The Spine Journal 19 (2019) 1057–1066

Review Article

Cross-cultural adaptation of the Neck Pain and Disability Scale: a methodological systematic review

Min Yao, PhD[a,b], Bao-ping Xu, MD[a,c,1], Zi-rui Tian, BS[a,b], Jie Ye, MD[d], Ying Zhang, PhD[e], Yong-jun Wang, PhD[a,b,**], Xue-jun Cui, PhD[a,b,*]

[a] *Spine Disease Institute, Longhua Hospital, Shanghai University of Traditional Chinese Medicine; 725 South Wanping Rd, Shanghai 200032, China*

[b] *Key Laboratory of Theory and Therapy of Muscles and Bones, Ministry of Education (Shanghai University of Traditional Chinese Medicine); 725 South Wanping Rd, Shanghai 200032, China*

[c] *Traditional Chinese Hospital of Lu'an; 76 Renmin Rd, Lu'an 237000, China*

[d] *Orthopedics and Traumatology, Longhua Hospital, Shanghai University of Traditional Chinese Medicine; 725 South Wanping Rd, Shanghai 200032, China*

[e] *Shanghai Changzheng Hospital; 415 Fengyang Rd, Shanghai 200003, China*

Received 19 November 2018; revised 25 January 2019; accepted 25 January 2019

FDA device/drug status: Not applicable.

Author disclosures: ***MY***: Grants: National Natural Science Foundation of China (D, paid directly to institution), Municipal Science and Technology Commission of Shanghai (F, paid directly to institution), Shanghai TCM Medical Center of Chronic Disease (F, paid directly to institution), National Thirteenth Five-Year Science and Technology Major Special Project for New Drug Innovation and Development (F, paid directly to institution), Program for Innovative Research Team of Ministry of Science and Technology of China (F, paid directly to institution), Innovation Team (F, paid directly to institution), Three Years Action to Accelerate the Development of Traditional Chinese Medicine Plan (F, paid directly to institution). ***BPX***: Grants: National Natural Science Foundation of China (D, paid directly to institution), Municipal Science and Technology Commission of Shanghai (F, paid directly to institution), Shanghai TCM Medical Center of Chronic Disease (F, paid directly to institution), National Thirteenth Five-Year Science and Technology Major Special Project for New Drug Innovation and Development (F, paid directly to institution), Program for Innovative Research Team of Ministry of Science and Technology of China (F, paid directly to institution), Innovation Team (F, paid directly to institution), Three Years Action to Accelerate the Development of Traditional Chinese Medicine Plan (F, paid directly to institution). ***ZRT***: Grants: National Natural Science Foundation of China (D, paid directly to institution), Municipal Science and Technology Commission of Shanghai (F, paid directly to institution), Shanghai TCM Medical Center of Chronic Disease (F, paid directly to institution), National Thirteenth Five-Year Science and Technology Major Special Project for New Drug Innovation and Development (F, paid directly to institution), Program for Innovative Research Team of Ministry of Science and Technology of China (F, paid directly to institution), Innovation Team (F, paid directly to institution), Three Years Action to Accelerate the Development of Traditional Chinese Medicine Plan (F, paid directly to institution). ***JY***: Grants: National Natural Science Foundation of China (D, paid directly to institution), Municipal Science and Technology Commission of Shanghai (F, paid directly to institution), Shanghai TCM Medical Center of Chronic Disease (F, paid directly to institution), National Thirteenth Five-Year Science and Technology Major Special Project for New Drug Innovation and Development (F, paid directly to institution), Program for Innovative Research Team of Ministry of Science and Technology of China (F, paid directly to institution), Innovation Team (F, paid directly to institution), Three Years Action to Accelerate the Development of Traditional Chinese Medicine Plan (F, paid directly to institution). ***YZ***: Grants: National Natural Science Foundation of China (D, paid directly to institution), Municipal Science and Technology Commission of Shanghai (F, paid directly to institution), Shanghai TCM Medical Center of Chronic Disease (F, paid directly to institution), National Thirteenth Five-Year Science and Technology Major Special Project for New Drug Innovation and Development (F, paid directly to institution), Program for Innovative Research Team of Ministry of Science and Technology of China (F, paid directly to institution), Innovation Team (F, paid directly to institution), Three Years Action to Accelerate the Development of Traditional Chinese Medicine Plan (F, paid directly to institution). ***YJW***: Grants: National Natural Science Foundation of China (D, paid directly to institution), Municipal Science and Technology Commission of Shanghai (F, paid directly to institution), Shanghai TCM Medical Center of Chronic Disease (F, paid directly to institution), National Thirteenth Five-Year Science and Technology Major Special Project for New Drug Innovation and Development (F, paid directly to institution), Program for Innovative Research Team of Ministry of Science and Technology of China (F, paid directly to institution), Innovation Team (F, paid directly to institution), Three Years Action to Accelerate the Development of Traditional Chinese Medicine Plan (F, paid directly to institution). ***ZJC***: Grants: National Natural Science Foundation of China (D, paid directly to institution), Municipal Science and Technology Commission of Shanghai (F, paid directly to institution), Shanghai TCM Medical Center of Chronic Disease (F, paid directly to institution), National Thirteenth Five-Year Science and Technology Major Special Project for New Drug Innovation and Development (F, paid directly to institution), Program for Innovative Research Team of Ministry of Science and Technology of China (F, paid directly to institution), Innovation Team (F, paid directly to institution), Three Years Action to Accelerate the Development of Traditional Chinese Medicine Plan (F, paid directly to institution).

* Corresponding author. Xue-jun Cui, Spine Disease Institute, Longhua Hospital, Shanghai University of Traditional Chinese Medicine; 725 South Wanping Rd, Shanghai 200032, China, Tel.: 18917763017. ** Yong-jun Wang, Spine Disease Institute, Longhua Hospital, Shanghai University of Traditional Chinese Medicine; 725 South Wanping Rd, Shanghai 200032, China, Tel.: 18917763018.

E-mail addresses: yjwang8888@126.com (Y.-j. Wang), 13917715524@139.com (X.-j. Cui).

[1] Bao-ping Xu contributed equally to this work and should be considered co-first author.

https://doi.org/10.1016/j.spinee.2019.01.007

1529-9430/© 2019 Elsevier Inc. All rights reserved.

Abstract

BACKGROUND: Neck pain is a common and uncomfortable symptom, adversely affecting the work and life of those affected. The Neck Pain and Disability Scale (NPDS) is widely used in neck pain assessment. It has been cross-culturally adapted into several languages to extend its reach to non–English-speaking countries. The aim of this study was to comprehensively evaluate the translation procedures and measurement properties of cross-cultural adaptations of the NPDS.

METHOD: We searched multiple databases, including PubMed, Embase, CINAHL, SciELO, PsycINFO, Medline, SinoMed, PsycINFO, Web of Science, and Scopus, using the keywords "Neck Pain and Disability Scale," "NPDS," "cross-cultural," and "translation". Cross-cultural adaptation and quality control of measurement properties of adaptation procedures were independently conducted by two reviewers in accordance with *Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures* and *Quality Criteria for Psychometric Properties of Health Status Questionnaire*.

RESULTS: There are 15 adaptations of NPDS in 11 different languages with multiple versions in Korean, simplified Chinese and Turkish with 19 studies. In about half of these studies, forward and back translations were conducted. Specially, they mostly focus on the synthesis phase of the translations. Only the simplified-Chinese-2011 adaptation meets the standards of composition due to the existence of an expert committee. Internal consistency, reliability, and construct validity have been evaluated in most existing eligible articles. Half of these articles tested ceiling and floor effects, and only a few included agreement responsiveness and interpretability.

CONCLUSION: The Italian (publication 1 and 2), Persian-Iranian, simplified-Chinese-2011, and Thai adaptations show better quality than others with regard to cross-cultural adaptation and measurement properties. Further studies should fully assess the measurement properties of the NPDS in the Dutch (publication 1 and 2), Hindi-Indian, Korean-2013, simplified-Chinese-2010, Turkish-2004, and Turkish-2007 adaptations. © 2019 Elsevier Inc. All rights reserved.

Keywords: Neck Pain and Disability Scale; Systematic review; Cultural adaptation; Translation; Measurement property; Validity

Introduction

The Neck Pain and Disability Scale (NPDS) is one of the most frequently used neck disability questionnaires for measuring neck pain and related disability [1]. It lists 20 items related to neck movements, neck pain intensity, the effect of neck pain on emotion and cognition, as well as level of interference with daily life activities. The NPDS is easy to be understood and used for both patients and clinical investigators [2]. In addition, the NPDS's reliability, face validity, and construct validity have been affirmed [2].

The NPDS has been translated into various languages, such as French [3], German [4,5], Portuguese [6], and Turkish [7]. Cultural and linguistic variation makes instrument translation problematic, and directly translating the original version of a questionnaire does not guarantee similar measurement properties because of cultural differences. Consequently, rough translations may lead to construct bias, method bias, and item bias, all of which impact the validity of cross-cultural comparisons [8]. The regular process of cross-cultural adaptation includes translation, synthesis, back translation, expert committee review, pretesting, and appraisal of the adaptation process [9]. The main measurement properties that should be assessed include internal consistency, criterion validity, construct validity, agreement, reliability, responsiveness, floor and ceiling effects, as well as interpretability [10].

A systematic review about translated versions of neck-specific questionnaires was published in 2011 [11]. However, due to the early publication of the systematic review, its conclusion is unsuitable for current application, as there were many adaptations published after the systematic review [11]. Consequently, an updated systemic review on the quality of the cross-cultural adaptations of the NPDS would be useful. Therefore, the aim of the present study was to make a comprehensive evaluation of the translation procedures and the measurement properties of cross-cultural adaptations of the NPDS.

Materials and methods

This study is a systematic review of all original published articles on the cross-cultural adaptations of the NPDS, following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses [12].

Search strategy

Search strategies for this systematic review were designed to assess the cross-cultural adaptation of the NPDS. An articles search was performed using PubMed, Embase, CINAHL, SciELO, PsycINFO, Medline, SinoMed, PsycINFO, Web of Science, and Scopus. The search terms were "Neck Pain and Disability Scale," "NPDS," "cross-cultural," "equivalence," "translation," "validation," and "adaptation". Additionally, hand searches of journals, references list, conference papers, and textbooks related to the NPDS were comprehensively performed. The initial search commenced in March 2017, and the final search was performed in July 2018.

Selection/eligibility criteria

The following study papers were included:

(1) Studies related to the cross-cultural adaptation development of the NPDS;
(2) Studies that report the process of cross-cultural adaptation;
(3) Studies on the quality assessment of at least one measurement property of a cross-cultural adaptation;
(4) Studies examining adults (aged >18 years) with neck pain who had provided signed informed consent.

No language restrictions were imposed. Studies published in any language could be selected if they meet the criteria. Studies for which full articles or the detailed adaptation process were unavailable owing to the authors being uncontactable were excluded.

Two reviewers checked the potentially relevant studies according to the inclusion and exclusion criteria and selected eligible studies independently. Any disagreement was resolved through discussion with the third reviewer.

Data extraction and quality assessment

The language, population, publication year, and application time of all included studies were extracted by two independent reviewers in a predefined form and the third reviewer verified the information to ensure study characteristics were accurately.

The translation and cross-cultural adaptation methods of each study were classified according to *Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures* [9]. This process of translating instruments into a new language and culture would involve forward translation, the synthesis of the translations, back translation, the use of an expert committee, a test on the prefinal version, and an appraisal of the adaptation process. Each step was classified as positive (+) (when it was taken in accordance with the criteria), doubtful (?) (when its description was unclear), negative (−) (when it was taken incorrectly), or zero (0) (when there was not enough information to evaluate each step). These procedures are described in more detail in Additional File 1.

Measurement properties were classified according to the *Quality Criteria for Psychometric Properties of Health Status Questionnaire*, which focuses on the assessment of psychometric properties [10]. The measurement properties evaluated in this study were data quality (ceiling and floor effects), internal consistency, criterion validity, construct validity, reproducibility (agreement and reliability), and responsiveness. These procedures are described in more detail in Additional file 2. The standard error of measurement (SEM), smallest detectable change (SDC), and minimally important difference (MIC) were also extracted.

The summary of quality assessments was made based on the overall assessment of the measurement properties checklist. Only the adaptations that meet half of the measurement properties with results meeting standards were recommended.

Reviewers' qualification

Quality assessment was performed by the two independent reviewers. The results were adopted on the premise that the weighted Kappa (Kw) was more than 0.75. Any disagreement was resolved by consensus. When consensus could not be reached, the third reviewer adjudicated.

The first and third reviewers arbitrated disagreements and were specialized in cross-cultural adaptation. These two reviewers were also responsible for training the other reviewers. After training, the reviewers were given five papers about cross-cultural adaptation. Only those reviewers who had a Kw over 0.75 with the first reviewers conducted data extraction and the quality appraisal process.

Results

In this study, a total of 215 publications were identified through keyword search. Up to 46 duplications were excluded. Among the remaining 169 articles, 46 studies were considered irrelevant, and 11 were reviews. Finally, 27 were identified as potentially relevant publications after screening their titles and abstracts (Fig. 1).

There were five excluded studies that were subsequent research of the original NPDS [2,13–16]. One study is a duplication of the French adaptation [16]. We were unable to retrieve the cross-cultural adaptation process of the Japanese and Finnish versions, even after attempting to contact the author, so these adaptations were excluded [17,18].

Among these 19 publications, 15 adaptations of the NPDS are currently available in 11 different languages [3–7,19–32]. Two adaptations are available in Korean [26,27], simplified-Chinese [29,30], German [4,5,22], and Turkish [7,32]. Moreover, multiple studies have examined Dutch (publication 1 and 2) [19–21], German (publication 1 and 2) [4,22], and Italian (publication 1 and 2) adaptations [24,25] (Table 1).

The sample size of the studies on validity ranged from 34 to 448. None reported sample size calculation or included patients consecutively. Only the Dutch (publication 1 and 2) adaptation was produced with the approval of the original author [19–21].

Quality assessment of the cross-cultural adaptations of the NPDS

The quality assessment of the adaptation process was made by independent reviewers who achieved a Kw of 0.886. Consensus was achieved on 100% of occasions, when disagreements among reviewers occurred (Table 2).

Most adaptations were reported to be forward translation. However, only five adaptations did not completely meet the requirement that the forward translation process should be completed by one translator with a medical background and another one with no medical background [5,6,26,30,31]. Both translators of the Dutch (publication 1 and 2) adaptation have a medical or clinical background

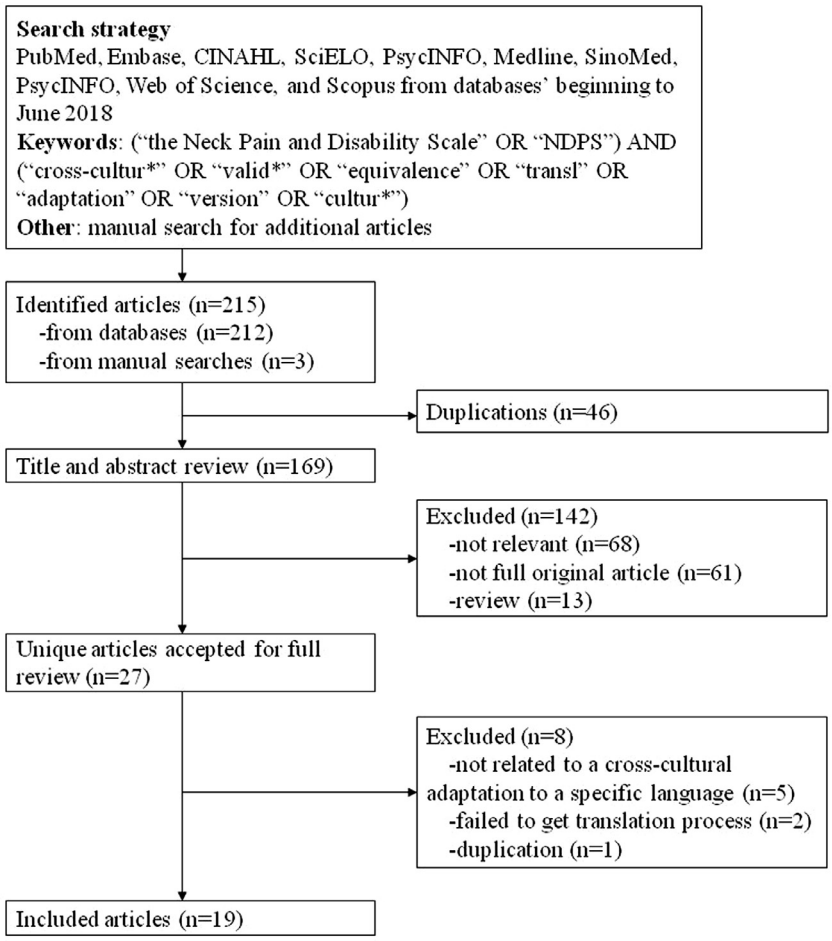

Fig. 1. The summary of the literature identification and selection process.

Table 1
Description of cross-cultural adaptation versions for the NPDS

Language-population	Year	Sample size	Sample size calculation	Agreement of original author	Consecutive or not	Time interval of test-retest reliability
Dutch (publication 1 and 2) [19–21]	2010/2012	34	No	Yes	No	18 d
French [3]	2001	101	No	No	No	1 d
German (publication 1 and 2) [4,23]	2008/2011	448	No	Yes	No	3 mo
German-Swiss [5]	2008	108	No	No	No	1–2 wk
Hindi-Indian [23]	2009	64	No	No	No	2 d
Italian (publication 1 and 2) [24,25]	2008	157	No	No	No	7 d
Korean-2006 [26]	2006	180	No	No	No	7 d
Korean-2013 [27]	2013	81	No	No	No	2 wk
Persian-Iranian [28]	2007	185	No	No	No	2 d
Portuguese-Brazilian [6]	2006	203	No	No	No	7 d
simplified-Chinese-2010 [29]	2010	125	No	No	No	2 separate d
simplified-Chinese-2011 [30]	2011	106	No	No	No	5–7d
Thai [31]	2011	185	No	No	No	2 d
Turkish-2004 [7]	2004	61	No	No	No	NA
Turkish-2007 [32]	2007	102	No	No	No	1–3 d

NA, not available; NPDS, Neck Pain and Disability Scale.
Dutch (publication 1 and 2), two Dutch publications with the same adaptation, the same with German (publication 1 and 2), Italian (publication 1 and 2).

[19–21], and they were aware of the concepts examined in the questionnaire that did not meet the guidelines at that time. One translator of the Turkish-2004 [7] adaptation is a native English speaker. None of the translators of the French [7] and Italian (publication 1 and 2) [24,25] adaptations were familiar with the NPDS. Whether the translators of the Korean-2013 [27] adaptation completed forward translation independently remains unclear. Moreover, we

Table 2
Cross-cultural adaptations of the NPDS that used the translation-based approach relate to the *Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures*

Language-population	Translation	Synthesis	Back translation	Expert committee review	Pretesting	Appraisal of the Adaptation Process
Dutch (publication 1 and 2) [19–21]	–	+	–	0	–	+
French [3]	–	+	–	–	?	+
German (publication 1 and 2) [4, 22]	?	+	?	0	0	+
German-Swiss [5]	+	+	+	–	–	+
Hindi-Indian [23]	?	0	?	0	0	+
Italian (publication 1 and 2) [24, 25]	–	+	–	–	0	+
Korean-2006 [26]	+	+	+	?	0	+
Korean-2013 [27]	?	+	+	–	+	+
Persian-Iranian [28]	?	+	?	–	+	+
Portuguese-Brazilian [6]	+	+	+	?	0	+
Simplified-Chinese-2010 [29]	?	+	?	+	–	+
Simplified-Chinese-2011 [30]	+	+	+	–	+	+
Thai [31]	+	+	+	–	–	+
Turkish-2004 [7]	–	0	0	0	0	+
Turkish-2007 [32]	?	+	?	0	–	+

+, positive rating; ?, doubtful design or method; –, negative rating; 0, no information available; NPDS, Neck Pain and Disability Scale.; Dutch (publication 1 and 2), two Dutch publications with the same adaptation, the same with German (publication 1 and 2), Italian (publication 1 and 2).

do not know whether the mother tongues of the translators of the Persian-Iranian [28] and simplified-Chinese-2010 [29] adaptations are the target language. The German (publication 1 and 2) [4,22] and Turkish-2007 [34] adaptations do not provide an explanation of their translators' specific backgrounds, and the design of the forward translation process for the Hindi-Indian [23] adaptation is doubtful.

Most of the adaptations introduced a synthesis stage of translation, thus they completely meet the requirements of the synthesis process. No relevant information was found about the Hindi-Indian [23] and Turkish-2004 [7] adaptations.

Back translation was an important stage of the translation process. Two translators fully completed the back translation process of the German-Swiss [5], Korean-2006 [26], Korean-2013 [27], Portuguese-Brazilian [6], simplified-Chinese-2011 [30], and Thai [31] adaptations. The Dutch (publication 1 and 2) [19–21], and French [3] adaptations had only one translator to complete the back translation process. One translator of the Italian (publication 1 and 2) [24,25] adaptation is not a native English speaker. As a result, the Persian-Iranian [28], simplified-Chinese-2010 [29], and Turkish-2007 [34] adaptations do not report whether the back translators are native English speakers. The German (publication 1 and 2) [4,22] and Hindi-Indian [23] adaptations do not provide any more details about back translation. The Turkish-2004 [7] adaptation does not claimed whether back translation were conducted.

Only one adaptation meets the standards of composition for the existence of an expert committee [29]. The

Table 3
The clinimetric testing of the NPDS adaptations relate to *Quality Criteria for Psychometric Properties of Health Status Questionnaire*

Language-population	Internal consistency	Construct validity	Agreement	Reliability	Responsiveness	Floor and ceiling effects	Interpretability
Dutch (publication 1 and 2) [19–21]	–	+	+	–	+	–	0
French [3]	0	+	0	+	0	+	0
German (publication 1 and 2) [4,22]	+	+	0	0	0	0	+
German-Swiss [5]	–	?	0	+	0	+	+
Hindi-Indian [23]	–	+	0	+	0	0	0
Italian (publication 1 and 2) [24,25]	+	+	0	+	+	0	+
Korean-2006 [26]	–	0	0	+	0	+	0
Korean-2013 [27]	–	+	0	+	0	0	0
Persian-Iranian [28]	+	+	0	+	0	+	0
Portuguese-Brazilian [6]	+	+	0	+	0	0	0
Simplified-Chinese-2010 [29]	+	+	0	–	0	0	0
Simplified-Chinese-2011 [30]	+	+	0	+	0	+	0
Thai [31]	+	+	+	+	0	+	0
Turkish-2004 [7]	–	+	0	0	0	0	0
Turkish-2007 [32]	?	+	0	+	0	0	0

+, positive rating; ?, doubtful design or method; –, negative rating; 0, no information available; NPDS, Neck Pain and Disability Scale.
Dutch (publication 1 and 2), two Dutch publications with the same adaptation, the same with German (publication 1 and 2), Italian (publication 1 and 2).

Portuguese-Brazilian [6] adaptation does not explain the specific composition of this committee. The Korean-2013 [27], simplified-Chinese-2011 [30], and Thai [31] adaptations lack methodologists. The German-Swiss [5], Italian (publication 1 and 2) [24,25], and Persian-Iranian [28] adaptations lack language professionals. The French [3] adaptation had only one member for its expert committee. No information was observed in the Dutch (publication 1 and 2) [19–21], German (publication 1 and 2) [4,22], Hindi-Indian [23], Korean-2006 [24], Turkish-2004 [7] and Turkish-2007 [32] adaptations.

Pretest is the final stage of the adaptation process. Only three adaptations meet the requirements [27,28,30]. There were not enough translators for the prefinal versions of the Dutch (publication 1 and 2) [19–21], German-Swiss [5], simplified-Chinese-2010 [29], Thai [31], and Turkish-2007 [32] adaptations. The French adaptation does not report its sample size [3]. The German (publication 1 and 2) [4,22], Hindi-Indian [23], Italian (publication 1 and 2) [24,25], Korean-2006 [26], Portuguese-Brazilian [6], and Turkish-2004 [7] adaptations lack information about the pretest stage.

All adaptations submitted a final version and keep track of the translated versions.

Methodology used for property measurement

The Kw of the two reviewers for property measurement was 0.852. The methodological quality and the measurement properties of the trials varied substantially (Table 3).

An analysis of internal consistency was conducted on most adaptations [4–7,19–32]. The German (publication 1 and 2) [4,22], Italian (publication 1 and 2) [24,25], Persian-Iranian [28], Portuguese-Brazilian [6], simplified-Chinese-2010 [29], simplified-Chinese-2011 [30], and Thai [31] adaptations show a good internal consistency with appropriate methods. The Turkish-2007 [32] adaptation does not fully meet the criteria of internal consistency for missing factor analysis. The Dutch (publication 1 and 2) [19–21] adaptation does not present any factor analysis. We gave a negative rating for the internal consistency of the Hindi-Indian [23], Korean-2013 [27], and Turkish-2004 adaptations [7], because their sample sizes are smaller than 100. Based on criteria regarding internal consistency, Cronbach's α should range from 0.70 and 0.95. However, Cronbach's α for the German-Swiss [5] and Korean-2006 [26] adaptations is 0.97 and 0.96, respectively. This reflects an unnecessary duplication of contents or items and indicates redundancy rather than homogeneity. There is no information about the internal consistency of the French adaptation [3] (Table 4). The Dutch (publication 1 and 2) [19–21], German (publication 1 and 2) [4,22], Hindi-Indian [23], Italian (publication 1 and 2) [24,25], Korean-2013 [27], Persian-Iranian [28], Portuguese-Brazilian [4], simplified-Chinese-2010 [29], simplified-Chinese-2011 [30], Turkish-2004 [7] and Turkish-2007 [32] adaptations show good correlation with other known indicators such as the Neck Disability Index, visual analogue scale and Short Form 36 Health Survey. Therefore, we gave a positive rating for construct validity to the (sub)groups with at least 50 patients. Although construct validity was covered in the German-Swiss [5] and Thai [31] adaptations, there is no detailed information regarding construct validity. As a result, we cannot judge whether they fully meet the criterion of construct validity. Only the Korean-2006 [26] adaptation has no relevant information (Table 4).

Only the Dutch (publication 1 and 2) [19–21] and Thai [31] adaptations describe the Bland and Altman method for agreement. This is no visible tendency towards unequal variance of the data and the rest of the adaptations do not provide any information about agreement.

Based on an intraclass correlation coefficient of 0.70 or more, most of the adaptations meet the criterion of reliability. The German (publication 1 and 2) [4,22] adaptation uses Cronbach's α to test reliability, which is considered as a wrong method. The criterion of reliability is that the sample size is at least 50 patients. Nevertheless, the sample sizes of the Dutch (publication 1 and 2) [19–21] and simplified-Chinese-2010 [29] adaptations are less than 50 (Table 4).

The analyses on receiver operating characteristics of the NPDS reveal that the area under the curve is 0.91 in the Italian (publication 1 and 2) [24,25] adaptation. This measurement shows the ability of a questionnaire to measure changes that occurred. The rest of the adaptations have no information about responsiveness.

The Dutch (publication 1 and 2) [19–21], French [3], German-Swiss [5], Korean-2006 [26], Persian-Iranian [28], simplified-Chinese-2011 [30] and Thai [31] adaptations assessed floor or ceiling effects. However, most of the adaptations do not cover floor or ceiling effects. The sample size of the Dutch (publication 1 and 2) [19–21] adaptation is only 34, with a negative rating for floor and ceiling effects.

Most adaptations do not report interpretability, except for the German-Swiss [5], German (publication 1 and 2) [4,21] and Italian (publication 1 and 2) [24,25] adaptations.

Other details

The Dutch [19–21] adaptation reported its SDC and MIC (11.5 and 31.7, respectively), and the German (publication 1 and 2) [4,22] adaptation reported SEM and SDC (0.9 and 3, respectively). While the German-Swiss [5] adaptation reported its SEM (3.80), and the Italian (publication 1 and 2) [24,25] adaptation reported its MIC with 10 using an anchor-based longitudinal approach (Table 4).

Discussion

Summary of evidence

This study aimed to assess the cross-cultural adaptation procedures and the measurement properties of each

Table 4
Supplementary details for the psychometric properties of the NPDS

Language-population	Construct validity	Internal consistency	Reliability	SEM	SDC	MIC
Dutch (publication 1 and 2) [19−21]	Pearson correlations: NDI 0.77; VAS $_{pain}$ 0.54; VAS $_{disability}$ 0.57; SF-36 −0.58 to −0.36	Cronbach's α: 0.93	ICC: 0.76	/	11.5	31.7
French [3]	Spearman rank correlation coefficient: NDI 0.793; NPQ 0.733; VAS $_{disability}$ 0.627; VAS $_{pain}$ 0.515; VAS $_{handicap}$ 0.667; HADS $_{anxiety}$ 0.399; HADS $_{depression}$ 0.493; Neck sensitivity 0.312; Kellgren 0.039; ROM $_{flexion-extension}$ −0.455; ROM $_{rotation}$ 0.278	/	ICC: 0.91	/	/	/
German (publication 1 and 2) [4,23]	Pearson's correlation coefficients: HADS $_{depression}$ 0.2034; HADS $_{anxiety}$ 0.1463; F-SozU −0.1290	Cronbach's α: 0.94	/	0.9	3	/
German-Swiss [5]	/	Cronbach's α: NPDS 0.97; NPDS $_{pain}$ 0.95; NPDS $_{disability}$ 0.97; NPDS $_{neck-specific\ function}$ 0.87	ICC: 0.84−0.97	3.80	/	/
Hindi-Indian [23]	Pearson's correlation coefficients: NPRS $_{neck}$ 0.30; NPRS $_{arm}$ 0.33; NPRS $_{max}$ 0.48; VAS $_{activity}$ 0.15; VAS $_{depression}$ −0.80	/	ICC: NPDS 0.98; NPDS $_{factor1}$ 0.96; NPDS $_{factor2}$ 0.95; NPDS $_{factor3}$ 0.98; NPDS $_{factor4}$ 0.98	/	/	/
Italian (publication 1 and 2) [24,25]	/	Cronbach's α: NPDS 0.942; NPDS $_{subscale1}$ 0.919, NPDS $_{subscale2}$ 0.856; NPDS $_{subscale3}$ 0.889	Spearman's Correlations NPDS 0.913; NPDS $_{subscale1}$ 0.894, NPDS $_{subscale2}$ 0.931; NPDS $_{subscale3}$ 0.925	/	/	10
Korean-2006 [26]	Pearson's correlation coefficients: GPE −0.42	Cronbach's α: 0.96	ICC: 0.90	/	/	/
Korean-2013 [27]	Pearson correlation coefficient: SF-36 −0.48 to −0.19	Cronbach's α: NPDS 0.934; NPDS $_{subscale1}$ 0.911, NPDS $_{subscale2}$ 0.867; NPDS $_{subscale3}$ 0.871	ICC: NPDS 0.90; NPDS $_{subscale1}$ 0.89, NPDS $_{subscale2}$ 0.92; NPDS $_{subscale3}$ 0.91	/	/	/
Persian-Iranian [28]	Pearson correlation coefficient: NPDS $_{neck\ dysfunction\ and\ disability}$ with SF-36 −0.67 to −0.41; NPDS $_{neck\ pain\ intensity}$ with SF-36 −0.54 to −0.24; NPDS $_{neck\ pain\ during\ movement}$ with SF-36 −0.54 to −0.17; NPDS $_{static\ neck\ pain\ problems}$ with SF-36 −0.63 to −0.18	/	ICC: NPDS $_{neck\ dysfunction\ and\ disability}$ 0.95; NPDS $_{neck\ pain\ intensity}$ 0.97; NPDS $_{neck\ pain\ during\ movement}$ 0.92; NPDS $_{static\ neck\ pain\ problems}$ 0.90	/	/	/
Portuguese-Brazilian [6]	Correlation coefficient: NPDS $_{cervical\ dysfunction\ related\ to\ general\ activities}$ with SF-36 −0.46 to −0.07; NPDS $_{cervical\ dysfunction\ related\ to\ activities\ of\ the\ cervical\ spine}$ with SF-36 −0.28 to 0.01; NPDS $_{pain}$ with SF-36 −0.28 to 0.45	Cronbach's α: NPDS $_{cervical\ dysfunction\ related\ to\ general\ activities}$ 0.89; NPDS $_{cervical\ dysfunction\ related\ to\ activities\ of\ the\ cervical\ spine}$ 0.81; NPDS $_{pain}$ 0.72	ICC: NPDS $_{cervical\ dysfunction\ related\ to\ general\ activities}$ 0.91; NPDS $_{cervical\ dysfunction\ related\ to\ activities\ of\ the\ cervical\ spine}$ 0.62; NPDS $_{pain}$ 0.45	/	/	/
simplified-Chinese-2010 [29]	Pearson correlation coefficient: NPDS $_{neck\ dysfunction\ and\ disability}$ with SF-36 −0.61 to −0.20; NPDS $_{pain\ intensity\ during\ movement}$ with SF-36 −0.58 to −0.19; NPDS $_{static\ neck\ pain\ intensity}$ with SF-36 −0.61 to −0.22	Cronbach's α: NPDS $_{neck\ dysfunction\ and\ disability}$ 0.91; NPDS $_{pain\ intensity\ during\ movement}$ 0.88; NPDS $_{static\ neck\ pain\ intensity}$ 0.82	ICC: NPDS $_{pain}$ 0.95; NPDS $_{disability}$ 0.94; NPDS $_{neck\ specific\ function}$ 0.92; NPDS $_{emotional\ and\ cognitive\ influences}$ 0.86	/	/	/

Table 4 (*Continued*)

Language-population	Construct validity	Internal consistency	Reliability	SEM	SDC	MIC
simplified-Chinese-2011 [30]	Correlation coefficient: SF-36 0.721; $NPDS_{pain}$ with SF-36 −0.742 to −0.051; $NPDS_{disability}$ with SF-36 −0.667 to −0.139; $NPDS_{neck\text{-}specific\ function}$ with SF-36 −0.655 to 0.021; $NPDS_{emotional\ and\ cognitive\ influences}$ with SF-36 −0.633 to 0.124	Cronbach's α: $NPDS_{pain}$ 0.935; $NPDS_{disability}$ 0.952; $NPDS_{neck\ specific\ function}$ 0.955; $NPDS_{emotional\ and\ cognitive\ influences}$ 0.910	ICC: $NPPS_{pain}$ 0.813; $NPDS_{pain}$ 0.886; $NPDS_{disability}$ 0.972; $NPDS_{neck\text{-}specific\ function}$ 0.912; $NPDS_{emotional\ and\ cognitive\ influences}$ 0.939	/	/	/
Thai [31]	Spearman rank correlation coefficient: VAS 0.76	Cronbach's α: $NPDS_{disability}$ 0.96; $NPDS_{disability}$ 0.94; $NPDS_{pain}$ 0.90; $NPDS_{neck\ specific}$ 0.92	ICC: $NPDS_{disability}$ 0.88; $NPDS_{disability}$ 0.81; $NPDS_{pain}$ 0.91; $NPDS_{neck\ specific}$ 0.74	/	/	/
Turkish-2004 [7]	Pearson correlation coefficient: PDI 0.51; VAS 0.45; $HAD_{depression}$ 0.35; $HAD_{anxiety}$ 0.33	Cronbach's α: 0.86	/	/	/	/
Turkish-2007 [32]	Pearson correlation coefficient: VAS_{pain} 0.73; $VAS_{disability}$ 0.69; $VAS_{physician's\ assessment}$ 0.46; muscle spasm 0.29; neck sensitivity 0.47; $ROM_{flexion\ and\ extension}$ -0.35; $ROM_{lateral\ flexion}$ −0.33; $ROM_{rotation}$ −0.39; pain with motion 0.49	Cronbach's α: 0.94	ICC: 0.81	/	/	/

NPDS, Neck Pain and Disability Scale; NDI, neck disability index; VAS, visual analogue scale; SF-36, short-form 36 health survey; NPQ, Northwick Park neck pain questionnaire; HADS, hospital anxiety and depression scale; ROM, range of motion; GPE, global perceived effect scale; PDI, pain disability index; SEM, standard error of measurement; SDC, smallest detectable change; MIC, minimally important difference; ICC, intraclass correlation coefficients.

/, no information available; Dutch (publication 1 and 2), two Dutch publications with the same adaptation, the same with German (publication 1 and 2), Italian (publication 1 and 2).

adaptation of the NPDS. We found that the quality assessment of the cross-cultural adaptations of the NPDS, as conducted by expert committee review and by testing the prefinal version were poor. The main reason for this finding is the lack of expert committee members and insufficient number of samples. In terms of several measurement properties including agreement, responsiveness, and interpretability, the quality of most adaptations is insufficient.

Internal consistency

The analysis of internal consistency was conducted on most adaptations and showed good internal consistency with appropriate methods. However, one point needs to be stressed, that is, the NPDS is a multidimensional scale, in accordance with the results of the French [3], German (publication 1 and 2) [4,22], German-Swiss [5], and Hindi-Indian [23] adaptation. For this reason, Cronbach's α for each unidimensional scale or subscale is needed, rather than the whole scale [33].

Construct validity

For hypothesis testing of construct validity, the study should formulate a set of hypotheses about expected relationships between the adaptation and other well-defined and high quality comparator instruments that are used in the field. If possible, also some hypotheses about expected differences between subgroups can also be formulated in advance. This way, the study results are compared against the same hypotheses. The expected direction (positive or negative) and magnitude (absolute or relative) of the correlations or differences should be included in the hypotheses. Without this specification, deciding later whether the results are in accordance with the hypothesis or not is difficult [33].

Reliability

Most of the adaptations reported test-retest reliability, while the time interval ranged from 1 day to 18 days. A short interval increases the risk of recall bias and a long interval increases the risk of a real change in patient status [34]. The method to avoid the risk of recall bias is to prolong the time interval; we consider over 3 days as appropriate. To avoid the risk of a real change, prohibition of treatment is one method, but this does not conform to medical moral principles and a humane approach under certain conditions, and only including patients without change (determined by asking patients if their neck pain condition had changed compared with the first measurement) would be a better method.

SDC, MIC, and responsiveness

The SDC, MIC, and responsiveness are important psychometric properties for a scale, while few adaptations have reported their SDC, MIC, or responsiveness. The statistical significance of a treatment effect is recognized as corresponding to the sample size, rather than the clinical

effect. In contrast, clinical effect refers to the benefits derived from that treatment, its impact upon the patient, and its implications for clinical management of the patient [35,36]. As a yardstick for clinical relevance, changes in scores exceeding the MIC are clinically relevant. These measurements are not only important in clinical decision making (individual level), but also important for power calculations, sample size estimates, and cost evaluations in clinical research (group level) [37]. Conversely, the SDC and SEM were used to estimate the sizes of potential random measurement errors caused by chance variations in case of score changes by repeating the same tests on individuals. Considering the SDC of an outcome measure can help determine whether the change is real, that is, that the change does not exist within the range of error [38,39].

Responsiveness has been defined as the ability of a questionnaire to detect clinically important changes over time, even if these changes are small [40]. Responsiveness is more important to clinical trials. Most of the adaptations did not test responsiveness and the Italian (publication 1 and 2) [24,25] adaptation is a good example.

Limitations of this review

We failed to retrieve the processes of the cross-cultural Japanese [17] and Finnish [18] adaptations, and therefore excluded them from analyses. This is a notable limitation for a systematic review aimed at assessing all original articles that report cross-cultural translations of the NPDS.

Conclusion

The low quality of the NPDS cross-cultural adaptations is due to poor completion of expert committee review and prefinal version test, as well as insufficient clinimetric testing of measurement properties for agreement, responsiveness, and interpretability. The Italian (publication 1 and 2), Persian-Iranian, simplified-Chinese-2011, and Thai adaptations show a better quality than the other adaptations based on the assessment of the two checklists for cross-cultural adaptations and measurement properties. Further studies are still needed to fully assess the measurement properties of the Dutch (publication 1 and 2), Hindi-Indian, Korean-2013, simplified-Chinese-2010, Turkish-2004, and Turkish-2007 adaptations of the NPDS.

Acknowledgment

The authors thank Liwen Bianji, Edanz Group China (www.liwenbianji.cn/ac), for editing the English text of a draft of this manuscript. This work was supported by the National Natural Science Foundation of China (81704096, 81603635), Municipal Science and Technology Commission of Shanghai-TCM key project (16401970100), Shanghai TCM Medical Center of Chronic Disease (2017ZZ01010), National Thirteenth Five-Year Science and Technology Major Special Project for New Drug Innovation and Development (2017ZX09304001), The program for innovative research team of Ministry of Science and Technology of China (2015RA4002), "Innovation Team" development projects (IRT1270), Three Years Action to Accelerate the Development of Traditional Chinese Medicine Plan (ZY(2018-2020)-CCCX-3003).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Supplementary materials

Supplementary material associated with this article can be found in the online version at https://doi.org/10.1016/j.spinee.2019.01.007.

References

[1] Wheeler AH, Goolkasian P, Baird AC, Darden BV, 2nd. Development of the Neck Pain and Disability Scale. Item analysis, face, and criterion-related validity. Spine (Phila Pa 1976) 1999;24:1290–4.
[2] Goolkasian P, Wheeler AH, Gretz SS. The Neck Pain and Disability Scale: test-retest reliability and construct validity. Clin J Pain 2002;18:245–50.
[3] Wlodyka-Demaille S, Poiraudeau S, Catanzariti JF, Rannou F, Fermanian J, Revel M. French translation and validation of 3 functional disability scales for neck pain. Arch Phys Med Rehabil 2002;83:376–82.
[4] Scherer M, Blozik E, Himmel W, Laptinskaya D, Kochen MM, Herrmann-Lingen C. Psychometric properties of a German version of the Neck Pain and Disability Scale. Eur Spine J 2008;17:922–9.
[5] Bremerich FH, Grob D, Dvorak J, Mannion AF. The Neck Pain and Disability Scale: cross-cultural adaptation into German and evaluation of its psychometric properties in chronic neck pain and C1-2 fusion patients. Spine (Phila Pa 1976) 2008;33:1018–27.
[6] Cook C, Richardson JK, Braga L, Menezes A, Soler X, Kume P, et al. Cross-cultural adaptation and validation of the Brazilian Portuguese version of the Neck Disability Index and Neck Pain and Disability Scale. Spine (Phila Pa 1976) 2006;31:1621–7.
[7] Bicer A, Yazici A, Camdeviren H, Erdogan C. Assessment of pain and disability in patients with chronic neck pain: reliability and construct validity of the Turkish version of the Neck Pain and Disability Scale. Disabil Rehabil 2004;26:959–62.
[8] van de Vijver FJR, Poortinga YH. Towards an integrated analysis of bias in cross-cultural assessment. Eur J Psych Asses 1997;13:21–9.
[9] Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 2000;25:3186–91.
[10] Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34–42.
[11] Schellingerhout JM, Heymans MW, Verhagen AP, de Vet HC, Koes BW, Terwee CB. Measurement properties of translated versions of neck-specific questionnaires: a systematic review. BMC Med Res Methodol 2011;11:87.
[12] Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;21:339:b2535.
[13] Goolkasian P. Neck Pain and Disability Scale: a critical evaluation. Expert Rev Pharmacoecon Outcomes Res 2003;3:379–82.

[14] Blozik E, Kochen MM, Herrmann-Lingen C, Scherer M. Development of a short version of the Neck Pain and Disability Scale. Eur J Pain (London, England) 2010;14:864:e861–7.
[15] En MC, Clair DA, Edmondston SJ. Validity of the Neck Disability Index and Neck Pain and Disability Scale for measuring disability associated with chronic, non-traumatic neck pain. Man Ther 2009;14:433–8.
[16] Wlodyka-Demaille S, Poiraudeau S, Fermanian J, Catanzariti JF, Rannou F, Revel M. French translation and validation of a functional disability scale for neck pain. Ann Readapt Med Phys 2001;44:132–42.
[17] Ono R, Otani K, Takegami M, Suzukamo Y, Goolkasian P, Wheeler AH, et al. Reliability, validity, and responsiveness of the Japanese version of the Neck Pain and Disability Scale. J Orthop Sci 2011;16:339–46.
[18] Salo P, Ylinen J, Kautiainen H, Arkela-Kautiainen M, Hakkinen A. Reliability and validity of the finnish version of the neck disability index and the modified neck pain and disability scale. Spine (Phila Pa 1976) 2010;35:552–6.
[19] Jorritsma W, de Vries GE, Dijkstra PU, Geertzen JH, Reneman MF. Neck Pain and Disability Scale and Neck Disability Index: validity of Dutch language versions. Eur Spine J 2012;21:93–100.
[20] Jorritsma W, de Vries GE, Geertzen JH, Dijkstra PU, Reneman MF. Neck Pain and Disability Scale and the Neck Disability Index: reproducibility of the Dutch Language Versions. Eur Spine J 2010;19:1695–701.
[21] Jorritsma W, Dijkstra PU, de Vries GE, Geertzen JH, Reneman MF. Detecting relevant changes and responsiveness of Neck Pain and Disability Scale and Neck Disability Index. Eur Spine J 2012;21:2550–7.
[22] Blozik E, Himmel W, Kochen MM, Herrmann-Lingen C, Scherer M. Sensitivity to change of the Neck Pain and Disability Scale. Eur Spine J 2011;20:882–9.
[23] Agarwal S, Allison GT, Agarwal A, Singer KP. Reliability and validity of the Hindi version of the Neck Pain and Disability Scale in cervical radiculopathy patients. Disabil Rehabil 2006;28:1405–11.
[24] Monticone M, Baiardi P, Nido N, Righini C, Tomba A, Giovanazzi E. Development of the Italian version of the Neck Pain and Disability Scale, NPDS-I: cross-cultural adaptation, reliability, and validity. Spine (Phila Pa 1976) 2008;33:E429–34.
[25] Monticone M, Ambrosini E, Vernon H, Brunati R, Rocca B, Foti C, et al. Responsiveness and minimal important changes for the Neck Disability Index and the Neck Pain Disability Scale in Italian subjects with chronic neck pain. Eur Spine J 2015;24:2821–7.
[26] Lee H, Nicholson LL, Adams RD, Maher CG, Halaki M, Bae SS. Development and psychometric testing of Korean language versions of 4 neck pain and disability questionnaires. Spine (Phila Pa 1976) 2006;31:1841–5.
[27] Lee JS, Suh KT, Kim JI, Lee HS, Goh TS. Validation of the korean version of the Neck Pain and Disability Scale. Asian Spine J 2013;7:178–83.
[28] Mousavi SJ, Parnianpour M, Montazeri A, Mehdian H, Karimi A, Abedi M, et al. Translation and validation study of the Iranian versions of the Neck Disability Index and the Neck Pain and Disability Scale. Spine (Phila Pa 1976) 2007;32:E825–31.
[29] Wu S, Ma C, Mai M, Li G. Translation and validation study of Chinese versions of the neck disability index and the Neck Pain and Disability Scale. Spine (Phila Pa 1976) 2010;35:1575–9.
[30] Chen Z, Zhao Y, Wang C, Li M, Zhu X. An adapted Chinese version of Neck Pain and Disability Scale: validity and reliability. Spine (Phila Pa 1976) 2011;36:E1322–7.
[31] Uthaikhup S, Paungmali A, Pirunsan U. Validation of Thai versions of the Neck Disability Index and Neck Pain and Disability Scale in patients with neck pain. Spine (Phila Pa 1976) 2011;36:E1415E1421.
[32] Kose G, Hepguler S, Atamaz F, Oder G. A comparison of four disability scales for Turkish patients with neck pain. J Rehabil Med 2007;39:358–62.
[33] Mokkink LB, Prinsen CA, Patrick DL, Alonso J, Bouter LM, De Vet HC, et al. COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs); user manual. Retrieved 2017 from http://www.cosmin.nl.
[34] Md FPM. Quality of life: the assesment, analysis and interpretation of patient-reported outcomes. 2nd ed. West Sussex, England: John Wiley & Sons Ltd, 2007.
[35] Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989;10:407–15.
[36] Jacobson N, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991;59:12–9.
[37] 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.
[38] Beninato M, Portney LG. Applying concepts of responsiveness to patient management in neurologic physical therapy. J Neurol Phys Ther 2011;35:75–81.
[39] Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. Phys Ther 2006;86:735–43.
[40] Guyatt GH, Deyo RA, Charlson M, Levine MN, Mitchell A. Responsiveness and validity in health status measurement: a clarification. J Clin Epidemiol 1989;42:403e8.