J Rehabil Med 2002; 34: 284–287

Taylor&Francis healthsciences

VALIDITY AND RELIABILITY OF A MODIFIED VERSION OF THE NECK DISABILITY INDEX

Birgitta Helmerson Ackelman and Urban Lindgren

From the Departments of Physiotherapy and Orthopedic Surgery, Karolinska Institutet, Huddinge University Hospital, Stockholm, Sweden

The Neck Disability Index was tested for validity and reliability. Fifty-nine Swedish patients (28 men, 31 women) were included. Twenty patients were in the acute phase after a neck sprain, 19 had chronic neck pain and 20 had no neck pain but had other musculoskeletal symptoms. On 5 occasions, the patients completed the Neck Disability Index, the Disability Rating Index, the MOS 36-item short-form health survey, 2 visual analogue scales, for pain and overall activity and some complementary questions. Levels of sensitivity, test-retest reliability and validity were acceptable. In order to increase specificity, we modified the Neck Disability Index by clarifying that the items only referred to the pain in the neck in 9 of 10 items. Thirty-eight patients (16 men, 22 women) were included in a study of the modified version. Twenty patients with acute neck sprain and 18 with other musculoskeletal symptoms filled out the modified version of the Neck Disability Index, which provided a more specific measure of disability due to neck pain.

Key words: neck pain, whiplash injuries, disability evaluation, outcome assessment.

J Rehabil Med 2002; 34: 284–287

Correspondence address: Birgitta Helmerson Ackelman, Department of Physiotherapy R 41, Huddinge University Hospital, SE-141 86, Stockholm, Sweden. E-mail: birgitta.ackelman@physio.hs.sll.se

Submitted June 13, 2001; accepted June 4, 2002

INTRODUCTION

Valid and reliable tests are cornerstones in clinical research. To evaluate a patient's functional status, several instruments are available, for example, the Sickness Impact Profile (SIP) (1) or the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) (2–4). The Neck Disability Index (NDI) described by Vernon & Mior (5) is based on the Oswestry Low Back Pain Disability Index (6) and specifically measures activity limitations due to neck pain. The NDI has been tested for face validity, test-retest reliability, internal consistency, construct validity and concurrent validity, but the authors suggested that larger group studies should be conducted to strengthen the overall relevance of the NDI (5). The Quebec Task Force also later suggested such studies (7). A translation of the NDI has been used at Huddinge University Hospital and at Linköping University Hospital in Sweden, but was never tested for validity.

The purpose of this study was to validate a Swedish version of the Neck Disability Index and to test a version that had been modified in order to improve its specificity.

METHODS

Subjects

The subjects studied were 2 groups of patients with neck pain and 1 group of patients without neck pain but with some other musculoskeletal disorder (Table I). One group of patients with neck pain was in the acute phase after a neck sprain and another group consisted of patients with chronic neck pain who had not received any form of therapy during the study period. A general inclusion criterion was the ability to read and speak Swedish. For the acute group, only patients with no previous neck disorder were included. For the chronic group, pain for 3 months or more was required.

The first part of the study consisted of 59 patients. Nineteen subjects in the acute phase after a neck sprain were referred from the emergency room on the day of the first visit, 20 subjects with chronic neck pain previously treated at the department of physiotherapy and 20 subjects with no neck pain but having other musculoskeletal disorders were recruited from the department of physiotherapy.

The second part of the study was designed to test a modified version of the NDI. In that part a different cohort of 38 subjects were included (20 with acute neck sprain and 18 with no neck pain).

Questionnaires

The NDI consists of 10 items; pain intensity, personal care, lifting, sleeping, driving, recreation, headaches, concentration, reading and work. The 10 items, with 6 possible answers in each are scored 0 (no activity limitations) to 5 (major activity limitations) and summed up to yield a total score. In the Swedish version used in the study the item concerning driving a car was provided with an additional alternative, "not applicable".

The SF-36 questionnaire (2–4), produces a profile of eight domain scores, including physical functioning (PF), physical role limitations, emotional role limitations, social functioning, bodily pain, general mental health, vitality and general health perception. Each domain is scored from 0 (poor health) to 100 (optimal health). The only domain studied by us was PF.

The Disability Rating Index (DRI) (8), consists of 12 items related to physical functioning: dressing, outdoor walks, climbing stairs, sitting for a longer time, standing bent over a sink, carrying a bag, making a bed, running, light work, heavy work, lifting heavy objects and participation in exercise programs or sports. On 100-mm visual analogue scales (VAS) with anchor points 0 (without difficulty) and 100 (unable to perform), the patient marks his/her presumed ability to perform the activity. The mean value of these measurements provides the DRI expressed as a percentage of the highest possible rating.

In addition to those instruments we distributed complementary questions concerning the adequacy and completeness of the NDI, items concerning neck disorders, cervical range of motion and the need for pain relievers. Two 100-mm visual analogue scales were also added: one concerning pain (anchor points "no" and "worst possible") and one concerning overall activity (anchor points "fully active" and "prevented all activities").

In the second part of the study we modified the NDI to clarify that the items referred specifically to neck conditions. Thus "neck pain" was used

 ISSN 1650–1977

Table I. *Patient groups in study*

Group	*n*	M/W	Age (years) Median (Range)	Duration of pain
Part 1				
Chronic neck pain	20	7/13	46.5 (25–63)	8 months–20 years
No neck pain	20	10/10	31 (18–70)	–
Acute neck pain	19	11/8	38 (21–62)	1–4 days
Part 2				
No neck pain	18	11/7	36.5 (19–57)	–
Acute neck pain	20	5/15	34 (22–58)	1–4 days

instead of "pain" or "due to neck pain" was added (not for the item concerning headache). This modification involved 9 of the 10 items.

Data collection

Data collection for the first part was carried out between 1997 and 1998. The studies were approved by the Ethical Committee at Huddinge University Hospital. In the first part of the study the subjects filled out the forms for the NDI, the SF-36, the DRI, the 2 VAS (pain and overall activity) and answered the complementary questions. Questionnaires were filled out on 5 occasions (Fig. 1); the first at the initial visit to the department of physiotherapy, the other 4 at home. However, all 5 questionnaires were completed at home for the previously treated chronic neck pain patients. All forms completed at home were returned individually in prepaid envelopes. In the second part of the study the subjects with acute neck sprain filled in the modified NDI and the DRI on 2 occasions; first on the day of their visit to the Emergency Room and second in the department of physiotherapy at least 2 hours later. The group with no neck pain having other musculoskeletal symptoms filled out the questionnaires once at the department of physiotherapy.

Data analysis

Since the NDI originally is scored from 0 (no activity limitations) to 50 (major activity limitations), and since the scoring scheme does not provide a strategy for dealing with questions that are unanswered, we transformed the scoring to a percentage score as described for the Oswestry Low Back Pain Index (6). If the alternative not applicable in the item concerning driving a car was chosen, that item was excluded.

If less than 8 of 10 items were answered the case was considered a dropout.

Statistics

The Spearman rank order correlation coefficient was used for the test-retest reliability of the NDI and for the correlations between different instruments. The Wilcoxon rank sum test was used for comparisons between instruments, groups and measurements within a group.

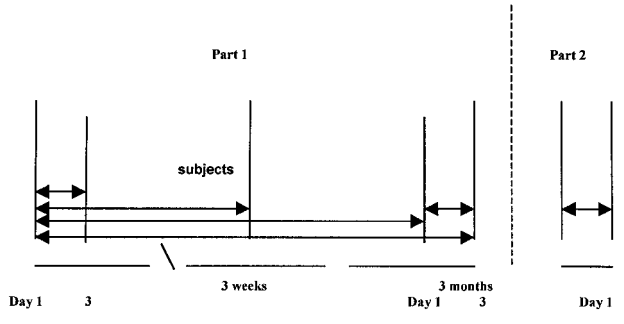

Fig. 1. Test-retest reliability conducted in the study.

RESULTS

In the first part of the study, the patients in the chronic group filled out 94% of the items satisfactorily; the percentage in the group without neck symptoms and in the acute group was 86% and 88%, respectively. In the second part of the study the items of the modified NDI were satisfactorily filled out by 92% of the group with no neck pain and by 88% of the group with acute neck pain.

First part

Face validity was estimated based on the subject's response during the initial evaluation as to whether the questionnaire was relevant to his/her disorder. Sixteen patients in the chronic neck pain group, 13 of the patients in the acute neck pain group and 2 in the no neck pain group felt that it was relevant or partially relevant (Table II).

The content validity was estimated based on the response to the question: "Is there something important you think should be added?". For results see Table III.

The concurrent validity was calculated as a rank correlation using the initial evaluation of NDI/DRI, NDI/PF, NDI/VAS pain and NDI/VAS activity. For results see Table IV.

The sensitivity was estimated based on the response to the statement in the initial complementary questionnaire, "I have a neck disorder". Twenty subjects in the chronic neck pain group, 19 subjects in the acute neck pain group answered "correct" or "partially correct". The NDI for the neck pain subjects was plotted in a diagram (Fig. 2).

The specificity was estimated for the initial evaluation using a diagram, in which the percentages of the NDI for the subjects with no neck pain were plotted (Fig. 3). Six subjects in the group with no neck pain had an NDI score over 10%, 4 of these over 20%. The items most frequently misunderstood were; "difficulty in lifting" (3 subjects), "working" (6 subjects), "sleeping" (5 subjects) due to pain and the item concerning pain (3 subjects). When analysing the individual answers, 2 subjects in the group without neck pain answered "partially correct" to the statement "I have a neck disorder". These 2 subjects did not mark any pain on the VAS.

The test-retest reliability coefficient for the subjects with chronic neck pain within 48 hours (initially/3 months), 3 weeks and 3 months were $r_s = 0.94–0.99$ (Table V). For the measurements within 48 hours for the subjects with acute neck pain $r_s = 0.81–0.89$ (Table V).

Second part

The modified NDI for the neck pain patients is presented in Fig. 4. The percentage scores of the modified NDI were 18–64% for the subjects with acute neck pain. The specificity of the modified NDI is presented in Fig. 5. The scores of subjects with no neck pain were all below 20% and there was a significant difference when comparing the modified NDI and the DRI ($p < 0.0001$). The test-retest correlation coefficient for the modified NDI was $r_s = 0.97$ ($n = 16$).

Table II. *Initial evaluation, face validity of the NDI. Answer to the question "Do you feel that the questions are relevant to your disorder?" Answers each group respectively*

Group	*n*	Yes, that's correct	Partly	No, not at all	No answer
Chr	20	9	7	0	4
Nl_1	20	1	1	13	5
Ac_1	19	11	2	1	5

Chr = chronic neck pain, Nl_1 = no neck pain, Ac_1 = acute neck pain.

DISCUSSION

When measuring disability or activity limitation one can choose between generic, condition specific and patient specific instruments. The SF-36 and the DRI belong to the generic group (8). Condition specific instruments like the NDI are thought to be more sensitive to changes in symptoms (9–11) or easier to use in clinical practice (12, 13). The study by Riddle & Stratford (13) describes similarities in results when using the SF-36 and the NDI (5). Westaway et al. (14) compared the Patient-Specific Functional Scale (PSFS) to the NDI. In the PSFS the patient chooses which 3 important personal activities are the most difficult to perform (14). The authors concluded that the PSFS is an excellent tool for working with individual patients and that it should be supplemented with a generic or condition-specific measure when group decision-making (i.e. quality assurance assessments or research) is the goal. They propose that when assessing the functional status of patients with cervical spine problems, either the physical component summary scale or the mental component summary scale of the SF-36 or the NDI can be used since there is considerable overlap between the two tests. The Swedish version of the NDI, which was used in the first part of our study, demonstrated good validity, sensitivity and test-retest reliability, but not optimal specificity. Therefore, in a modified version, we made it clear that the items specifically referred to neck pain. In the study by Vernon & Mior (5), test-retest reliability of the NDI was measured on a subset of 17 subjects with neck pain of different origin. The subjects completed the NDI within 48-hours. The correlation coefficient was $r_p = 0.89$. In the first part of our study we chose groups of neck pain patients who represented an acute or a chronic stage. The results showed a similar test-retest reliability within 48

Table III. *Part one. Content validity of the Neck disability index. All answers, collected at time-points 1–3 to the question "Is there something important that you think should be added?" are listed. Each statement refers to 1 subject*

Group Chronic neck pain
- "my life is almost destroyed because of pain"—"when one has pain even the family suffers"
- "I have numbness in my fingers and hands which I believe is connected to the neck disorder"
- "back problems"
- "not the pain itself but it's consequences-isolation , numbness, being trapped"
- "dizziness, memory"
- "the items are to distinct"
- "if numbness appears in different situations. Expand the item on lifting also to include carrying"
- "the alternatives are not entirely relevant"
- "which kind of pain, the degree of self confidence, dizziness when moving my head"
- "hard to find correct answer, a place to write my own comments"

Group Acute neck pain.
- "visual disturbances"
- "sore muscles when I bend forward and to the right"
- "locking sensation when opening my mouth"
- "depression …, no use …, a sense of not fitting in …, pain when lifting heavy objects from the floor"

Table IV. *Rank correlations on initial evaluation between Neck Disability Index (NDI) and Disability Rating Index (DRI) or Physical Functioning of the SF-36 (PF) or VAS pain or VAS activity*

Group	NDI/DRI	NDI/PF	NDI/VAS pain	NDI/VAS activity
Chr	0.95	−0.88	0.60	0.86
Nl_1	0.66	−0.17 n.s.	0.17 n.s.	0.35 n.s.
Ac_1	0.84	−0.86	0.91	0.82

Chr = chronic neck pain, Nl_1 = no neck pain, Ac_1 = acute neck pain. n.s. = non significant, all other correlations significant $p < 0.01$.

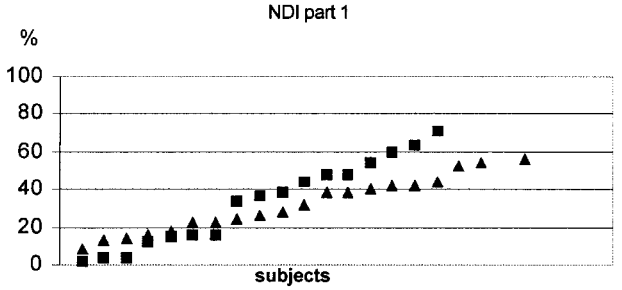

Fig. 2. Distribution of the Neck Disability Index (NDI) for the chronic neck pain group (squares $n = 20$) and the acute neck pain group (triangles $n = 17$) in part 1. Every dot represents 1 subject.

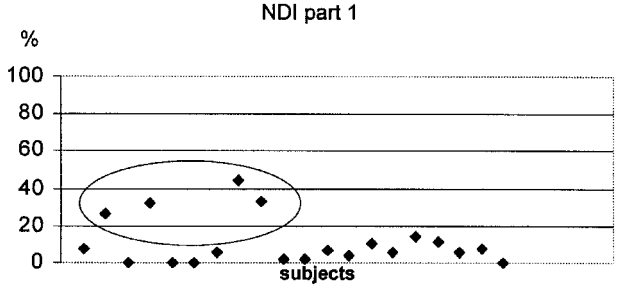

Fig. 3. Neck Disability Index (NDI) for the group with no neck pain ($n = 20$) in the first part. Every dot represents 1 subject. Four subjects reach an NDI over 20%.

Table V. *Neck disability index reliability. The Spearman correlation coefficient (r_s) between evaluation 1 and 2 (48 hours), 1 and 3 (3 weeks), 1 and 4 (3 months), 4 and 5 (48 hours after 3 months) for the chronic group. Correlation between evaluation 1 and 2 (48 hours) and 4 and 5 (48 hours after 3 months) for the acute neck pain patients*

	Measurement			
Group	1–2	1–3	1–4	4–5
Chr	0.97***	0.95***	0.94***	0.99***
	$n = 18$	$n = 18$	$n = 19$	$n = 18$
Ac_1	0.89***			0.81**
	$n = 16$			$n = 15$

Chr = chronic neck pain, Ac_1 = acute neck pain.
** $p < 0.01$. *** $p < 0.001$.

hours for the group with acute neck pain (correlation coefficient of $r_s = 0.81–0.89$). For the group with chronic neck pain we found $r_s = 0.97$ within 48 hours and within 3 months $r_s = 0.94$. The NDI for the neck pain subjects was well distributed and neither ceiling nor floor-effects could be seen. Because of the results from the first part of the study, we made the modified version. This version was clearly better, since non-specific ailments or co-morbidity did not produce false increases in the scores. Co-morbidity/other ailments may still be a problem when patients try to estimate the level of a particular disability. For example, the item dealing with headache is not necessarily a part of the neck pain syndrome. As for the concurrent validity, a high degree of pain for the chronic neck pain patients did not in most cases correspond with high levels of the NDI. This could indicate that the pain remains but the patient has learned to compensate for his/her disorder in different kinds of activities. Finally, in this study the suggestions for additional items indicated that the social consequences of the pain is an important part of the subjects' situation which is not covered when using the NDI.

In conclusion, it should be clearly stated for each item that only disability due to neck pain is of interest. This modification of the NDI is a valid and reliable instrument to measure disability due to neck pain.

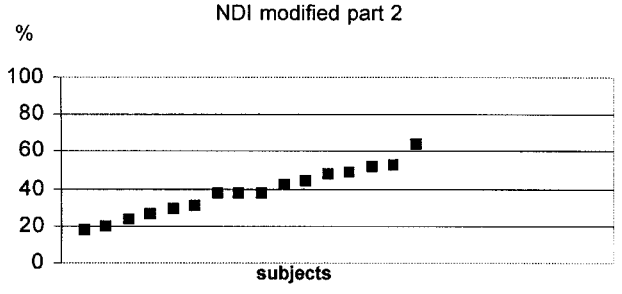

Fig. 4. Neck Disability Index (NDI) modified for the acute neck pain group in part 2, $n = 16$. Every dot represents 1 subject.

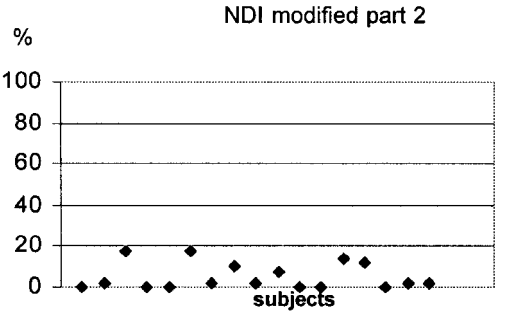

Fig. 5. Modified Neck Disability Index (NDI) for the group with no neck pain in part 2 ($n = 17$). Every dot represents 1 subject.

ACKNOWLEDGEMENT

This work was supported by grants from Karolinska Institutet, Board of Research for Health and Caring Sciences.

REFERENCES

1. Bergner M, Bobbitt R, Carter W, Gilson B. The sickness impact profile: development and final revision of a health status measure. Med Care 1981; 8: 787–805.
2. Sullivan M, Karlsson J, Ware J. The Swedish health survey—I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden. Soc Sci Med 1995; 10: 1349–1358.
3. Ware JE, Donald Sherbourne C. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992; 30: 473–483.
4. Ware JE, Kosinski M, Bayliss M, McHorney C, Rogers W, Raczek A. Comparison of methods for the scoring and statistical analysis of SF-36 health profile and summary measures: summary of results from the Medical Outcomes Study. Med Care 1995; 33: AS264–279.
5. Vernon H, Mior S. The neck disability index: a study of reliability and validity. J Manipulative Physiol Ther 1991; 14: 409–415.
6. Fairbank JT, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. Physiother 1980; 66: 271–273.
7. Quebec Task Force on WAD. Scientific monograph of the Quebec Task Force on Whiplash-Associated Disorders: Redefining 'whiplash' and its management. Spine 1995; 20: 1–273.
8. Salén BA, Spangfort EV, Nygren ÅL, Nordemar R. The disability rating index: an instrument for the assessment of disability in clinical settings. J Clin Epidemiol 1994; 47: 1423–1434.
9. Kopec JA, Esdaile JM, Abrahamowicz M, Abenhaim L, Wood-Dauphinee S, Lamping DL, Williams JI. The Quebec Back Pain Disability Scale. Measurement properties. Spine 1995; 20: 341–352.
10. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. Med Care 1989; 27: S217–232.
11. Stucki G, Liang M, Fossel A, Katz J. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. J Clin Epidemiol 1995; 48: 1369–1378.
12. Feinstein AR, Josephy BR, Wells CK. Scientific and clinical problems in indexes of functional disability. Ann Intern Med 1986; 105: 413–420.
13. Riddle DL, Stratford PW. Use of generic versus region-specific functional status measures on patients with cervical spine disorders. Phys Ther 1998; 78: 951–963.
14. 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: 331–338.