Scandinavian Journal of Pain 5 (2014) 28–33

Contents lists available at ScienceDirect

Scandinavian Journal of Pain

journal homepage: www.ScandinavianJournalPain.com

Clinical pain research

Reliability and responsiveness of the Norwegian version of the Neck Disability Index

Jan Borre Johansen[a], Cecilie Roe[a,b], Eva Bakke[c], Anne Marit Mengshoel[c], Nada Andelic[a,*]

[a] Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Norway
[b] Faculty of Medicine, University of Oslo, Norway
[c] Department of Health Sciences, Institute of Health and Society, University of Oslo, Norway

HIGHLIGHTS

- The Neck Disability Index (NDI)in Norwegian has good test-retest reliability.
- The Norwegian NDI reliably assesses changes in pain-related disability.
- Continued use of the Norwegian NDI is recommended.

ARTICLE INFO

Article history:
Received 23 May 2013
Received in revised form 8 October 2013
Accepted 17 October 2013

Keywords:
Neck pain
NDI
Validity
Test–retest reliability
Responsiveness

ABSTRACT

Background and aim: The Norwegian version of the Neck Disability Index (NDI) has been widely used in previous studies. To our knowledge, the test–retest reliability and responsiveness of the NDI have not been investigated. Thus, the aim of the present study was to investigate the test–retest reliability and responsiveness of the Norwegian version of the NDI in neck pain patients seen in a specialized outpatient clinic.

Methods: This study included patients referred to the neck and back outpatient clinic at Oslo University Hospital. A total of 255 patients were included in the study, of which 42 participated in the test–retest portion of the study. The intraclass correlation coefficient (ICC) was used to assess test–retest reliability. A total of 113 patients participated in the responsiveness analyses. Based on their responses on the Global Rating Scale of Change (GRS), patients were categorized into the following groups: worsened ($n = 24$), unchanged ($n = 7$) and improved ($n = 62$). The minimal detectable change (MDC) for the NDI was calculated. Responsiveness was assessed by constructing a Receiver Operating Characteristic curve (ROC curve) to distinguish patients who had improved or worsened from those who remained unchanged. The minimum clinically important difference (MCID) was estimated.

Results: The test–retest reliability between the baseline scores and the retest NDI scores was very good (ICC = 0.84; 95% CI 0.72–0.91). The ability of the NDI to discriminate between improved and unchanged patients (responsiveness) over time was acceptable based on the ROC curve analysis (AUC = 0.70; 95% CI 0.58–0.82). The estimated MDC for the Norwegian version of the NDI is 12.3%, and the MCID is 16.6%.

Conclusion: The Norwegian version of the NDI proved to be an instrument with good test–retest reliability and acceptable responsiveness for assessing neck pain-related disability among neck pain patients in a specialized outpatient clinic.

1. Introduction

An increasingly large number of adults experience musculoskeletal pain, which is responsible for nearly 50% of all sick leave due to neck and low back pain [1]. Neck pain is often characterized by a chronic course that leads to substantial functional limitations [2,3]. Several instruments have been developed to assess disability among patients suffering from neck pain [4]. Of these, the Neck Disability Index (NDI) [5] is the most extensively used instrument worldwide. Valid and reliable instruments are cornerstones of clinical research as they are necessary for the results to be accurately applied and interpreted [6–8]. The NDI has been validated in several international studies [9–13]. The preservation of the properties of translated measurements can be questioned when the translated versions have not been properly validated [14]. Indeed, the cultural context of an environment should be considered for all items

DOI of refers to article: http://dx.doi.org/10.1016/j.sjpain.2013.11.002.
* Corresponding author at: Department of Physical Medicine and Rehabilitation, Oslo University Hospital, PB 4950 Nydalen, 0424 Oslo, Norway. Tel.: +47 22 11 86 87.
E-mail address: nandelic@online.no (N. Andelic).

1877-8860/$ – see front matter
http://dx.doi.org/10.1016/j.sjpain.2013.10.001

of the instrument, e.g., driving a car, participating in work and leisure activities; therefore, adapting the measurements to meet national and cultural requirements is important [9,14]. Although the Norwegian version of NDI has been widely used [15–17], to our knowledge its reliability and responsiveness properties have not been investigated. Reliability can be defined as the degree to which the measurement is free from measurement error [6]. Responsiveness is defined as the ability of an instrument to detect change over time in the construct measured [6].

Internationally, the responsiveness of the NDI has proven to be fairly good at detecting change over time. However, there are relatively few studies that have calculated clinically important differences. A change of approximately 10% in the total score or a 5 or more point improvement in the total item score seems to be clinically relevant [18–20]. Although some studies have reported smaller differences when evaluating patients with acute neck pain in general practice [12], studies assessing patients with radiculopathy indicate larger differences [21]. These findings may reflect the fact that patient expectations of change could be context-related or that some conditions are more responsive to treatment than others [20].

With the assumption that patients with neck pain referred to a specialized clinic have tried various treatments without lasting effects, the question remains if the NDI is capable to capture changes in a presumed "treatment-resistant populations". Hence, the aim of the present study was to investigate the test–retest reliability and responsiveness of the Norwegian version of the NDI and to evaluate clinically important differences in neck pain among patients presenting to a specialized outpatient clinic.

2. Materials and methods

2.1. Study design and study population

This prospective study was conducted at the neck and back outpatient clinic of Oslo University Hospital between December 2007 and December 2009 with follow-up in April 2011. Patients are referred to the neck and back specialty clinic from primary care and other hospitals in the Southeast Health Region of Norway. The study is a part of the larger project on patients with neck pain. For additional details of the study population and research questions, see two recently published articles [16,22].

Briefly, the inclusion criteria were as follows: 16 years of age or older, the presence of neck pain and the ability to communicate in Norwegian. The exclusion criteria were as follows: presence of neck fractures, inflammatory rheumatic disease and systemic disease causing neck pain. The Regional Committee for Medical Research Ethics in Health Region East approved the study. We obtained written informed consent from all of the participants.

During the recruitment period, the questionnaire was sent to 600 referred patients. Thirty percent were unable to complete the questionnaire due to language barriers, and approximately 10% did not meet the inclusion criteria. Of the 360 eligible patients, 30% refused to participate. A total of 255 patients were included in the study. As shown in the inclusion flow chart (Fig. 1), 42 patients completed a second NDI within a specific time period of 3–14 days and were thus included in the test–retest reliability assessment. The short time interval was chosen in order to reduce major real changes in the patients functioning to influence the results. The aim was to test all patients after one week; but due to weekends, and various other factors the real time interval for retesting was 3–14 days.

Furthermore, 113 participants completed the NDI and the Global Rating Scale of Change (GRS) at the 2-year follow-up and were included in the responsiveness assessment.

2.2. Measures

Demographic factors, which were recorded at inclusion, included gender, age, marital status, the number of years of formal education (dichotomized into ≤12 years vs. >12 years) and the duration of the current sick leave episode (in weeks). To further describe patients' baseline clinical characteristics, we assessed pain intensity and emotional distress. We asked patients to rate their strongest pain intensity during the past week using a numeric rating scale (NRS) ranging from 1 ("no pain") to 10 ("the worst pain imaginable") as used in our hospital registry.

Emotional distress was assessed using the Norwegian version [23] of the Hopkins symptom checklist 25 (HSCL-25) [24]. The assessment is a 25-item questionnaire that includes items addressing depression (9 items), anxiety (9 items) and somatization (7 items). Each item is scored on a scale from 1 ("not at all") to 4 ("very much"). The sum of the scores is divided by the number of answered items. A general score equal to or greater than 1.75 is consistent with an increased level of emotional distress.

The NDI is a self-rated disability questionnaire developed for patients with neck pain [20]. It consists of 10 items: pain intensity, personal care, lifting, reading, headaches, concentration, work, driving, sleeping and leisure activities. In the present study, the Norwegian translation of the NDI was used (http://www.mapi-institute.com/questionnaires-and-translation/ourcatalog/133-specific-questionnaires-rheumatology). Each item is scored from 0–5, (no disability to total disability), with the maximum score being 50. We report the total scores as percentages, with the maximum score corresponding to 100%.

To determine patients' impression of improvement, we used the five-item Norwegian Global Rating Scale of Change (GRS), which includes the responses of "a lot worse, a bit worse, no change, somewhat better and much better" and is scored from −2 to +2.

3. Data analysis and statistics

Descriptive data are presented as proportions and means with standard deviations (SD). Independent sample *t*-tests were used to compare patient groups with respect to continuous variables. Paired samples *t*-tests were used for comparisons within the groups. Categorical variables were analyzed using a chi square test.

3.1. Reliability

Test–retest reliability assesses the extent to which scores for patients who have not changed remain the same for repeated measurements [6]. In this study, the test–retest reliability of the baseline scores and the second NDI scores ($n = 42$) were assessed using a 2-way random effects intraclass correlation coefficient (ICC). According to Rosner, ICC values from 0.40 to 0.75 indicate fair to good reproducibility, while an ICC value >0.75 shows excellent reproducibility [25].

In addition, Cronbach's alpha was used to estimate the internal consistency. The values >0.9 were regarded as excellent, 0.8–0.9 good and >0.7 acceptable [26].

3.2. Responsiveness

The anchor used in this analysis was the GRS scale, which was categorized into the following categories: worsened (response of a lot worse and a bit worse; $n = 24$), unchanged (response of no change; $n = 27$) and improved (response of somewhat better and much better; $n = 62$).

Similar to Pool et al. [27], we defined minimum detectable change (MDC) as the smallest difference in a score that can be detected, given the variation in changes on the NDI observed in

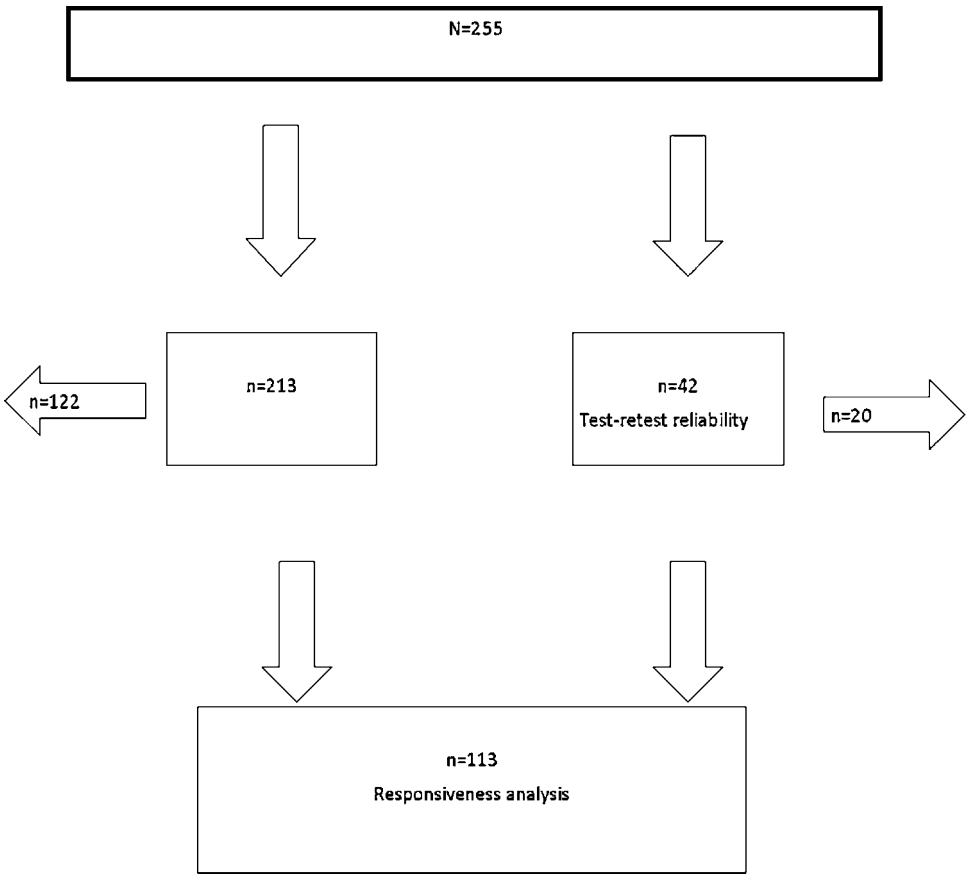

Fig. 1. Flow-chart of included patients.

patients who remained unchanged on the GRS ($n = 27$). The MDC was determined by first calculating the standard error of the mean (SEM) of the change in NDI scores. We used the formula SEM = SD $\sqrt{(1 - r)}$, where SD is the standard deviation and r is the reliability coefficient, and then determined the 95% CI for MDC using the equation MDC = $1.96 \times \sqrt{2} \times$ SEM [28,29]. Furthermore, we assessed responsiveness in patients who noted improvement on the GRS. Receiver Operating Characteristic curves (ROC curve) were plotted with the sensitivity values (true positive) on the *y*-axis against the 1-specificity values (false positive) on the *x*-axis for improved versus unchanged patient scores. The area under the curve (AUC) and 95% confidence intervals were calculated. The AUC value was considered acceptable if it was $\geq$0.70 (Hosmer and Lemeshow, 2000) [30]. A separate calculation was conducted for worsened versus unchanged patients scores. The minimum clinically important difference (MCID) was determined to be the magnitude of change associated with the uppermost left-hand corner of the curve, where both sensitivity and 1-specificity are maximized (Fig. 2) [28]. We performed the analyses using Predictive Analytics Soft Ware (PASW, version 18.0, SPSS, Chicago, IL, USA).

4. Results

The demographic and clinical characteristics of the 255 included patients (173 males and 82 females) are presented in Table 1. The majority of included patients suffered from chronic pain (lasting > 3 months). There were no statistically significant differences between the characteristics of the patients (e.g., demographics, percentages of patients on sick leave and with increased emotional distress, the pain intensity), who participated in the test–retest reliability analysis of the NDI ($n = 42$) and those individuals who did not participate ($n = 213$) (p-values ranges between 0.24 and 0.82, respectively). The patients included in the responsiveness analysis ($n = 113$) had slightly higher education levels ($p = 0.02$) and function as measured by the NDI, when compared with subjects who did not

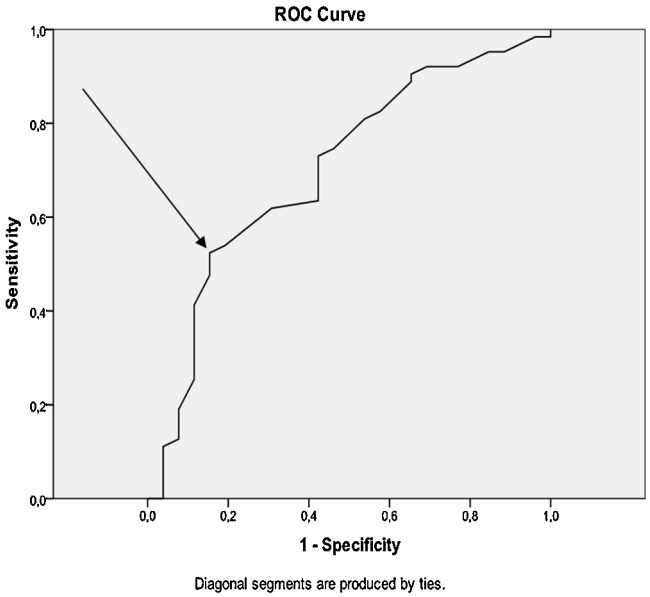

Fig. 2. Receiver operating characteristics curve on the Neck Disability Index (NDI) for those having improved according to the Global Rating Scale of Change (GRS). The arrow marks the data point nearest the uppermost left hand corner of the graph and represent the Minimum Clinically Important Difference (MCID) of the NDI.

Table 1
Demographic and clinical data for the included patients and the patients participating in the reliability and responsiveness sub-studies.

Variables	*N* = 255 All included patients	*N* = 42 Reliability study	*N* = 113 Responsiveness study
Male/female (*n*)	115/140	15/27	53/60
Age, mean (SD), years	45.81 (11.68)	46.33 (11.75)	47.27 (12.20)
Percent of patients with chronic pain	96%	95%	95%
Education level (percent <12 years)	64% Missing = 1	64%	58%*
Percent on sick leave	61%	38%	63%
HSCL-25 ≥1.75	49% Missing 5	37% Missing 1	43%
Maximal pain, mean (SD)	7.27 (2.08) Missing 24	6.95 (1.93)	7.16 (2.07)
NDI mean (SD)	41.72 (15.97)	36.48 (13.41)	39.89 (15.13)*

HSCL-25: Hopkins symptom checklist 25.
* Statistically significant difference between participants in responsiveness study and non-participants.

complete follow-up assessments ($n = 142$) (mean NDI score 39.89 (15.13) vs. 43.26 (16.56) respectively, $p = 0.02$).

4.1. Reliability

The mean NDI total score at baseline ($n = 42$) was 37.48 (SD 13.51), and the mean NDI total score after 3–14 days was 35.36 (SD 14.82). The mean difference between these 2 scores was 2.12%, which was not statistically significant.

The test–retest reliability between the baseline NDI total scores and the second NDI score was very good (ICC = 0.84; 95% CI 0.72–0.91). Table 2 shows the test–retest estimates of each of the 10 NDI items, with the ICC values in the range of good-to-excellent reproducibility (0.5–0.9).

The internal consistency of NDI assessed by Cronbach's alpha was good-to-excellent in both the baseline and second administration (0.83 and 0.91, respectively).

4.2. Responsiveness

The mean change in NDI scores noted at follow-up for the responsiveness study (improved vs. unchanged patients) was 10.4 (SD 15.4). The SEM was calculated to be 4.44, while the MDC value was 12.3% (or 6.15 out of 50).

The difference in the ability of the NDI to discriminate between improved and unchanged patients (responsiveness) over time was significant based on the ROC curve analysis (AUC = 0.70; 95% CI 0.58–0.82) (see Fig. 2). Similar results were found when computing ROC curves for those reporting to have worsened, with an AUC value of 0.72 (95% CI 0.58–0.87).

Table 2
The test–retest reliability between the baseline items scores and second items scores by using a 2-way random effects intraclass correlation coefficient (ICC).

NDI items	ICC	95% CI
1. Pain	0.48	0.27–0.72
2. Personal care	0.75	0.54–0.87
3. Lifting	0.51	0.09–0.74
4. Reading	0.89	0.78–0.94
5. Headaches	0.85	0.72–0.92
6. Concentration	0.76	0.56–0.87
7. Work	0.77	0.56–0.88
8. Driving	0.83	0.67–0.91
9. Sleeping	0.72	0.48–0.85
10. Leisure activities	0.79	0.61–0.89

In the present study, a change of 8.3 points (if referencing 50-point scale) or 16.6% was the estimated MCID for the Norwegian version of the NDI.

5. Discussion

Using conventional statistical methods, we found the NDI to have good test–retest reliability and responsiveness for assessing neck pain-related disability among patients with neck pain in a specialized neck and back outpatient clinic.

5.1. Reliability

The test–retest reliability of the total NDI scores and separate NDI items in this study may be regarded well, based on the ICC values. The values were within the range reported in other studies (0.50 and 0.98) [9,12,31]. Variance in different studies is expected and is dependent on the time intervals between the test–retest scores, the different anchors used to define the "stable" state, the different subgroup classifications used, the presence of acute versus chronic conditions and the origin of the pain (i.e., muscular versus neurologic) [9].

Previous studies of NDI have reported Cronbach's alpha values in the range of 0.74 and 0.93 [31–34]. Our findings are consistent with other international studies in this respect, as we noted a Crohnbach's alpha value of 0.83 and 0.91. In this context, it is important to note that it has been stated that a meaningful Cronbach's alpha value demands a unidimensional scale [8]. However, previous studies questioned the unidimensionality of the NDI when the Rasch analyses were applied to test its unidimensionality and scaling [11,16]. Nevertheless, in the present study, the alpha value was very high, and we believe that it would not be significantly altered in a way that changes the test–retest reliability. Of note, our study population consisted of patients with chronic pain whose pain intensity did not fluctuate significantly from day to day.

5.2. Responsiveness

Responsiveness has been evaluated using 2 methods, the Minimum Detectable Change (MDC) and Receiver Operating Characteristics curve (ROC), both of which integrate an anchor-based and distribution-based approach. The instrument can be considered responsive when the MDC exceeds the SEM, which was the case in our study.

Based on the current AUC value, the Norwegian version of the NDI seems to be a responsive instrument for detecting changes in clinical state. However, the figures are not quite convincing, as the AUC value did not exceed 0.70. The ability to detect worsening was somewhat improved, although the difference was small and only suggestive. This result is consistent with findings from other studies, which have found that the instrument in question is somewhat better in detecting worsening than improvement [35].

Responsiveness can also be reported using the Minimum Clinically Important Difference (MCID), which is the minimal amount of change in an instrument score associated with subjective improvement. As mentioned before, this value is related to the sensitivity and specificity of the instrument. Furthermore, it is also dependent on the anchor being used and on the patients to whom it is applied (e.g., acute and chronic pain, with and without referred pain, patients seen in general practice versus specialized clinics). In our study, the MCID (8.3 points or 16.6%) was out of the bounds of measurement error as indicated by the MDC (6.15 points or 12.3%). Our MCID value was similar to that of Young et al. [28] which studied improved versus stable patient groups with mechanical neck pain, but their MCID value was within the bounds of measurement error. However, the MCID was lower than that noted by Cleland

et al. [36] who reported an MCID value of 10 of 50 points and 19 of 50 points in two different studies of mechanical neck pain and radiculopathy, respectively. Stratford et al. [37] reported a MCID value of 5 of 50 points but they used an a priori assessment by the evaluating or treated therapist to assess meaningful change. Vos et al. [12] reported a value as low as 1.66 when including stable patients with acute neck pain from a general practice. Some of these differences may be explained by the different estimation methods and anchors used and the fact that different groups of patients may have different prognoses with respect to improvement. For example, patients with a higher baseline NDI score must experience a greater improvement to feel that it is relevant, which is reflected by a higher MCID value [20]. Furthermore, the MCID is an outcome measure that depends on the patient's evaluation, the clinicians' definition of important change values and the context in which it will be used. In addition, it should be interpreted in the context of empirical experience.

In the present study, chronic pain was defined as neck pain lasting 3 months or more. Thus, it is questionable to what extent these patients really can change or if the condition is an expression of a steady-state situation. However, three quarters of patients either improve or worsen, indicating that the use of the NDI as an outcome measure in these patients is reasonable.

In several other studies, the anchor-based method of evaluating the responsiveness of the NDI has been used [12,28,36–38]. However, there has been discussion as to whether it is appropriate to use the GRC as a comparison to the NDI and regarding the time delay for responses and the impossibility of recalling a subjective state condition at different times for comparison. In the current study, the patients had not been evaluated during or after a specified treatment session. Indeed, they had only been asked to evaluate their present neck pain status, and this assessment was compared to a created gold standard (i.e., the GRS), first presented by Jaeschke et al. [39]. The participants were presented with a set of questionnaires at the first and last distribution; thus, the ability to recall any exact answers on earlier occasions was diminished. We note that the patients completing the GRS had a significantly higher educational level and improved functioning as measured by the NDI.

The present results give further incentives to study the responsiveness of the NDI in different clinical settings, as well as with different patient groups and treatment options, e.g., studying patients in need of surgery versus patients seeking conservative treatment options. Furthermore, the NDI may be useful in other clinical settings as well due to the fact that it is a well-distributed instrument within the International Classification of Functioning (ICF) perspective [40] and the possibilities that this offers for the instrument as an outcome measurement to describe functioning and change in function, for example, with respect to work ability measurements.

Finally, it is worth mentioning the potential shortcomings of this study. In the test–retest reliability, the "practice effect" – respondents "learn" to answer the same questions in the first test and this affects their responses in the second test. Thus, the NDI scores may tend to be higher in the second test. However, the baseline NDI total score in this study was somewhat higher than the second scores. The other shortcoming is that only 44% of the patients participated in the responsiveness analysis. The external validity of the presented data may, therefore, be limited.

6. Conclusion

The Norwegian version of the NDI proved to be an instrument with good test–retest reliability and acceptable responsiveness for assessing neck pain-related disability among neck pain patients in a specialized outpatient clinic. Further studies should be conducted regarding the responsiveness of the NDI in different clinical settings and treatment interventions.

Conflict of interest

The authors report no conflicts of interest.

References

[1] Brage S, Ihlebaek C, Natvig B, Bruusgaard D. Musculoskeletal disorders as causes of sick leave and disability benefits. Tidsskr Nor Laegeforen 2010;130:2369–70.
[2] Borghouts JA, Koes BW, Bouter LM. The clinical course and prognostic factors of non-specific neck pain: a systematic review. Pain 1998;77:1–13.
[3] Hill J, Lewis M, Papageorgiou AC, Dziedzic K, Croft P. Predicting persistent neck pain: a 1-year follow-up of a population cohort. Spine (Phila Pa 1976) 2004;29:1648–54.
[4] Pickering PM, Osmotherly PG, Attia JR, McElduff P. An examination of outcome measures for pain and dysfunction in the cervical spine: a factor analysis. Spine (Phila Pa 1976) 2011;36:581–8.
[5] Vernon H, Mior S. The Neck Disability Index: a study of reliability and validity. J Manipulative Physiol Ther 1991;14:409–15.
[6] Mokkink LB, Terwee CB, Gibbons E, Stratford PW, Alonso J, Patrick DL. Inter-rater agreement and reliability of the COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments) checklist. BMC Med Res Methodol 2010;10:82.
[7] Angst F. The new COSMIN guidelines confront traditional concepts of responsiveness. BMC Med Res Methodol 2011;11:152.
[8] Terwee CB, Schellingerhout JM, Verhagen AP, Koes BW, de Vet HC. Methodological quality of studies on the measurement properties of neck pain and disability questionnaires: a systematic review. J Manipulative Physiol Ther 2011;34:261–72.
[9] MacDermid JC, Walton DM, Avery S, Blanchard A, Etruw E, McAlpine C. Measurement properties of the Neck Disability Index: a systematic review. J Orthop Sports Phys Ther 2009;39:400–17.
[10] Nieto R, Miro J, Huguet A. Disability in subacute whiplash patients: usefulness of the Neck Disability Index. Spine (Phila Pa 1976) 2008;33:E630–5.
[11] Van der Velde Beaton D, Hogg-Johnston S, Hurwitz E, Tennant A. Rasch analysis provides new insights into the measurement properties of the Neck Disability Index. Arthritis Rheum 2009;61:544–51.
[12] 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 practice. Eur Spine J 2006;15:1729–36.
[13] Pietrobon R, Coeytaux RR, Carey TS, Richardson WJ, DeVellis RF. Standard scales for measurement of functional outcome for cervical pain or dysfunction: a systematic review. Spine (Phila Pa 1976) 2002;27:515–22.
[14] 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.
[15] Kaale BR, Krakenes J, Albrektsen G, Wester K. Whiplash-associated disorders impairment rating: Neck Disability Index score according to severity of MRI findings of ligaments and membranes in the upper cervical spine. J Neurotrauma 2005;22:466–75.
[16] Johansen JB, Andelic N, Bakke E, Holter EB, Mengshoel AM, Roe C. Measurement properties of Norwegian version of the Neck Disability Index in chronic neck pain. Spine (Phila Pa 1976) 2012;38:851–6.
[17] Andelic N, Johansen JB, Bautz-Holter E, Mengshoel AM, Bakke E, Roe C. Linking self-determined functional problems of patients with neck pain to the International Classification of Functioning, Disability, and Health (ICF). Patient Prefer Adherence 2012;6:749–55.
[18] Riddle DL, Stratford PW. Use of generic versus region-specific functional status measures on patients with cervical spine disorders. Phys Ther 1998;78:951–63.
[19] Stratford PW, Gill C, Westaway M, Binkley J. Assessing disability and change on individual patients: a report of a patient specific measure. Physiother Can 1995:258–63.
[20] Vernon H. The Neck Disability Index: state-of-the-art, 1991–2008. J Manipulative Physiol Ther 2008;31:491–502.
[21] Cleland JA, Fritz JM, Whitman JM, Palmer JA. The reliability and construct validity of the Neck Disability Index and patient specific functional scale in patients with cervical radiculopathy. Spine (Phila Pa 1976) 2006;31:598–602.
[22] Johansen JB, Roe C, Bakke ES, Mengshoel AM, Storheim K, Andelic N. The determinants of function and disability in neck patients referred to a specialized outpatient clinic. Clin J Pain 2013;29:1029–35.
[23] Sandanger I, Moum T, Ingebrigtsen G, Dalgard OS, Sorensen T, Bruusgaard D. Concordance between symptom screening and diagnostic procedure: the Hopkins Symptom Checklist-25 and the Composite International Diagnostic Interview I. Soc Psychiatry Psychiatr Epidemiol 1998;33:345–54.
[24] Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L. The Hopkins Symptom Checklist (HSCL). A measure of primary symptom dimensions. Mod Probl Pharmacopsychiatry 1974;7:79–110.
[25] Rosner B. Fundamentals of biostatistics. Belmont, CA: Duxbury Press; 2005.
[26] Cortina JM. What is coefficient alpha? An examination of theory and applications. J Appl Psychol 1993;9:4–104.

[27] Pool JJ, Ostelo RW, Hoving JL, Bouter LM, de Vet HC. Minimal clinically important change of the Neck Disability Index and the Numerical Rating Scale for patients with neck pain. Spine (Phila Pa 1976) 2007;32:3047–51.
[28] Young BA, Walker MJ, Strunce JB, Boyles RE, Whitman JM, Childs JD. Responsiveness of the Neck Disability Index in patients with mechanical neck disorders. Spine J 2009;9:802–8.
[29] Stauffer ME, Taylor SD, Watson DJ, Peloso PM, Morrison A. Definition of nonresponse to analgesic treatment of arthritic pain: an analytical literature review of the smallest detectable difference, the minimal detectable change, and the minimal clinically important difference on the pain visual analog scale. Int J Inflam 2011;2011:231926.
[30] Hosmer DW, Lemeshow S. Applied Logistic Regression. New York: Wiley; 2000.
[31] 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.
[32] 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.
[33] Mousavi SJ, Parnianpour M, Montazeri A, Mehdian H, Karimi A, Abedi M. 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.
[34] McCarthy MJ, Grevitt MP, Silcocks P, Hobbs G. 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.
[35] Beaton DE. Understanding the relevance of measured change through studies of responsiveness. Spine (Phila Pa 1976) 2000;25:3192–9.
[36] Cleland JA, Childs JD, Whitman JM. 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.
[37] Stratford PW, Riddle DL, Binkley JM, Spadoni G, Westaway MD, Padfield B. Using the Neck Disability Index to make decisions concerning individual patients. Physiother Can 1999;51:107–12.
[38] Westaway MD, Stratford PW, Binkley JM. The patient-specific functional scale: validation of its use in persons with neck dysfunction. J Orthop Sports Phys Ther 1998;27:331–8.
[39] Jaeschke R, Singer J, Guyatt GH. A comparison of seven-point and visual analogue scales. Data from a randomized trial. Control Clin Trials 1990;11:43–51.
[40] Ferreira ML, Borges BM, Rezende IL, Carvalho LP, Soares LP, Dabes RA. Are neck pain scales and questionnaires compatible with the international classification of functioning, disability and health? A systematic review. Disabil Rehabil 2010;32:1539–46.