa-alueesta.

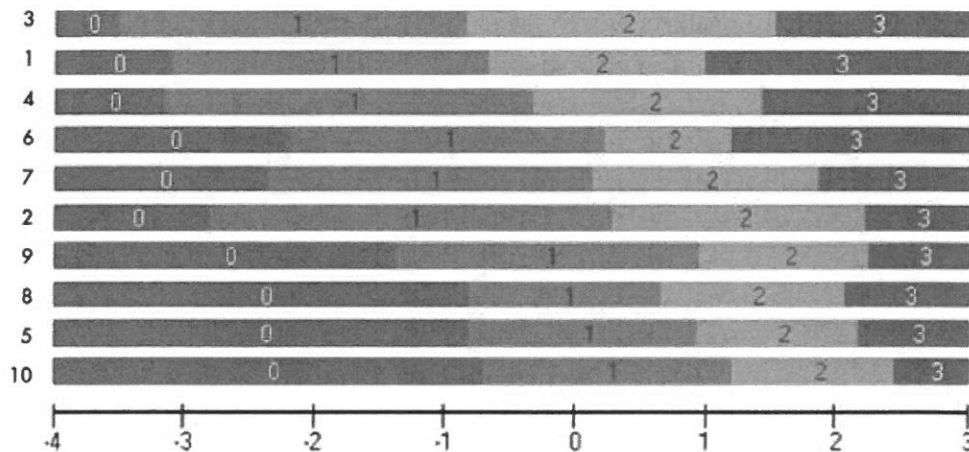
Pohdinta

Tämän tutkimuksen tulos tukee DEPS-testin soveltuvuutta aikuisille selkäpotilaille valikoimattomassa populaatiossa. Testi täytti oleellimmat Raschin mallin vaatimukset ja mittaria voidaan pitää päte-

A

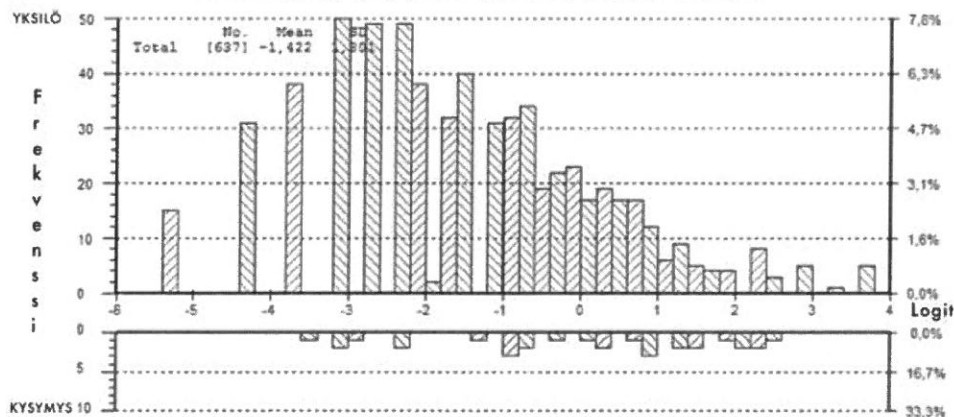
Kysymyksen numero

Vastauskategoriat



B

Yksilön ja kysymysten raja-arvojen jakautuminen



Kuva 1. 1A: Vastausten raja-arvokartta kaikille DEPS-testin kysymyksille. X-akselilla rakenne (masennuksen taso). Palkeissa vastauskategoriat: 0: Ei lainkaan (sininen); 1: Jonkin verran (punainen); 2: Melko paljon (vihreä); 3: Erittäin paljon (violetti).

1B: Testin kohdentaminen otokseen, yksilön ja kysymysten raja-arvojen jakautuminen. X-akselilla rakenne (masennuksen taso), jossa masennuksen taso lisääntyy vasemmalle. Y-akselilla yksilön mittauskohdan tiheys (ylempi histogrammi) ja kysymyksen sijaintia (alempi histogrammi).

vänä ja luotettavana aikuisten selkäpotilaiden masennusoireiden kartoittamiseen.

Vain kolme kysymystä ylitti asetetut sopivuuden jäännösarvojen rajat. Jäännösarvojen ylittävissä kysymyksissä voidaan epäillä olevan turhaa toistoa ja ne voivat olla kokonaisuuden kannalta tarpeettomia. Tämä nyt raportoitava tutkimus vahvisti hypoteesin siitä, että DEPS-testi toimii yksiulotteisesti aikuisilla, joilla on pitkäaikainen selkäsairaus, sillä merkitsevät t-testit eivät

ylittäneet raja-arvoa (5 %). Tulos antaa oleellista lisätietoa mittarin rakenneliditeetistä ja vahvistaa sen mittaavan yhtä rakennetta – masennusoireita.

DIF analyysillä voidaan arvioida kysymysten ennakointia käytöstä eri sukupuolilla. Naisilla oli hieman suurempi todennäköisyys vastata paremmin (saada korkeammat pisteet) kysymyksessä 3 ”Tunsin itseni surumieliseksi”. Miehillä oli sen sijaan suurempi todennäköisyys saada korkeammat pisteet kysy-

myksessä 7 (En nauttinut elämästäni) 1.5 alemman luokkavälin kohdalla, mutta ylempien luokkavälien kohdalla tulokset tasoittuvat. Tulkittaessa potilaiden vastauksia tulee ottaa huomioon, että kysymyksissä 3 ja 7 on eroa sukupuolelten välisissä vastauksissa.

Tutkimuksen vahvuuksia oli suuri otoskoko. Siitä pystyttiin erinomaisesti tekemään riittävällä tilastolaskennallisella voimalla analyysijä ja mittaamaan DIF:ä sukupuolelle. Tämän tutkimuksen tarkoituksena oli arvioida DEPS-testin psykometrisia ominaisuuksia nimenomaan degeneratiivisesta selkäsairautta sairastavilla potilailla. Oros koostui 72,9 %:sta erikoissairaanhoidon selkäpoliklinikalle pitkittyneen degeneratiivisen selkäsairauden vuoksi vuoden aikana konsultaatioon lähetetyistä potilaista. Tutkimukseen osallistumisprosenttia voidaan pitää merkittävänä ja se todennäköisesti kattaa riittävältä osin selkäpoliklinikalle lähetettävän potilasmateriaalin. Tämän nyt raportoitavan tutkimuksen tuloksia voidaan yleistää pitkittyntä selkäsairautta poteville. Tutkimuksen rakenne oli poikkileikkaustutkimus ja se voidaan tulkita myös tutkimuksen heikkoudeksi.

Yhteenveto

DEPS-testin todettiin täyttävän oleellimmat Raschin mallin vaatimukset ja mittaria voidaan pitää pätevänä ja luotettavana aikuisten selkäpotilaiden masennusoireiden kartoittamiseen. Tutkittu selkäpotilasjoukko on edustava otos yhden sairaanhoitopiirin väestöstä aikuisten pitkittyntä selkäsairautta potevien potilasryhmässä, joten tulos on yleistettävissä tähän populaatioon.

Viitteet

1. Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. Danish Institute for Education Research, reprinted: MESA Press, 1993.
2. Salokangas RK, Poutanen O, Stengard, E. Screening for depression in primary care. Development and validation of the Depression Scale, a screening instrument for depression. *Acta Psychiatrica Scandinavica*, 1995; 92, 10-16.
3. Poutanen O, Koivisto AM, Salokangas RK. The Depression Scale (DEPS) as a case finder for depression in various subgroups of primary care patients. *Eur Psychiatry*. 2008; 23:580-6.
4. Poutanen O, Koivisto AM, Salokangas RK. Applicability of the DEPS Depression Scale: assessing format and individual items in subgroups of patients. *Nord J Psychiatry*. 2010; 64:384-90.
5. Tuomainen I, Pakarinen M, Aalto T, Sinikallio S, Kröger H, Viinamäki H, Airaksinen O. Depression is associated with the long-term outcome of lumbar spinal stenosis surgery: a 10-year follow-up study. *Spine J*. 2017; 16: S1529-9430. doi: 10.1016/j.spinee.2017.08.228. [Epub ahead of print]
6. Wahlman M, Häkkinen A, Dekker J, Marttinen I, Vihtonen K, Neva MH. The prevalence of depressive symptoms before and after surgery and its association with disability in patients undergoing lumbar spinal fusion. *Eur Spine J*. 2014; 23:129-34.
7. Kyrölä K, Repo J, Mecklin JP, Ylinen J, Kautiainen H, Häkkinen A. Spinopelvic Changes Based on the Simplified SRS-Schwab Adult Spinal Deformity Classification: Relationships with Disability and Health-related Quality of Life in Adult Patients with Prolonged Degenerative Spinal Disorders. *Spine (Phila Pa 1976)* 2017 Aug 1. doi: 10.1097/BRS.0000000000002370. [Epub ahead of print]
8. Pekkanen L, Kautiainen H, Ylinen J, Salo P, Hakkinen A. Reliability and validity study of the Finnish version 2.0 of the Oswestry disability index. *Spine (Phila Pa 1976)* 2011;36:332-8.
9. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. *Ann Med* 2001;33:350-7.
10. Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain* 1983;17:45-56.
11. Tesio L. Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation research. *J Rehabil Med* 2003; 35: 105-115.
12. 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.
13. Bond TG, Fox CM. Applying the Rasch model: fundamental measurement in the human sciences, 2nd edn. Mahwah: Lawrence Erlbaum Associates; 2007.
14. 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 Dec;6(12):1094-105.
15. 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.
16. 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): 1-177.
17. Ehlan AH, Kucukdeveci AA, Tennant A. The Rasch Measurement Model. In: Franco Franchignoni (Ed). *Research Issues in Physical & Rehabilitation Medicine*. Pavia: Maugeri Foundation, 2010; p. 89-102.
18. Wright BD, Stone MH. Reliability in Rasch measurement. In *Measurement Essentials*, 2nd edition. 1999; 157-166. Wide Range, INC. Wilmington, Delaware.
19. Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychol Bull*. 1955;52(4):281-302.