

Neck Pain and Disability Scale and Neck Disability Index: validity of Dutch language versions

Wim Jorritsma · Grietje E. de Vries ·
Pieter U. Dijkstra · Jan H. B. Geertzen ·
Michiel F. Reneman

Received: 28 September 2010/Revised: 25 January 2011/Accepted: 9 July 2011/Published online: 4 August 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract

Purpose To investigate the validity of the Neck Pain and Disability Scale Dutch Language Version (NPAD-DLV) and the Neck Disability Index (NDI)-DLV.

Methods NPAD-DLV, NDI-DLV, Short-Form-36 Health Survey (SF-36)-DLV, visual analog scale (VAS)_{pain} and VAS_{disability} were administered to 112 patients with non-specific chronic neck pain in an outpatient tertiary rehabilitation setting. Twenty seven hypotheses were formulated regarding validity. NPAD-DLV and NDI-DLV were evaluated for content validity (normal distribution total scores, missing items, floor and ceiling effects), internal consistency (Cronbach's alpha and Spearman Item-total correlations), construct validity (Pearson correlations with SF-36 domains, VAS_{pain} and VAS_{disability} and Pearson correlation between total scores of NPAD-DLV and NDI-DLV).

Results NPAD-DLV and NDI-DLV scores were distributed normally. Missing items were negligible. Floor and ceiling effects were absent in NPAD-DLV and in NDI-DLV two items had floor effects and one item had a ceiling effect. Cronbach's alpha of NPAD-DLV was 0.93 and of NDI-DLV 0.83. Item-total correlations ranged for NPAD-DLV from 0.45 to 0.73 and for NDI-DLV from 0.40 to 0.64. The correlation between, respectively, NPAD-DLV and NDI-DLV and: SF-36 domains ranged from -0.36 to -0.70 and from -0.34 to -0.63; VAS_{pain} was 0.54 and 0.43; VAS_{disability} was 0.57 and 0.52. The correlation between the total scores of NPAD-DLV and NDI-DLV was 0.77. Twenty six hypotheses were not rejected and one hypothesis was rejected.

Conclusion The NPAD-DLV and NDI-DLV are valid measures of self-reported neck-pain related disability.

Keywords Validation study · Assessment · Chronic pain · Short-Form-36 health survey

W. Jorritsma (✉) · P. U. Dijkstra · J. H. B. Geertzen ·
M. F. Reneman
Department of Rehabilitation Medicine,
Center for Rehabilitation,
University Medical Center Groningen,
University of Groningen,
PO Box 30002, 9750 RA Haren,
The Netherlands
e-mail: w.jorritsma@cvr.umcg.nl

G. E. de Vries
Department of Pulmonary Diseases,
University Medical Center Groningen,
University of Groningen, Groningen,
The Netherlands

P. U. Dijkstra
Department of Oral and Maxillofacial Surgery,
University Medical Center Groningen,
University of Groningen,
Groningen, The Netherlands

Introduction

Neck pain is a common musculoskeletal complaint in western societies [11]. In the majority of cases the pathological basis for neck pain is unclear and complaints are labeled as 'non-specific' or 'mechanical' [4]. Neck pain may result in disability, limitations in activities and restrictions in participation in daily living and work [32, 34]. Self-reported disability in patients with neck pain is often measured by means of region-specific and generic questionnaires [25]. Questionnaires should have good psychometric qualities, including validity [25, 27]. Three aspects of validity will be tested in this study. Content validity is the extent to which items of the questionnaire

reflect all aspects of the construct to be measured [25, 27]. Internal consistency is the extent to which all items measure the same construct [25, 27]. Construct validity is the extent to which a questionnaire is convergent and/or divergent correlated with other tests that are presumed to measure a similar or different construct [25, 27].

The most frequently used neck disability questionnaires are the Neck Pain and Disability Scale (NPAD) [34] and the Neck Disability Index (NDI) [32], which are validated in several languages [2, 3, 6, 8, 16, 17, 19, 21, 22, 29, 35]. The validity of the Dutch Language Versions (DLV) of the NPAD and NDI has not been studied. The aim of this study was to investigate the validity of the NPAD–DLV and the NDI–DLV in patients with non-specific chronic neck pain (CNP) in an outpatient tertiary rehabilitation setting. A priori hypotheses were defined (Text box 1) and outlined in “Materials and methods”.

Materials and Methods

Study sample

Patients with CNP were recruited from referrals by general practitioners or medical specialists for rehabilitation treatment in the Center for Rehabilitation at the

University Medical Center Groningen, The Netherlands. Inclusion criteria for this study were non-specific chronic neck pain (>3 months duration), admitted for outpatient rehabilitation, age between 18 and 65 years, and sufficient knowledge of the Dutch language (to complete questionnaires). Exclusion criteria were status post surgery in the cervical region, cardiovascular or pulmonary diseases severely diminishing physical capacity, pregnancy, addiction to drugs, and extensive psychological or behavioral problems.

Procedures

Prior to the first visit patients filled out a baseline questionnaire assessing clinical characteristics including visual analog scale (VAS)_{pain} and VAS_{disability}. During the first visit a review of the medical history and a physical examination was performed. A second visit was scheduled, depending on the length of the waiting list and patient availability, 2–9 weeks after the first visit, but prior to the start of the rehabilitation program. During the second visit the patients filled out the NPAD–DLV, the NDI–DLV and the Short-Form-36 Health Survey (SF-36). All patients signed informed consent for their data to be used for research purposes. Data were gathered between November 2006 and October 2009.

Text box 1. Hypothesis for examining validity of the NPAD–DLV and NDI–DLV

The validity is not rejected when:

Content validity

1. The total scores are normally distributed
2. The percentage of missing items is <5%
3. Floor and ceiling effects in item responses are not present
4. Total scores on the NDI of patients with CNP in a tertiary rehabilitation setting are significantly higher than patients with neck pain in a primary care setting

Internal consistency

5. The Cronbach's alphas are ≥ 0.7
6. The strength of the relationship of the single items with the total scale is fair to moderate ($0.25 \leq r < 0.75$)

Construct validity

7. The strength of the relationship with all eight SF-36 domains is fair to moderate ($0.25 \leq r < 0.75$)
8. The strength of the relationship with VAS_{pain} is fair to moderate ($0.25 \leq r < 0.75$)
9. The strength of the relationship between the NPAD and VAS_{pain} is higher than the strength of the relationship between the NDI and VAS_{pain}
10. The strength of the relationship with VAS_{disability} is moderate ($0.50 \leq r < 0.75$)
11. Differences in total scores between two age groups (below and above mean age of the study population) are not significant
12. Differences in total scores between males and females are not significant
13. Total scores of patients who are in litigation because of CNP are significantly higher than patients who are not in litigation
14. Total scores of patients who are receiving workers compensation because of CNP are significantly higher than patients who are not receiving workers compensation
15. The strength of the relationship between the NPAD and the NDI is moderate to good ($0.50 \leq r \leq 1.00$)

All hypotheses are operative for both questionnaires with exception of hypotheses 4, 9 and 15

Measurements

The NPAD consists of 20 items divided into 4 dimensions; neck problems; pain intensity; emotion and cognition; and interference with life activities [34]. Each item has a VAS of 100 mm with numeric anchors at 0, 1, 2, 3, 4 and 5 (each 20 mm apart). Item scores range from 0 (no pain or activity limitation) to 5 (as much pain as possible or maximal limitation). The total NPAD score ranges from 0 to 100 points. Higher scores indicate greater disability [34]. The NPAD has shown to be a valid and responsive measure of disability in other languages [3, 6, 8, 17, 19, 21, 22, 29, 34, 35]. The NPAD–DLV was used in this study; the reproducibility is acceptable [15].

The NDI consists of ten items: pain intensity, personal care, lifting, reading, headaches, concentration, work, driving, sleeping, and recreation [32]. Each item has six different assertions expressing progressive levels of pain or limitation in activities. Item scores range from 0 (no pain or limitation) to 5 (as much pain as possible or maximal limitation). The total NDI score ranges from 0 to 50 points. Higher scores indicate greater disability [32]. The NDI has shown to be a valid and responsive measure of disability in different languages [2, 8, 17, 19, 20, 22, 26, 32, 33, 35]. The NDI–DLV [16] was used in this study; the reproducibility [15, 33] and responsiveness are acceptable [26, 33].

The SF-36 is a questionnaire assessing general health of the past 4 weeks in 8 domains: physical functioning, physical role restriction, bodily pain, general health, vitality, social functioning, emotional role restriction, and mental health [12]. Scores for each domain range from 0 to 100, with higher scores indicating higher levels of functioning or well-being. The Dutch language version of the SF-36 has shown to be reliable and valid [1].

The VAS_{pain} is a horizontal line, 100 mm in length, anchored by word descriptors at each end (0: no pain, 100: worst pain possible). Patients are asked to draw a vertical mark across the horizontal line that best represents the pain level. The VAS_{pain} is a commonly used assessment instrument with proven reliability and validity [9].

The VAS_{disability} was evaluated by the question ‘how much does your neck pain restrict you in your daily activities?’ (ADL, housekeeping, work, hobby, recreation, sport and social activities). The scoring procedures are similar to the VAS_{pain}. The anchoring word descriptors are 0: no restriction and 100: worst possible restriction. The reliability and validity of the VAS_{disability} were assessed in patients with chronic musculoskeletal pain [5].

Hypotheses

Hypotheses are listed in Text box 1 and for the most part based on previous studies as described below.

Content validity

A normal distribution of the total scores of the NPAD–DLV and NDI–DLV was expected (Hypothesis 1), a good completeness of item responses (Hypothesis 2), and no floor and ceiling effects in item responses were expected (Hypothesis 3) [6, 7, 17, 19, 34, 35]. It was expected that scores on the NDI in a tertiary rehabilitation setting would be significantly higher than those in a Dutch primary care setting (Hypothesis 4) [6, 14, 19, 20, 26, 33, 34]. No Dutch data are available for comparison of the NPAD–DLV.

Internal consistency

It was expected that Cronbach’s alphas of the NPAD–DLV and NDI–DLV would be ≥ 0.70 (Hypothesis 5) and that Item–total score correlations would be fair to moderate (Hypothesis 6) [6, 8, 17, 19–22, 29, 32, 34, 35].

Construct validity

A fair to moderate correlation with all eight SF-36 domains was expected (Hypothesis 7) [8, 20–22]. It was expected that the NPAD–DLV and NDI–DLV had a fair to moderate correlation with VAS_{pain} [2, 3, 7, 13, 17, 20–22, 35] and a moderate correlation with VAS_{disability} [17, 35] (Hypotheses 8 and 10). Because four questions of the NPAD are pain-oriented a stronger correlation between the NPAD–DLV and the VAS_{pain} was expected than between the NDI–DLV and the VAS_{pain} (Hypothesis 9). No significant differences between sexes or age groups (below and above mean age of the study population) were expected (Hypotheses 11 and 12) [20, 32]. Significantly higher NPAD–DLV and NDI–DLV scores were expected for patients who were in litigation or who were receiving workers compensation because of their neck problems than for patients who were not in litigation or who received no workers compensation (Hypotheses 13 and 14) [18, 28]. A moderate-to-good correlation between the total scores of the NPAD–DLV and NDI–DLV was expected (Hypothesis 15) [2, 10, 22, 35]. All hypotheses are operative for both the NPAD–DLV and NDI–DLV with exception of hypotheses 4, 9 and 15; in total this results in 27 hypotheses.

Data analyses and criteria

Normality of the total scores was analyzed using the Kolmogorov–Smirnov test and PP plots. Floor and ceiling effects were considered to be present if more than 15% of respondents achieved the lowest or highest possible score for items [6]. When $\geq 75\%$ of the items did not have floor or ceiling effects, these questionnaires were considered to have no floor or ceiling effects. Internal consistency was

assessed with Cronbach's alphas and values ≥ 0.7 are considered adequate [24]. Standardized item- total score Spearman correlations of the NPAD-DLV and NDI-DLV were analyzed by calculating correlation coefficients between each item and the sum of all other items excluding the item investigated. Independent *t*-tests were used to analyze differences NPAD-DLV and NDI-DLV total scores between tertiary and primary care patients, patients younger or older than the mean age of the study population, men and women, patients with or without litigation, and with or without workers compensation. Pearson correlations were used to determine the strength of the relationship between the total scores of the NPAD-DLV and NDI-DLV and the SF-36 domain scores, VAS_{pain} and $VAS_{\text{disability}}$ and also between the total scores of NPAD-DLV and NDI-DLV. The construct validity was interpreted as good when at least 75% of the results corresponded with the hypotheses [30]. Correlations were interpreted as follows: $0.75 \leq r \leq 1.0$ as good, $0.50 \leq r < 0.75$ moderate, $0.25 \leq r < 0.50$ fair, and $0.00 \leq r < 0.25$ little or no [27]. All statistical analyses were performed with SPSS software, version 16.0. The critical values for significance were set at $p < 0.05$.

Results

A total of 391 patients with CNP were referred to the Center for Rehabilitation between November 2006 and October 2009 of which 129 were admitted for rehabilitation. A total of 125 patients fulfilled inclusion criteria. During the waiting period 13 patients decided not to start with the rehabilitation program because of lack of time, waiting period too long, problems with insurance company, and further diagnostic procedures. Clinical characteristics of the patients ($n = 112$) are presented in Table 1.

Content validity

NPAD-DLV and NDI-DLV were normally distributed. Therefore, hypothesis 1 was not rejected. Mean scores for individual items for the NPAD-DLV ranged from 1.7 to 4.2 (Table 2) and for the NDI-DLV from 0.7 to 2.8 (Table 3). In total 22 (1%) of 2,240 NPAD-DLV items and 15 (1%) of 1,120 NDI-DLV items were missing; therefore hypothesis 2 was not rejected (Tables 2, 3). Floor effects were $<10\%$ for all NPAD-DLV items. Ceiling effects were $<13\%$ for all NPAD-DLV items; therefore hypothesis 3 was not rejected (Table 2). For the NDI-DLV the items 'personal care' and 'sleeping' had floor effects, with respectively 44 and 19% of the patients scoring the lowest possible value. A ceiling effect was present for 'headaches' (19% of patients scored highest). Because 8 out of 10 NDI-

Table 1 Patient characteristics ($n = 112$)

	Mean (SD) <i>n</i> (%)
Age (years)	38.8 (11.4)
Duration of chronic pain (months)	18.0 (8.0–48.0) ^a
Sick leave in the past year (weeks)	15.6 (18.1)
NPAD-DLV (scale 0–100)	53.1 (16.6)
NDI-DLV (scale 0–50)	21.5 (7.4)
VAS_{pain} (0–100)	53.2 (21.4)
$VAS_{\text{disability}}$ (0–100)	54.0 (23.5)
Female	70 (63)
Pain radiating to,	
Shoulder(s)	94 (84)
Upper arm(s)	55 (49)
Forearm(s)	36 (32)
Hand/fingers	33 (30)
Between shoulder blades	54 (50)
Pins and needles below elbow	36 (34)
Concomitant complaints	
Headache	81 (73)
Dizziness	38 (34)
Concentration problems	20 (18)
Nausea	13 (12)
Fatigue	69 (62)
Low back pain	44 (40)
Self reported cause of neck pain	
Motor vehicle accident	47 (42)
Other trauma	16 (14)
Spontaneously/unknown	11 (10)
Stress	5 (5)
Work related	12 (11)
Other	21 (19)
Previously been treated for neck pain	102 (92)
Education	
Low	4 (4)
Intermediate vocational education	82 (75)
High	23 (21)
Work status (employed)	94 (84)
Workers compensation	62 (55)
Involved in litigation	34 (31)

NPAD-DLV Neck Pain and Disability Scale Dutch Language Version, NDI-DLV Neck Disability Index Dutch Language Version, VAS Visual Analog Scale

^a Median and Interquartile Range

DLV items did not have floor effects and 9 out of 10 did not have ceiling effects, hypothesis 3 was not rejected (Table 3). The total NDI-DLV score was 21.5 (Table 1). This score is significantly higher than the total scores in a Dutch primary care setting ($t(293) = 8.2$ (95% CI 5.3–8.7)

Table 2 Descriptive data and distribution of responses for each item in the NPAD-DLV ($n = 112$) and Spearman correlation between item scores and total score

Item	Present study Mean (SD)	% Individuals with lowest score	% Individuals with highest score	Numbers of missing	Item–total Correlation
1. Pain intensity	2.6 (1.0)	1	2	1	0.66
2. Average pain	2.9 (0.9)	0	1	1	0.58
3. Worst pain	4.2 (0.7)	0	13	1	0.45
4. Sleeping	2.5 (1.5)	9	3	1	0.60
5. Standing	2.1 (1.2)	2	1	1	0.66
6. Walking	2.0 (1.1)	5	1	1	0.68
7. Driving/riding	2.6 (1.3)	5	2	1	0.67
8. Social activities	2.9 (1.2)	3	4	1	0.73
9. Recreational activities	3.0 (1.1)	2	2	1	0.72
10. Working	3.4 (1.1)	1	7	1	0.57
11. Personal care	1.7 (1.3)	10	0	1	0.70
12. Personal relationships	2.1 (1.3)	6	0	1	0.70
13. Outlook on life and future	2.2 (1.6)	9	5	1	0.57
14. Emotions	2.5 (1.4)	4	1	1	0.61
15. Thinking/concentration	2.7 (1.5)	7	4	1	0.55
16. Neck stiffness	2.5 (1.2)	3	1	1	0.50
17. Turning neck	2.4 (1.3)	5	1	1	0.57
18. Looking up/down	2.3 (1.4)	6	2	2	0.52
19. Working overhead	3.5 (1.2)	1	8	1	0.50
20. Effect of pain pills	2.7 (1.4)	5	6	2	0.47

All correlations significant at the 0.01 level (2-tailed)

Table 3 Descriptive data and distribution of responses for each item in the NDI-DLV ($n = 112$) and Spearman correlations between item score and total score

Item	Present study mean (SD)	% Individuals with lowest score	% Individuals with highest score	Numbers of missing	Item–total Correlation
1. Pain	2.2 (0.8)	0	1	1	0.48
2. Personal Care	0.7 (0.8)	44	0	1	0.50
3. Lifting	2.3 (1.3)	5	2	1	0.50
4. Reading	2.4 (1.0)	6	0	1	0.40
5. Headaches	2.8 (1.6)	13	19	1	0.54
6. Concentration	2.0 (1.3)	12	2	2	0.53
7. Work	2.5 (1.2)	9	6	2	0.55
8. Driving	2.4 (1.2)	8	3	2	0.48
9. Sleeping	1.9 (1.3)	19	1	2	0.56
10. Recreation	2.4 (1.0)	3	1	2	0.64

All correlations significant at the 0.01 level (2-tailed)

[26] and $t(297) = 8.3$, (95% CI 5.3–8.7) [33]); therefore hypothesis 4 was not rejected.

Internal consistency

The Cronbach’s alphas of the NPAD-DLV and the NDI-DLV were respectively 0.93 and 0.83; therefore hypothesis 5 was not rejected. The strength of all Item–total correlations ranged from $r = 0.45$ to $r = 0.73$ (NPAD-DLV) and from $r = 0.40$ to $r = 0.64$ (NDI-DLV) (Tables 2, 3). Because all Item–total correlations fell within the hypothesized ranges, hypothesis 6 was not rejected.

Construct validity

Correlations between the total scores and SF-36, VAS_{pain} , and $VAS_{\text{disability}}$ are presented in Table 4. Differences between age groups, sexes, litigation status, and workers’ compensation are presented in Table 5. Hypotheses 7–13 were not rejected. Hypothesis 14 was rejected for the NPAD-DLV and not rejected for the NDI-DLV. The relationship between total scores of NPAD-DLV and NDI-DLV is presented in Fig. 1. The strength of the correlation between the NPAD-DLV and NDI-DLV was $r = 0.77$ (Table 4); therefore hypothesis 15 was not rejected.

Table 4 Construct validity of the NPAD–DLV and NDI–DLV (Pearson correlations)

	NPAD	95% CI	NDI	95% CI
NDI	0.77	0.68 to 0.84	–	
VAS _{pain}	0.54	0.39 to 0.66	0.43	0.27 to 0.57
VAS _{disability}	0.57	0.43 to 0.68	0.52	0.37 to 0.64
SF-36 Physical functioning	–0.58	–0.69 to –0.44	–0.49	–0.62 to –0.33
SF-36 Role physical	–0.36	–0.51 to –0.19	–0.38	–0.53 to –0.21
SF-36 Bodily pain	–0.70	–0.78 to –0.59	–0.63	–0.73 to –0.50
SF-36 General health	–0.44	–0.58 to –0.28	–0.47	–0.60 to –0.31
SF-36 Vitality	–0.50	–0.63 to –0.35	–0.51	–0.64 to –0.36
SF-36 Social functioning	–0.58	–0.69 to –0.44	–0.61	–0.71 to –0.48
SF-36 Role emotional	–0.39	–0.54 to –0.22	–0.37	–0.52 to –0.20
SF-36 Mental health	–0.45	–0.59 to –0.29	–0.34	–0.49 to –0.16

All correlations significant at the 0.01 level (2-tailed)

NPAD–DLV Neck Pain and Disability Scale Dutch Language Version, NDI–DLV Neck Disability Index Dutch Language Version, VAS Visual Analog Scale, SF-36 Short Form Health Survey

Table 5 Results of independent *t*-tests for the comparison of age ≤ 39 versus age >39 , male versus female, litigation versus no litigation, workers compensation (WC) versus no WC

	NPAD		<i>P</i> -value	NDI		<i>P</i> -value
	Mean (SD)	95% CI		Mean (SD)	95% CI	
Age						
≤39	51.1 (15.8)	–10.63 to 1.86	0.167	20.7 (6.3)	–4.62 to 1.11	0.227
>39	55.5 (17.3)			22.4 (8.5)		
Gender						
Male	54.3 (17.5)	–4.57 to 8.36	0.562	22.6 (8.2)	–1.12 to 4.65	0.228
Female	52.4 (16.1)			20.8 (6.9)		
Litigation ^a						
Yes	57.6 (16.9)	–13.68 to –3.5	0.020	25.4 (6.8)	–8.76 to –3.09	<0.001
No	50.5 (15.8)			19.5 (6.9)		
WC ^a						
Yes	55.3 (15.8)	–11.17 to 1.35	0.062	23.5 (6.9)	–7.21 to –1.82	<0.001
No	50.4 (17.2)			19.0 (7.4)		

Note: ^a1-tailed

Discussion

In this study the validity of the DLV of the NPAD and the NDI was tested with the use of pre-defined hypotheses. Because 26 of the 27 (96%) pre-defined hypotheses were not rejected the validity of the NPAD–DLV and NDI–DLV was interpreted as good. The current study was conducted in a university setting and is therefore representative for patients with CNP in a tertiary referral center. The sample size in our study was similar to those of other validity studies [6, 17, 35]. In the current study more women (63%) than men (37%) were included; this is similar to other validity studies [6, 10, 14, 17, 20–22, 29, 32, 34, 35], where the female proportions ranged from 54 to 83%. The mean age in our study was relatively young (39 years) in comparison with other validity studies, where the mean age of patients ranged from 38 to 65 years [6, 10, 14, 17, 20–22, 29, 32, 34, 35].

The normality of the total scores and the completeness of item responses were similar to other studies [6, 19, 21,

29, 35]. Floor and ceiling effects were not found in two studies [22, 35], while in two other studies floor effects for NPAD (6 items [19] and 14 items [6]) and NDI (3 items [19]) and ceiling effects for the NDI (1 item [19]) were found. The lower scores for most of the items for the NPAD [6, 19] and NDI [6] in those studies may explain the differences in floor effects with the present study. It is of interest that in the German study [6] ($n = 108$ of which $n = 80$ after atlantoaxial screw fixation and $n = 28$ with CNP) in the subgroup of patients with CNP much less items (3 in stead of 14) had floor effects. The Korean study [19] ($n = 180$) consisted of patients treated in physiotherapy departments of private hospitals or clinics.

We calculated a single Cronbach's alpha for the NPAD–DLV and NDI–DLV because their factor structure (1, 2, 3, or 4 factors for NPAD and 1 or 2 factors for NDI) is unclear and because in the original English versions single Cronbach's alphas for the total scales were calculated [6, 8, 21–23, 29, 31, 32, 35]. In the present study Cronbach's alpha for the NPAD–DLV was high (0.93). Other studies

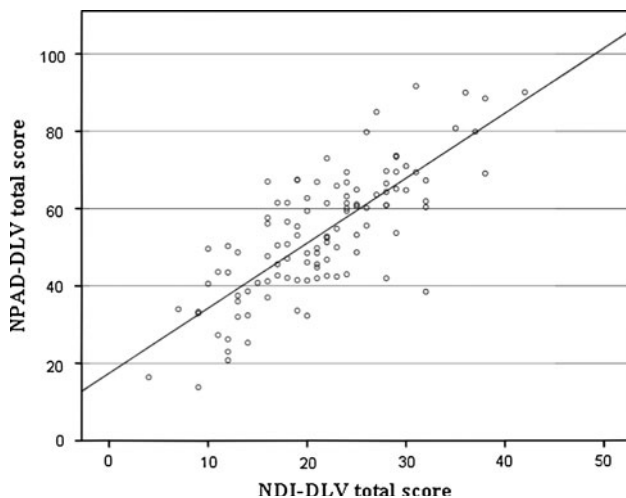


Fig. 1 Scatterplot showing total scores of NDI-DLV and NPAD-DLV

also found high values of Cronbach's alpha (range: 0.93–0.97) [6, 17, 19, 21, 22, 29, 34] indicating redundancy of items. Cronbach's alpha for the NDI-DLV in the present study (0.83) also falls within the range (0.74–0.92) reported by others [8, 17, 19, 20, 22, 32]. The variation in the Item-total score correlations for the NPAD-DLV and the NDI-DLV observed in the present study is similar with the variation found in other language versions (0.45–0.91 for the NPAD [6, 29, 34] and 0.45–0.84 for the NDI [20, 32]).

There is no established gold standard for assessment of neck pain disability. Therefore, criterion validity of the NPAD and NDI could not be analyzed [24]. To test the construct validity, comparisons were made with other constructs known to be associated with neck pain, neck pain related disability or generic health. The differences in the strength of the relationship between NPAD-DLV and NDI-DLV and all eight SF-36 domains with previous studies may be explained by differences such as patient setting, nature of neck condition, pain duration, and amount of neck pain related disability of the study samples [8, 20–22]. In the present study the correlation between the NPAD-DLV and VAS_{pain} was slightly higher than for the NDI-DLV and VAS_{pain} as hypothesized [17, 35]. The correlation of the NPAD-DLV and NDI-DLV with $VAS_{\text{disability}}$ in the present study was similar with that of other studies [17, 35]. The correlation between NPAD-DLV and NDI-DLV ($r = 0.77$) was similar with other studies (0.66–0.86), suggesting that these questionnaires measure comparable constructs [2, 10, 22, 35].

A potential limitation of this study was that the sample consisted largely of patients with moderate neck pain and disability. Although this may be expected in this tertiary rehabilitation setting, the validity of the NPAD-DLV and NDI-DLV should also be tested in general practice

populations. Furthermore, the period between the baseline assessment and the second assessment was variable and the stability of VAS_{pain} and $VAS_{\text{disability}}$ between first and second assessment was assumed but not formally assessed. All our patients with CNP started rehabilitation after completing the waiting period, indicating that their health status had not changed substantially [15]. Therefore, although we cannot be sure, this suggests that the potential impact of this weakness is unlikely to be substantial [15]. Finally, the hypotheses and the cut-off points that were used in the current study were based on previous studies without a methodically and qualitatively analysis of the validity of these studies.

A strength of this study is that to the author's knowledge for the first time a validity study is performed for the NPAD as well as the NDI in relation with SF-36 domain scores, VAS_{pain} and $VAS_{\text{disability}}$. Another strength is that the validity of the questionnaires is tested using explicit pre-defined hypotheses. The advantage of this method is its explicitness and transparency. Because the results are presented in detail, readers can develop and test their own hypotheses and perhaps interpret the same results differently. Further study with the NPAD-DLV is necessary to assess other measurement properties, such as responsiveness and minimally important change.

Conclusion

The NPAD-DLV and NDI-DLV are valid questionnaires to measure self-reported disability in patients with CNP within an outpatient tertiary rehabilitation setting.

Acknowledgments The authors thank Marleen Speller for assistance in data collection.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Aaronson NK, Muller M, Cohen PD et al (1998) Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. *J Clin Epidemiol* 51(11):1055–1068
2. Aslan E, Karaduman A, Yakut Y et al (2008) The cultural adaptation, reliability and validity of neck disability index in patients with neck pain: a Turkish version study. *Spine (Phila Pa 1976)* 33(11):E362–E365
3. Bicer A, Yazici A, Camdeviren H, Erdogan C (2004) 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* 26(16):959–962

4. Bogduk N (1999) The neck. *Baillieres Clin Rheumatol* 13(2):261–285
5. Boonstra AM, Schiphorst Preuper HR, Reneman MF, Posthumus JB, Stewart RE (2008) Reliability and validity of the visual analogue scale for disability in patients with chronic musculoskeletal pain. *Int J Rehabil Res* 31(2):165–169
6. Bremerich FH, Grob D, Dvorak J, Mannion AF (2008) 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* 33(9):1018–1027
7. Clair D, Edmondston S, Allison G (2004) Variability in pain intensity, physical and psychological function in non-acute, non-traumatic neck pain. *Physiother Res Int* 9(1):43–54
8. Cook C, Richardson JK, Braga L et al (2006) Cross-cultural adaptation and validation of the Brazilian Portuguese version of the neck disability index and neck pain and disability scale. *Spine* 31(14):1621–1627
9. Crossley KM, Bennell KL, Cowan SM, Green S (2004) Analysis of outcome measures for persons with patellofemoral pain: which are reliable and valid? *Arch Phys Med Rehabil* 85(5):815–822
10. En MC, Clair DA, Edmondston SJ (2009) Validity of the Neck Disability Index and Neck Pain and Disability Scale for measuring disability associated with chronic, non-traumatic neck pain. *Man Ther* 14(4):433–438
11. Ferrari R, Russell AS (2003) Regional musculoskeletal conditions: neck pain. *Best Pract Res Clin Rheumatol* 17(1):57–70
12. Garratt AM, Ruta DA, Abdalla MI, Buckingham JK, Russell IT (1993) The SF36 health survey questionnaire: an outcome measure suitable for routine use within the NHS? *BMJ* 306(6890):1440–1444
13. Gay RE, Madson TJ, Cieslak KR (2007) Comparison of the neck disability index and the neck bournemouth questionnaire in a sample of patients with chronic uncomplicated neck pain. *J Manipulative Physiol Ther* 30(4):259–262
14. Hoving JL, O’Leary EF, Niere KR, Green S, Buchbinder R (2003) Validity of the neck disability index, Northwick Park neck pain questionnaire, and problem elicitation technique for measuring disability associated with whiplash-associated disorders. *Pain* 102(3):273–281
15. Jorritsma W, de Vries GE, Geertzen JH, Dijkstra PU, Reneman MF (2010) Neck pain and disability scale and the neck disability index: reproducibility of the dutch language versions. *Eur Spine J* 19(10):1695–1701
16. Köke AJA, Heuts PHTG, Vlaeyen JWS et al (1996) Neck disability index. Pain knowledge center, university medical center maastricht, measurement instruments chronic pain. *Pijn Kennis Centrum, Academisch Ziekenhuis Maastricht, Meetinstrumenten chronische pijn, Maastricht*
17. Kose G, Hepguler S, Atamaz F, Oder G (2007) A comparison of four disability scales for Turkish patients with neck pain. *J Rehabil Med* 39(5):358–362
18. Landers MR, Cheung W, Miller D et al (2007) Workers’ compensation and litigation status influence the functional outcome of patients with neck pain. *Clin J Pain* 23(8):676–682
19. Lee H, Nicholson LL, Adams RD et al (2006) Development and psychometric testing of Korean language versions of 4 neck pain and disability questionnaires. *Spine* 31(16):1841–1845
20. McCarthy MJ, Grevitt MP, Silcocks P, Hobbs G (2007) 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* 16(12):2111–2117
21. Monticone M, Baiardi P, Nido N et al (2008) Development of the Italian version of the neck pain and disability scale, NPDS-I: cross-cultural adaptation, reliability, and validity. *Spine* 33(13):E429–E434
22. Mousavi SJ, Parnianpour M, Montazeri A et al (2007) Translation and validation study of the Iranian versions of the neck disability index and the neck pain and disability scale. *Spine* 32(26):E825–E831
23. Nieto R, Miro J, Huguet A (2008) Disability in subacute whiplash patients: usefulness of the neck disability index. *Spine (Phila Pa 1976)* 33(18):E630–E635
24. Offringa M, Assendelft WJJ, Scholten RJPM (2008) Introduction to evidence-based medicine (Inleiding in evidence-based medicine), Houten
25. Pietrobon R, Coeytaux RR, Carey TS, Richardson WJ, DeVellis RF (2002) Standard scales for measurement of functional outcome for cervical pain or dysfunction: a systematic review. *Spine* 27(5):515–522
26. Pool JJ, Ostelo RW, Hoving JL, Bouter LM, de Vet HC (2007) Minimal clinically important change of the neck disability index and the numerical rating scale for patients with neck pain. *Spine (Phila Pa 1976)* 32(26):3047–3051
27. Portney LG, Watkins MP (2000) *Foundations of Clinical Research*, 2nd Ed. Prentice Hall Health., Upper Saddle River, New Jersey
28. Riddle DL, Stratford PW (1998) Use of generic versus region-specific functional status measures on patients with cervical spine disorders. *Phys Ther* 78(9):951–963
29. Scherer M, Blozik E, Himmel W et al (2008) Psychometric properties of a German version of the neck pain and disability scale. *Eur Spine J* 17(7):922–929
30. Terwee CB, Bot SD, de Boer MR et al (2007) Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 60(1):34–42
31. Trouli MN, Vernon HT, Kakavelakis KN et al (2008) Translation of the neck disability index and validation of the Greek version in a sample of neck pain patients. *BMC Musculoskelet Disord* 9:106
32. Vernon H, Mior S (1991) The neck disability index: a study of reliability and validity. *J Manipulative Physiol Ther* 14(7):409–415
33. Vos CJ, Verhagen AP, Koes BW (2006) Reliability and responsiveness of the Dutch version of the neck disability index in patients with acute neck pain in general practice. *Eur Spine J* 15(11):1729–1736
34. Wheeler AH, Goolkasian P, Baird AC, Darden BV (1999) Development of the neck pain and disability scale. Item analysis, face, and criterion-related validity. *Spine* 24(13):1290–1294
35. Wlodyka-Demaille S, Poiraudou S, Catanzariti JF et al (2002) French translation and validation of 3 functional disability scales for neck pain. *Arch Phys Med Rehabil* 83(3):376–382