Spine

SPINE Volume 38, Number 10, pp E609–E615

HEALTH SERVICES RESEARCH

Cross-cultural Adaptation, Reliability, and Validity of the Arabic Version of Neck Disability Index in Patients With Neck Pain

Afaf Ahmed Mohamed Shaheen, PT, PhD,*‡ Mohammed Taher Ahmed Omar, PT, PhD,†‡ and Howard Vernon, DC, PhD§

Study Design. Translation and psychometric testing.

Objective. To adapt the neck disability index (NDI) cross-culturally to Arabic language and to investigate the reliability and validity of the Arabic version of NDI in an Arabic-speaking sample with neck pain.

Summary of Background Data. Although largely used, no previous reports exist on the translation process or the testing of the psychometric properties of the Arabic version of the NDI.

Methods Cross-cultural adaptation of an outcome questionnaire. The English version of the NDI was translated into Arabic (NDI-Ar) and back-translated according to established guidelines. Sixty-five patients with neck pain completed the NDI -Ar twice during a 1-week period, to assess its test-retest reliability. Further psychometric testing was done by assessing internal consistency, construct validity (factor structure), and responsiveness.

Results The internal consistency value (Cronbach α) for the NDI-Ar was 0.89. The test-retest reliability (intraclass correlation coefficient) was excellent at 0.96 (95% confidence interval from 0.93 to 0.97). There was a significant correlation ($r = 0.92$, $P < 0.05$) between the scores obtained from the first administration of the NDI-Ar and the second administration. Factor analysis demonstrated a 2-factor structure, explaining 67.58% of total variance. The analysis of responsiveness was calculated with an unpaired t test after 1 week of treatment and demonstrating a statically significant difference between stable and improved patients ($P < 0.05$). The Spearman correlation coefficient (rS = 0.81; $P = 0.000$) revealed strong relation between the change in score in the NDI-Ar and global rating of change. No ceiling or floor effects were detected in the NDI-Ar.

Conclusion. The Arabic version of the NDI has a 2-factor 10-item structure and is a reliable, valid, and responsive tool that can be used to assess neck pain in Arabic-speaking patients with neck pain. Therefore, it can be recommended for clinical and research purposes.

Key words: neck pain and disability, neck disability index, reliability, validity, responsiveness.

Spine 2013;38:E609–E615

Neck pain is a highly prevalent condition among the general population. Data from cross-sectional studies show that point estimates are approximately 10% to 35%.[1,2] Moreover, half of these patients had persistent pain and disability at the 5-year follow-up.[3] In the large majority of these cases, the pathological basis for the neck pain is unclear, and the complaints are labeled as "nonspecific" or "mechanical" neck pain.[4] The persistence of nonspecific neck pain may have a significant impact on patient's health status, activity of daily living, and work-related activities resulting in poor quality of life.[5,6]

Self-administered questionnaires are an integral part of patient evaluation. They are useful to quantify the level of disability, pain perception, and relevant outcomes.[7,8]

The neck disability index (NDI) is the most widely used and validated instrument to assess the impact of neck pain on the patients' functional activities and to measure outcomes in clinical practice and research.[5] The NDI questionnaire has been translated properly, culturally adapted, and validated to be used in different language and social environments.[9–24] This offers a standard measure for use in international studies and clinical practice and allows clinicians and researchers to share information and direct comparison across countries.[25,26]

As an Arabic version has not been developed with full cross-culture adaptation and psychometric properties, we chose to translate the NDI questionnaires into Arabic, as it is less costly and time consuming than generating a new measure, and to investigate its validity and reliability.

The purposes of this study were to translate, and cross-culturally adapt the NDI to Arabic language and to investigate

From the *Basic Sciences Department, Faculty of Physical Therapy; †Physical Therapy Department for Surgery, Faculty of Physical Therapy, Cairo University, Giza, Egypt; ‡Rehabilitation Research Chair, King Saud University, Riyadh, Kingdom of Saudi Arabia; and §Canadian Memorial Chiropractic College, Toronto, Canada.

Acknowledgment date: December 5, 2012. Revision date: January 15, 2013. Acceptance date: February 1, 2013.

The manuscript submitted does not contain information about medical device(s)/drug(s).

No funds were received in support of this work.

Relevant financial activities outside the submitted work: grants.

Address correspondence and reprint requests to Afaf Ahmed Mohamed Shaheen, PT, PhD, Department of Basic Sciences, Faculty of Physical Therapy, Cairo University, Egypt, 7 Ahmed Elziat St, Giza, Egypt; E-mail: afaf_pt@yahoo.com

DOI: 10.1097/BRS.0b013e31828b2d09

the reliability, factor structure and validity of the Arabic version of NDI (NDI-Ar) in Arabic-speaking patients with neck complaints. The ultimate goal was to develop an instrument in Arabic that would facilitate international research in musculoskeletal disorders as well as to serve health practitioners in their everyday clinical practice.

MATERIALS AND METHODS

Instruments

Neck Disability Index

The NDI is adapted from Oswestry Low Back Pain Disability Questionnaire.[5] It consists of 10 questions: pain intensity, personal care, lifting, reading, headaches, concentration, work, driving, sleeping, and recreation.[5] Each item is scored from 0 (no disability) to 5 (total disability). The maximum possible score is 50. However, the NDI is frequently normalized to 100 and reported as percent. Because the questionnaire is straightforward, the majority of the patients need approximately 5 minutes to complete it.[27]

Global Rating of Change

The 15-point global rating of change (GRC) scale was used for both patients and clinicians to assess overall perceptions of improvement since the initiation of treatment.[28] The GRC scale ranges from −7 (a very great deal worse) to 0 (about the same) to 7 (a very great deal better). Incremental descriptors of worsening and improving are assigned values from −1 to −7 and from 1 to 7, respectively. The following classifications have been proposed regarding the GRC score: 0, 1, or −1 signifies no change; ±2 to 3 signifies minimal change; ±4 to 5 signifies moderate change; and ±6 to 7 signifies a large change in a patient's condition. The GRC has been well validated and extensively used as an external reference standard to compare outcome measures.[29–32]

Translation and Cross-cultural Adaptation

The translation and cross-cultural adaptation of the NDI were based on the guidelines established by Beaton *et al*.[25] The entire process is made up of 6 steps (Figure 1).

Step I: The NDI was translated from English into Arabic by 2 independent translators whose language was Arabic. They provided a written report.

Step II: The 2 forward translations were compared and single-consensus Arabic NDI was then constructed by the expert committee.

Step III: A backward translation into English was undertaken by 2 independent native English speakers. They were blinded to the original version of the NDI. Their 2 versions of the back-translation were submitted to the committee.

Step IV: An expert committee reviewed all reports and agreed by consensus to a prefinal version of the Arabic NDI. This team of 10 experts was made up of methodologists, health professionals, language professionals, and translators. The original developer of the English version of the NDI (H.V.) also participated in an advisory capacity.

Step V: The prefinal version of NDI-Ar was tested on a sample of 20 patients for face validity. The patients gave their general impression on the clarity of the items, the relevance of the content to their situation, the comprehensiveness of the instructions and their ability to complete it on their own. The same questions were asked for every single item, and they were able to make suggestions whenever necessary. Finally, a

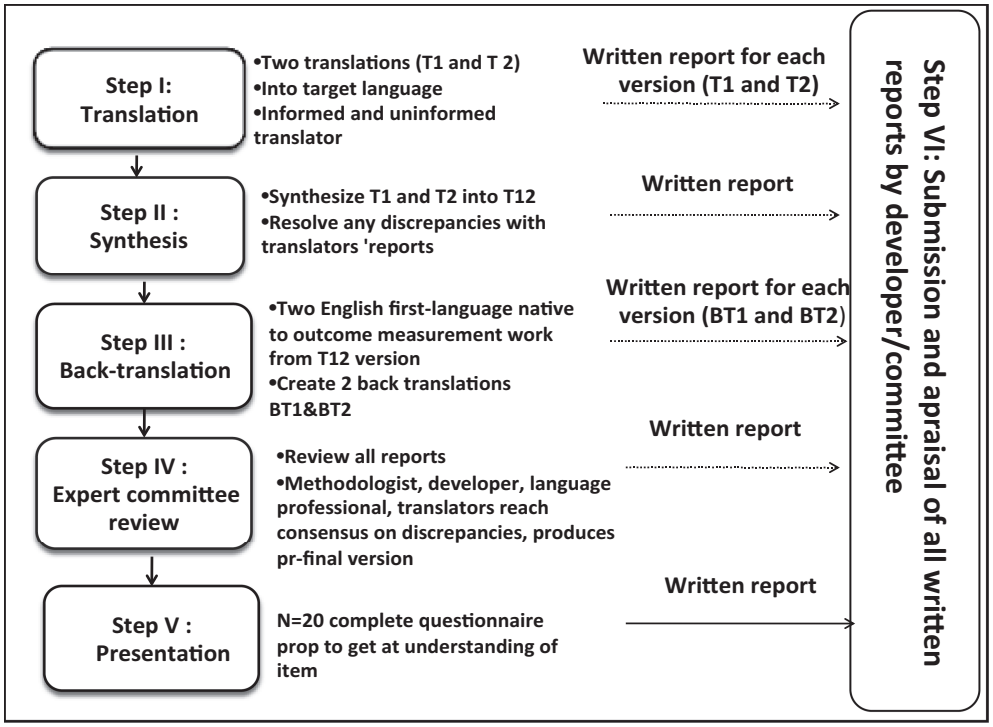

Figure 1. Steps of translation and cross-cultural adaptation.

debriefing summary, including all participant interviews, and a final debriefing decisions grid were sent to the developer for comments.

Step VI: Based on the results of prefinal version of NDI specifically, equivalence was reached between the source and target version in 4 areas: semantic equivalence (*i.e.*, ensuring that the words mean the same thing), idiomatic equivalence (*i.e.*, formulation of equivalent expressions for colloquialisms), experiential equivalence (ensuring that each item properly captured the experience of daily life in the target culture), and conceptual equivalence (ensuring that items hold the same conceptual meaning).[25,33] The committee made minor changes and completed the final Arabic version of NDI (see the Supplemental Digital Content Appendix I, available at http://links.lww.com/BRS/A749).

Testing the Psychometric Properties

To explore the psychometric properties of the NDI-Ar, the questionnaire was administered to 65 patients with neck pain. Patients eligible for the study were consecutively recruited from 3 primary heath care centers in Riyadh, Kingdom of Saudi Arabia from March 2011 to July 2012. Eligibility criteria were: age more than 18 years, a written consent of the patient, and neck pain lasting more than 3 months. Patients were excluded if they had severe cervical radiculopathy, neck pain related to vertebral fracture, neck surgery, cognitive impairment, infectious disease, neurological deficits, cancer, or other systemic diseases with possible effect on the musculoskeletal system. The study was approved by the Ethical Review Board at Department of Rehabilitation Heath Sciences, College of Applied Medical Science, King Saud University.

Eligible patients were informed about the purposes of the study and the confidentiality and anonymity of the process. After giving written consent, they completed a questionnaire on demographic and clinical characteristics of the NDI-Ar. Patients visited the physiotherapy clinics and/or orthopedic clinics 1 week later to complete the NDI-Ar with the changed item order. Patients also completed the GRC used as a criterion for "stable" conditions.

Data Analysis

Floor/Ceiling Effects

Floor/ceiling effects were considered to be present if more than 15% of respondents achieved the lowest or highest possible total score.[34,35]

Internal Consistency and Test-Retest Reliability

The internal consistency is considered to be good if the value of the Cronbach α is more than 0.70. Test-retest reliability was tested using a 2-way analysis of variance random-effect intraclass-correlation ($ICC_{[2,1]}$),[36] with good and excellent reliability being, respectively, indicated by values of 0.60 to 0.80 and more than 0.80. It was also assessed using the Pearson correlation coefficient and interpreted as excellent (0.81–1.0), very good (0.61–0.80), good (0.41–0.60), fair (0.21–0.40), and poor (0–0.20).[16] The sample size was estimated based on a method developed to calculate the required number of subjects in a reliability study.[17,37]

Construct Validity (Factor Analysis)

Factors structures were assessed using a Varimax rotation exploratory factor analysis to determine the number of extracted factors with eigenvalues greater than 1. Varimax rotation was applied, and the items with a factor loading of more than 0.40 were included in the factor; the expected explained variance was more than 0.50%.[27,38]

Responsiveness (Sensitivity to Change)

Responsiveness was analyzed in 2 ways. First, the patients were classified into 2 groups comprising those with stable (GRC <3 to > -3) and improved (GRC ≥ 3) scores after 1 week of treatment.[39] Next, the NDI change scores between the 2 groups were compared using an unpaired t test.[24] Patients with a deteriorated status were excluded from the analysis (n = 4). Responsiveness was also assessed by correlating the change score of the questionnaire to the GRC using the Spearman correlation coefficient.[17]

All the analyses were conducted using Statistical Package of Social Science (SPSS, Inc., Chicago, IL) version 18.0. The level of significant was set at $P < 0.05$.

RESULTS

Participants

Seventy patients with neck complaints visited the Health Centers. Five patients did not meet the eligibility criteria and were excluded from the study. All eligible subjects agreed to participate in the study and returned to complete the questionnaires for a second time (100% response rate). Descriptive statistics for missing patterns revealed 20 patients (30.76%) with one missing item. Demographic and clinical characteristics of patients and item-level descriptive statistics of NDI are presented in Tables 1 and 2, respectively.

Translation and Cross-cultural Adaptation

The developer comments on the translated tool concerned replacement of "pain" with "neck pain." This was applicable for items pain intensity, personal care, and lifting. The general impression of the patients was that the questionnaire and the instructions were easy to understand and that the items were important to their situation.

Psychometric Scale Properties

Floor/Ceiling Effects

The NDI-Ar had no floor/ceiling effects as less than 15% of patients achieved the minimum (1 patient, 1.5%) or maximum (2 patients, 3.1%) possible scores.

Internal Consistency and Test-Retest Reliability

Cronbach α of NDI-Ar was 0.89. Test-retest reliability was measured in all of the subjects and was excellent (ICC = 0.96;

TABLE 1. Demographic and Clinical Characteristics of the Participants

Variables		Mean	SD
Age (yr)		41.32	10.21
		Frequency	%
Sex	Male	45	69.2
	Female	20	30.8
Level of education	High	55	84.6
	Secondary	3	4.6
	Primary	7	10.8
Occupation	Office worker	30	46.2
	Manual	24	36.9
	Homemaker	5	7.7
	Heavy manual	6	9.2
Residence	Urban	61	93.8
	Rural	4	6.2
Duration of last episode	Acute (1–7 d)	4	6.2
	Subacute (7 d to 7 wk)	15	23.1
	Chronic (>7 wk)	46	70.8

SD indicates standard deviation.

95% confidence interval: (0.93–0.97). Test-retest reliability was also assessed with Pearson correlation ($r = 0.92$; $P < 0.05$). Table 3 shows the full results.

Construct Validity (Factor Analysis)

Factor analysis with varimax rotation revealed a 2-factor structure based on eigenvalues greater than 1. The first factor had an eigenvalue of 5.49 explaining 54.89% of the variance, whereas the second had an eigenvalue of 1.29 explaining an additional 12.95% of the variance. The factors were named activities of daily living (ADL, factor 1 items 2, 3, 7, 9, and 10) and pain and concentration (factor 2 items 1, 4, 5, 6, and 8). The results are shown in Table 4.

Responsiveness

The analysis of responsiveness was calculated with an unpaired *t* test comparing the NDI change scores between improved and stable patients after 1 week of treatment and was statistically significant ($P < 0.05$). The Spearman correlation coefficient, as calculated for stable and improved patients, was 0.81 ($P = 0.000$).

DISCUSSION

Arabic is the sixth most widely spoken language in the world. To our knowledge, this study is the first to translate, and cross-culturally adapt the NDI into Arabic language and test the psychometric properties of the NDI-Ar.

The results of the adaptation process indicate that it was successfully developed following international guidelines, and the difficulties encountered were overcome by careful wording and consensus decisions. The adapted questionnaire is self-administered and simple to use in clinical practice.

In our study, 20 patients (30.76%) did not answer item 8 (driving) in the NDI-Ar. This result was nearly similar with both Turkish and Polish version of the NDI, where 23.87% and 25% of the participants did not answer this section.[16,36] These missing values were higher when compared with those of Thai and French version of NDI, where 2.2% and 5% of the patients did not answer this item.[9,23] Moreover, these results were lower than the results of Greek and Japanese version of NDI where 44% and 38.2% of the patients did not answer the driving item of the NDI.[17,24] One explanation for the discrepancies may be that the driving is restricted to male population rather than the female in Saudi Arabia.

TABLE 2. Descriptive Statistics of NDI Items

	Test (Day 1)					Retest (Day 8)				
Items	Missing	Min	Max	Mean	SD	Missing	Min	Max	Mean	SD
Pain intensity	0.00	0.00	4.00	2.17	0.99	0.00	0.00	4.00	1.72	0.89
Personal care	0.00	0.00	3.00	0.63	0.69	0.00	0.00	3.00	0.68	0.69
Lifting	0.00	0.00	4.00	1.42	1.26	0.00	0.00	4.00	1.51	1.15
Reading	0.00	0.00	4.00	1.82	1.17	0.00	0.00	4.00	1.54	1.05
Headache	0.00	0.00	5.00	1.46	1.31	0.00	0.00	5.00	1.42	1.38
Concentration	0.00	0.00	4.00	1.01	1.23	0.00	0.00	4.00	0.94	1.14
Work	0.00	0.00	4.00	1.22	1.11	0.00	0.00	4.00	1.11	0.95
Driving	20	0.00	4.00	1.22	1.22	20	0.00	4.00	1.13	1.01
Sleeping	0.00	0.00	4.00	1.43	0.98	0.00	0.00	4.00	1.37	0.88
Recreation	0.00	0.00	5.00	1.66	1.32	0.00	0.00	5.00	1.52	1.28

Min indicates minimum; max, maximum; SD, standard deviation.

TABLE 3. Reliability of NDI-Ar by Internal Consistency and Test-Retest

	Internal Consistency (α)	Test-Retest	
		Correlation (*r*)	ICC (95% CI)
NDI-Ar	0.89	0.92	0. 96 (0.93–0.97)
NDI-Ar Factor 1	0.86	0.90	0.86 (0.79–0.52)
NDI-Ar Factor 2	0.77	0.92	0.77 (0.66–0.85)

NDI-Ar indicates the Arabic version of neck disability index; CI, confidence interval.

These authors concluded that this problem was not a translation issue. So, we did not think that it was necessary to make any changes. Missing values concerning the same item, driving, have been reported in several other studies.[9,12,13] In many cases where 1 item is missing, several authors have chosen to report the total score as a percentage of completed items.[40] This is also the instruction in the original article by Fairbank *et al*.[41]

The NDI-Ar has no serious floor/ceiling effects and better score distribution. Therefore, it can be reasonably concluded that NDI-Ar can assess the full range of severity related to neck pain. This result was in agreement with previous studies.[15,23,24,34,38] However, the result contradicted with the Finnish version of NDI, where 2 items in the Finnish version of NDI questionnaire did reach the floor value of 15%.[21]

The Cronbach α coefficient of NDI-Ar (0.89) indicated an excellent internal consistency, which is nearly equivalent to value obtained by Vernon and Mior (0.80)[5] in the original version of NDI. Our results are comparable with results from the previous studies (0.74–0.92).[11–15,17,20–24,35,36,42–44]

As for test-retest reliability, the ICC of the Arabic version was 0.96 at the 1-week interval. This result is comparable with original version of NDI (0.90), and adapted results (0.90–0.98) from previous reports.[5,9,15–17,20–21,24,44] Furthermore, this result is higher than the Brazilian (0.48), Thai (0.85), and Italian versions of NDI (0.84).[12,23,27] In addition there was a high significant correlation ($r = 0.92$, $P < 0.05$) between pre- and postscores of NDI-Ar. This result is in line with the Nigerian version of NDI estimates ($r = 0.96$).[45]

The variability among these reports may be attributed to the various intervals that have been used to determine test-retest reliability. Deyo *et al*[32] recommended 1- to 2-week intervals to avoid memory effects, whereas Dawson *et al*[46] recommended 2 to 3 days to ensure that changes in patient condition were minimal. In this study, we used a 1-week interval to avoid the memorizing effect and the ICC values have been similar to several studies (0.88–0.98).[16,17,21,24,27] According to Deyo *et al*[32] assessing reproducibility by retest at 1- to 2-week intervals (rather than a short interval) may result in more realistic estimate of the variability to be observed among control subjects in a longitudinal study. In summary, the internal consistency and test-retest results of this study indicated that the NDI-Ar is reliable, useable, and highly adapted instrument in the Arabic culture.

TABLE 4. Factor Analysis Loadings

Items	Factors	
	1	2
1. Pain	0.427	0.486*
2. Personal care	0.741*	0.342
3. Lifting	0.590*	0.330
4. Reading	0.183	0.829*
5. Headache	0.225	0.858*
6. Concentration	0.065	0.932*
7. Work	0.740*	0.558
8. Driving	0.468	0.813*
9. Sleeping	0.651*	0.148
10. Recreational activities	0.803*	0.182

**Factor loading of 0.4 or more.*

Factor analysis revealed a 2-factor structure: 1 subscale describing ADL and the other pain and concentration. These factors explained 67.58% of the total variance, and the percentage of the variance was comparable with the Japanese version (explained 61.8%)[24] and relatively higher than that was found in previous studies.[9,27,43,44]

The items of the NDI-Ar were classified in more or less in the same manner as the study results of other countries that had 2-factor solutions. Item 8 was classified under pain and concentration label similar to Japanese version, but this was because driving itself does not involve using the whole body.[24] Furthermore, the classification of item 6 under the label of pain and concentration is in agreement with the Italian, Catalan, and Korean versions of NDI[27,43,44] but in contrast to French and Japanese versions.[9,24]

In relation to factor analysis and factor structure, some controversies still exists. Hains *et al*[47] found a positive result for a 1-factor model, accounting for 59% of variance. As have other adaption,[12,17,20,21,23] with explained variance (41%–51.3%). This 1- or 2-factor solution has been seen in other region-specific questionnaires that contain questions about pain and disability[48,49] and may reflect the different conditions and samples.

Because the NDI is a condition-specific instrument, it is considered responsive to changes and thus appropriate for evaluative purposes. It is often used as an outcome measure in studies exploring the effectiveness of interventions; in patients with neck pain.[50] Unpaired *t* test revealed statistically significant differences between improved and stable patients. This result is similar to Japanese version.[24] As well as there was a significant correlation between NDI-Ar change scores and the GRC. This is in line with Greek version of NDI,[17] but it is in contrast with the study by Cleland *et al*,[51] where they evaluated the responsiveness of the instrument in patients with

cervical radiculopathy. Nevertheless, sensitivity to change of the questionnaire must also be explored in greater intervals because long-term outcomes are essential in estimating the effectiveness of interventions.

Study Limitations

There are some limitations to consider in this study. The relationships between neck disability and physical tests were not considered because only NDI questionnaires were used. Hence, further validation studies are suggested to investigate NDI-Ar psychometric properties (*e.g.*, the correlations with pain and quality of life) in subjects with other neck conditions. In addition further researches are required to investigate the convergent validity of NDI-Ar using health-related pain/disability and quality of life questionnaires.

Another limitation of our study is that global ratings do not represent a standard way of assessing changes in functional status. Therefore, definitions of clinically important changes could be inaccurate.

CONCLUSION

The Arabic version of the NDI has a 2-factor 10-item structure and is a reliable, valid, and responsive tool. NDI-Ar consists of simple and easily understood words therefore it may be used to assess neck pain in Arabic-speaking patients with neck pain in different Arabic countries. Therefore it can be recommended for clinical and research purposes.

➢ Key Points

- ❑ The NDI was translated and culturally adapted into Arabic following the proposed guidelines.
- ❑ Our version is the first to be published in Arabic countries and has proven to be reliable, valid, and sensitive to change.
- ❑ The NDI-Ar can be used to assess Arabic patients with neck pain

Supplemental digital content is available for this article. Direct URL citations appearing in the printed text are provided in the HTML and PDF version of this article on the journal's web site (www.spinejournal.com).

References

1. Andersson HI, Ejlertsson G, Leden I, et al. Chronic pain in a geographically defined general population: studies of differences in age, gender, social class and pain localization. *Clin J Pain* 1993;9:174–82.
2. Cote P, Cassidy JD, Carroll L. The Saskatchewan health and back pain survey: the prevalence of neck pain and related disability in Saskatchewan adults. *Spine* 1998;23:1689–98.
3. Enthoven P, Scargren E, Oberg B. Clinical course in patients seeking primary care for back or neck pain: a prospective 5-year follow-up of outcome and health care consumption with subgroup analysis. *Spine* 2004;29:2458–65.
4. Bogduk N. The neck. *Baillieres Clin Rheumatol* 1999;13:261–85.
5. Vernon H, Mior S. The neck disability index: a study of reliability and validity. *J Manipulative Physiol Ther* 1991;14:409–15.
6. Wheeler AH, Goolkasian P, Baird AC, et al. Development of the neck pain and disability scale. Item analysis, face, and criterion-related validity. *Spine* 1999;24:1290–4.
7. Rubinstein SM, Peerdeman SM, Van Tulder MW, et al. A systematic review of the risk factors for cervical artery dissection. *Stroke* 2005;36:1575–80.
8. Nordin M, Carragee EJ, Hogg-Johnson S, et al. Assessment of neck pain and its associated disorders: results of the bone and joint decade 2000–2010 task force on neck pain and its associated disorders. Best evidence on assessment and intervention for neck pain. *Spine* 2008;33:S101–22.
9. Wlodyka-Demaille S, Poiraudeau S, Catanzariti JF, et al. French translation and validation of 3 functional disability scales for neck pain. *Arch Phys Med Rehabil* 2002;83:376–86.
10. Ackelman B, Lindgren U. Validity and reliability of a modified version of the neck disability index. *J Rehabil Med* 2002;34:284–7.
11. Vos CJ, Verhagen AP, Koes BW. Reliability and responsiveness of the Dutch version of the Neck Disability Index in patients with acute neck pain in general practices. *Eur Spine J* 2006;15:1729–36.
12. Cook C, Richardson JK, Braga L, et al. Cross-cultural adaptation and validation of the Brazilian Portuguese version of the neck disability index and the neck pain and disability scale. *Spine (Phila Pa1976)* 2006;31:1621–7.
13. Lee H, Nicholson LL, Adams RD, et al. Development and psychometric testing of Korean language versions of 4 neck pain and disability questionnaires. *Spine (Phila Pa 1976)* 2006;31:1841–5.
14. Agarwal S, Allison GT, Agarwal A, et al. Reliability and validity of the Hindi version of the Neck Pain and Disability Scale in cervical radiculopathy patients. *Disabil Rehabil* 2006;28:1405–11.
15. Mousavi SJ, Parnianpour M, Montazeri A, 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.
16. Aslan E, Karaduman A, Yakut Y, et al. The cultural adaptation, reliability and validity of neck disability index in patients with neck pain. A Turkish version study. *Spine (Phila Pa 1976)* 2008;33:E362–5.
17. Trouli MN, Vernon HT, Kakavelakis KN, et al. Translation of the neck disability index and validation of the Greek version in a sample of neck pain patients. *BMC Musculoskelet Disord* 2008;9:106–13.
18. Bremerich FH, Grob D, Dvorak J, et al. 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* 2008;33:1018–27.
19. Scherer M, Blozik E, Himmel W, et al. Psychometric properties of a German version of the neck pain and disability scale. *Eur Spine J* 2008;17:922–9.
20. Andrade Ortega JA, Delgado Martínez AD, Almecija Ruiz R. Validation of the Spanish version of the neck disability index. *Spine (Phila Pa 1976)* 2010;35:E114–8.
21. Salo P, Ylinen J, Kautiainen H, et al. 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.
22. Wu S, Ma C, Mai M, et al. 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.
23. 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:E1415–21.
24. Nakamaru K, Vernon H, Aizawa J, et al. Cross-Cultural adaptation, reliability and validity of the Japanese version of the neck disability index. *Spine (Phila Pa 1976)* 2012;36:E1415–21.
25. Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000;25:3186–91.
26. Pietrobon R, Coeytaux RR, Carey TS, et al. Standard scales for measurement of functional outcome for cervical pain or dysfunction: a systematic review. *Spine* 2002;27:515–22.
27. Monticone M, Ferrante S, Vernon H, et al. Development of the Italian version of the Neck Disability Index. Cross-cultural adaptation, factor analysis, reliability, validity, and sensitivity to change. *Spine* 2012;37:1038–44.

28. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10:407–15.
29. Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. *Spine* 2005;30:1331–4.
30. Whitman JM, Flynn TW, Childs JD, et al. A comparison between two physical therapy treatment programs for patients with lumbar spinal stenosis: a randomized clinical trial. *Spine* 2006;31: 2541–9.
31. Cleland JA, Childs JD, Whitman J M. Psychometric properties of the Neck Disability Index and Numeric Pain Rating Scale in patients with mechanical neck pain. *Arch Phys Med Rehabil* 2008;89:69–74.
32. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures: statistics and strategies for evaluation. *Control Clin Trials* 1991:142–58.
33. 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;46:1417–32.
34. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.
35. Misterska E, Jankowski R, Glowacki M. Cross-cultural adaptation of the Neck Disability Index and Copenhagen Neck Functional Disability Scale for patients with neck pain due to degenerative and discopathic disorders. Psychometric properties of the Polish versions. *BMC Musculoskeletal Disord* 2011;12:84.
36. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull* 1979;86:420–8.
37. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med* 1998;17:101–10.
38. Bowling A. *Research Methods in Health: Investigating and Health Services*. 2nd ed. Philadelphia: University Press; 2002.
39. Young BA, Walker MJ, Strunce JB, et al. Responsiveness of the Neck Disability Index in patients with mechanical neck disorders. *Spine* 2009;9:802–8.
40. Vernon H. The Neck Disability Index: state-of-the-art, 1991–2008. *J Manipulative Physiol Ther* 2008;31:491–502.
41. Fairbank JC, Couper J, Davies JB, et al. The Oswestry Low Back Pain Disability questionnaire. *Physiotherapy* 1980;66:271–3.
42. McCarthy MJ, Grevitt MP, Silcocks P, et al. The reliability of the Vernon and Mior neck disability index, and its validity compared with the short form-36 health survey questionnaire. *Eur Spine J* 2007;16:2111–7.
43. Nieto R, Miró J, Huguet A. Disability in subacute whiplash patients: usefulness of the neck disability index. *Spine* 2007;33:E630–5.
44. Song KJ, Choi BW, Choi BR, et al. Cross-cultural adaptation and validation of the Korean version of the neck disability index. *Spine* 2010;35:E1045–9.
45. Odole AC, Adegoke BO, Akomas NC. Validity and test re-test reliability of the neck disability index in the Nigerian clinical setting. *Afr J Med Sci* 2011;40:135–8.
46. Dawson J, Fitzpatrick R, Carr A. Questionnaire on the perceptions of patients about shoulder surgery. *J Bone Joint Surg Br* 1996;78:593–600.
47. Hains F, Waalen J, Mior S. Psychometric properties of the neck disability index. *J Manipulative Physiol Ther* 1998;21:75–80.
48. Monticone M, Baiardi P, Ferrari S, et al. Development of the Italian version of the Oswestry Disability Index (ODI-I) : a cross-cultural adaptation, reliability, and validity study. *Spine* 2009;34: 2090–5.
49. Pekkanen L, Kautiainen H, Ylinen J, et al. Reliability and validity study of the Finish version 2.0 of the Oswestry Disability Index. *Spine* 2011;36:332–8.
50. White P, Lewith G, Prescott P, et al. Acupuncture versus placebo for the treatment of chronic mechanical neck pain: a randomized controlled trial. *Ann Intern Med* 2004;141:911–9.
51. Cleland JA, Fritz JM, Whitman JM, et al. The reliability and construct validity of the Neck Disability Index and patient specific functional scale in patients with cervical radiculopathy. *Spine (PhilaPa 1976)* 2006;31:598–602.