Development, validity and reliability of the Italian version of the Copenhagen neck functional disability scale

Domenico Angilecchia[1]*, Maura Mezzetti[2], Alessandro Chiarotto[3,4], Antonella Daugenti[5], Giuseppe Giovannico[6] and Francesca Bonetti[7]

Abstract

Background: Valid and reliable patient-reported outcome measures support health professionals in evaluating the results of clinical research and practice. The Copenhagen Neck Functional Disability Scale (CNFDS) has shown promising measurement properties to measure disability in patients with neck pain, but an Italian version of this questionnaire is not available. The objective of this study was to cross-culturally adapt the CNFDS into Italian (CNFDS-I), and to assess its validity and reliability in patients with neck pain.

Methods: The CNFDS-I was developed according to well-established guidelines for cross-cultural adaptation of patient-reported outcome measures. A cross-sectional clinimetric study was conducted to evaluate its validity and reliability. Patients with chronic neck pain (pain > 3 months) participated in this study. The following measurement properties (defined by the COSMIN initiative) were assessed: structural validity (exploratory factor analysis), internal consistency (Cronbach's α), construct validity [by testing hypotheses on expected correlations with the Neck Disability Index (NDI), the Neck Bournemouth Questionnaire (NBQ), and pain Visual Analogue Scale (VAS)]. Test-retest reliability [Intraclass Correlation Coefficient for agreement (ICCagreement)], and measurement error [Smallest Detectable Change (SDC)] were also assessed in 50 clinically stable patients. Floor/ceiling effects and acceptability were calculated.

Results: One-hundred and sixty-two patients (mean age = 47.9 ± 14.5 years, 70% female) were included. The CNFDS-I exhibited sufficient unidimensionality (one factor explained 83% of the variability) and internal consistency (α = 0.83). Construct validity was sufficient as all correlations with the other questionnaires were as expected ($r = 0.846$ with NDI, $r = 0.708$ with NBQ, $r = 0.570$ with VAS). Test-retest reliability was excellent (ICCagreement = 0.99, 95% CI from 0.995 to 0.999), while measurement error was equal to 8.31 scale points (27% scale range). No floor/ceiling effects were detected. The average time for filling the questionnaire was two minutes.

Conclusions: The CNFDS-I proved to be a valid and reliable outcome measure to assess disability in patients with chronic neck pain. Head-to-head comparison studies on the CNFDS-I measurement properties against other disability measures for neck pain (e.g. NDI and NBQ) are required to determine the relative merits of these different measures.

Keywords: Chronic neck pain, Copenhagen neck functional disability scale, Outcome measure, Validity, Reliability, Italian

* Correspondence: angilecchia@gmail.com
[1]Department of Physical Medicine and Rehabilitation OSMAIRM, via Cappuccini 9, 74014 Laterza, Taranto, Italy
Full list of author information is available at the end of the article

Background

Neck pain (NP) is a very prevalent musculoskeletal disorder worldwide; almost half of the population will experience a NP episode during the lifetime [1]. Most acute episodes of NP resolve with or without treatment, but almost 50% of people will continue to experience a certain degree of pain [2]. When symptoms associated to NP persists over three months, it is defined chronic NP [3]. A Neck Pain Task Force [4] has highlighted that about 10–20% of the European population displays chronic or persistent NP; it has also been shown that NP is the second cause of absence from work [5].

Neck pain is a multifactorial condition which can be related to sex, age, bad posture, poor state of health, other comorbidities, repetitive strain injuries, psychological factors, sleep disorder, and lifestyle [5, 6]. Therefore, patients with this condition present with a complex dysfunctional framework that clinicians should regularly manage the best possible manner. Within this context, it is useful to measure neck-related disability as perceived and described by the patient over time, irrespective of its etiology; this measurement can allow to monitor the patients' outcomes, and to study the relationship between this outcome and other health-related and environmental factors [7].

The Copenhagen Neck Functional Disability Scale (CNFDS) was developed by Jordan et al., with the aim to make a new instrument to investigate the patient's neck-related disability [8]; in contrast with other questionnaires (e.g. Neck Disability Index), it was not derived from already existing questionnaires [9, 10]. The functional disabilities experienced by patients with NP were used as a starting point by an interdisciplinary team of physiotherapists, rheumatologists and other professionals who aimed to develop this questionnaire as a self-reported tool, to avoid interviewer bias. The developers decided also not to include pain questions in the CNFDS, because these were considered to measure separate domains [11, 12]; their inclusion could also lead to problems with the unidimensionality of this questionnaire [8].The CNFDS has been shown to be strongly reliable and internally consistent, and to have excellent construct validity [8], similar to other instruments used to evaluate neck-related disability [13].

Considering that an Italian version of the CNFDS is not available for research and clinical purposes, this study aimed to cross-culturally adapt the CNFDS into the Italian language and culture, and to assess its measurement properties in patients with chronic NP. Making the CNFDS available in Italian will also allow to conduct head-to-head comparisons with other available neck-related disability questionnaires (e.g. Neck Disability Index [14], Neck Bournemouth Questionnaire [15]), to establish which one displays better measurement properties in NP patients.

Materials and methods

This clinimetric study was approved by the ethics committee of the Brindisi ASL (Italy). All procedures were conducted according to the declaration of Helsinki, and all patients provided informed consent prior to study inclusion. Authorization to adapt the CNFDS into Italian was obtained from the original developer.

Copenhagen neck functional disability scale

The CNFDS consists of 15 items that evaluate the impact of NP on headache, ability to sleep, concentration, activities related to work, daily activities and leisure activities. It includes also questions of psychosocial nature, such as decreased social contact, influence on emotional relationships with family members and attitudes toward the future. Each item can be answered as 'yes' (0 points), 'occasionally' (1 point), and 'no (2 points). To avoid repetitive answering, response options are reversed after the fifth question. The total score can range from 0 to 30, with higher scores indicating worse disability.

Translation and cross-cultural adaptation

The cross-cultural adaptation process was carried out in accordance with the recommendations proposed by Beaton et al. [16].

Step 1: Forward translation to Italian. Two Italian native speakers with good English knowledge independently translated the questionnaire into Italian. One translator was an economist with no medical background, the other was a speech therapist. Translators aimed at the conceptual equivalent of a word or phrase and used natural and acceptable language for the broadest audience range. The two translations were called T1 and T2.

Step 2: Synthesis. The two translators and four authors (DA, FB, AD, GG) discussed the translated questionnaire (T1 and T2) in a consensus meeting, to disentangle any discordant or ambiguous word and to develop a consensus-based version (T1–2).

Step 3: Backward translation to English. Two English native speakers without medical background, independently back-translated T1–2 into English. These translators were unfamiliar with the study purpose and were blinded to the original English version. Two back-translations were obtained (BT1 and BT2).

Step 4: Expert committee. The expert committee consisted of the four translators and four physiotherapists/authors (DA, FB, AD, GG). The committee reviewed all the translations (T1, T2, BT1, BT2) and compared the Italian version of the scale (T1–2) with the original version of the scale. Consensus in terms of semantic equivalence (ensuring that the words mean the same thing), idiomatic equivalence (formulation of equivalent expressions for colloquialisms), experiential equivalence

Angilecchia *et al. BMC Musculoskeletal Disorders*   (2018) 19:409

Page 3 of 9

(ensuring that each item properly captured the experience of daily life in target culture), and conceptual equivalence (ensuring that items hold the same conceptual meaning) was sought and achieved in the prefinal CNFDS version. The committee made only one change: in the first question "Can you sleep at night without neck pain interfering?" translated into Italian as "Riesce a dormire la notte senza che il dolore al collo possa disturbarla?" was changed into "Riesce a dormire la notte senza che il dolore al collo la disturbi?"

Step 5: Pretesting. The prefinal CNFDS was administered to 30 participants responding to the study inclusion criteria and who were asked about any misunderstanding, conflicting, or ambiguous word or sentence. These participants did not have any questions about the scale and all the questions were well understood. There was no multiple answer question or missing answer. After the pretesting, the final version of the scale was obtained (i.e. CNFDS-I), Additional file 1: Table S1.

Participants

Consecutive outpatients seeking treatment or evaluation between March 2015 and February 2016 at the Physical Medicine and Rehabilitation Unit of "Organizzazione Sanitaria Meridionale Assistenza Inabili Recupero Minori" (OSMAIRM) were assessed for inclusion.

The inclusion criteria were: non-specific chronic NP (lasting > 12 wk) with or without arm pain, at least 18 years old, and ability to read and speak Italian fluently. The NP area was defined according to the IASP definition [17] as the area bounded by the nuchal fold on the top, an imaginary transverse line passing through the tip of the first chest spinous process on the bottom, and a sagittal suture tangential to the lateral edges of the neck on both sides.

Patients with neurological signs in the arms were excluded because there was a lack of instrumental examination for confirming the diagnosis (i.e. cervical radiculopathy) [18]. Other exclusion criteria were: lawsuits, cognitive impairment, fractures, cancer central neurological signs and severe psychiatric disorders.

Other measurement instruments

A booklet asking information about demographic (e.g. age, sex) and clinical characteristics (e.g. pain duration) was administered to each patient. The same booklet contained the Italian versions of the Neck Disability Index (NDI-I) [14], of the Neck Bournemouth Questionnaire (NBQ-I) [15], and of the CNFDS-I; a Visual Analogue Scale (VAS) to assess pain intensity was also included [19].

The NDI is the most commonly tested and translated neck-related disability scale [20]. It consists of 10 items and each question is scored on a 6-point scale ranging from 0 (no disability) to 5 (full disability). Total score ranging from 0 to 50 which can also be expressed as a percentage. The NBQ is a short questionnaire consisting of seven items representing aspects of the biopsychosocial model relevant to patients with NP [21]. Each item is scored on a 0–10 numerical scale, where zero represents absence of limitation, for a total score ranging from 0 to 70 points. The NBQ-I has displayed acceptable construct validity and responsiveness in Italian patients with chronic NP [15].

The VAS is a horizontal line, 100 mm in length, asking patients to rate their pain intensity at the moment, "no pain" and "worst pain" are the extremes. The VAS has exhibited satisfactory test-retest reliability and construct validity in patients with chronic pain [22].

Measurement analysis

The COnsensus-based standards for the Selection of health Measurement INstruments(COSMIN) definitions of measurement properties were adopted in this study [23].

Acceptability

The time needed to answer the questionnaire was recorded. Once completed, the patients were asked about any problems they encountered and the professionals administering the questionnaire inspected for the presence of missing or multiple responses.

Structural validity

Structural validity is the degree to which scores of an instrument adequately reflect the dimensionality of the construct to be measured [23]. Since this measurement property was not previously assessed for the CNFDS, an exploratory factor analysis was performed on the 15 items of the questionnaire, using a principal component estimator. The eigenvalue of each extracted factor was calculated, and a scree plot was drawn. A ratio between the first and the second eigenvalue larger than 4 was considered as an indication of unidimensionality [24]. Factor loadings for each item were also calculated, and loadings smaller than 0.4 were considered for item reduction of the questionnaire [25].

Internal consistency

This property reflects how much the items of a questionnaire are intercorrelated and it was assessed using Cronbach's Alpha. The value is considered satisfactory if it is higher than 0.80 [26]. Correlation of each item with the total score of remaining items (item-total correlation) and inter-items correlation were also computed.

Construct validity

Construct validity was assessed by means of hypotheses testing, as suggested by the COSMIN initiative and the

International Society for Quality of Life Research (ISO-QOL) [27]. Hypotheses on expected correlations between CNFDS-I and the other questionnaires (i.e. NDI, NBQ and VAS) were tested. The correlation with NDI and NBQ was expected to be ≥0.60 as these questionnaires are considered to measure the similar construct (i.e. neck-related disability); the correlation with the VAS was expected to be ≥0.30 and < 0.60 as it measures a related (but not the same) construct. These hypotheses were in line with those formulated for similar and related constructs in a previous study in patients with NP [28]. Since data were normally distributed (i.e. Shapiro-Wilk test, p-value > 0.05), the correlations were assessed by means of Pearson's correlation coefficient (rs). Construct validity was considered satisfactory if all three hypotheses were met [29].

Test-retest reliability

For the test-retest reliability, the questionnaire was administered twice to patients not receiving any treatment, with an interval period of 3 days. As suggested by Holt et al. [30], a long interval period may be inappropriate for a test-retest study of health measures because too many changes in the patient's health status can occur. Before the second measurement, patients were asked if their NP condition had changed compared to the first measurement and only those reporting no change were re-assessed. The intraclass correlation coefficient (ICC) with a two-way mixed effect model for absolute agreement (ICCagreement) was calculated. An ICCagreement ≥0.70 indicated sufficient test-retest reliability [31].

Measurement error

Measurement error was assessed in stable patients that undertook the second measurement. The Standard Error of Measurement for agreement (SEM$_{agreement}$) was computed by undertaking the square root of the within-people ANOVA sum of squares (i.e. sum of systematic and random error). The Smallest Detectable Change (SDC) was calculated with the following formula: SEM$_{agreement}$ * 1.96 * √2. To judge whether a SDC is sufficiently small to allow measurement of a 'real' change, it should be compared with the minimal important change of an instrument [29]; however, since no minimal important change has been estimated for the CNFDS, this comparison was not feasible in this study.

Floor/ceiling effects

Floor and ceiling effects represented the percentage of patients with the minimal score (i.e. 0) and the percentage of patients with the maximum score (i.e. 30). Ceiling and floor effect were considered to be present if more than 15% of respondents achieved the lowest or highest possible total score [32].

Results

Among the examined patients, 162 were eligible for this study. Fifty patients were (31%) clinically stable at the second measurement (after at least 3 days) and completed the CNFDS-I twice to assess reliability. The mean age of patients was 47,9 (standard deviation 14,5) years, most were female (69,7%), married (66,6%) and had completed secondary or university school (82,7%). Table 1 shows sociodemographic characteristics and it reports descriptive statistics for the total scores of CNFDS-I, NDI-I, NBQ-I and VAS. Descriptive statistics, item total correlations and the item rest correlations are illustrated for each CNFDS-I item in Table 2.

Acceptability

All the questions were well accepted, there were no problems in the instrument's comprehensibility. The questionnaire was completed with a mean of 120 s (standard deviation = 40). No missing responses or multiple answers were found.

Structural validity

Exploratory factor analysis revealed that the first factor explained 83% of the total variability (eigenvalue = 4.122), while the second factor explained only 13% (eigenvalue = 0.648). The ratio between the first and the second eigenvalues was equal to 6.36. The scree plot (Fig. 1) also indicated a clear unidimensional pattern for the CNFDS-I. Factor loadings are presented in Table 2 and show as the item 9 appear deflected. We decided not to delete this item because we consider its content (Have you been bothered by headaches during the time that you have had neck pain?) very relevant to patients with NP, because it presented a sufficient item total correlation (Table 2), and because we would not like to create a CNFDS Italian version that differs from other versions in the number of items.

Internal consistency

The Cronbach α for the scale was 0.83, indicating satisfactory results. Cronbach's α when an item was excluded ranged from 0.12 to 0.69, while correlation of each item with the total score ranged from 0.24 to 0.76 (Table 2).

Construct validity

CNFDS-I correlations with the NDI-I and the NBQ-I were equal to 0.846 and 0.708, respectively. The correlation between the CNFDS-I and the VAS was 0.570. All these correlations met the expected hypotheses and the CNFDS-I construct validity was considered satisfactory.

Summary in Table 3.

Angilecchia et al. BMC Musculoskeletal Disorders (2018) 19:409

Page 5 of 9

Table 1 Sociodemographic characteristics of the population (n = 162)

Variable	Value	N	%
Sex	F	113	69.8
	M	49	30.2
Married/Helpmate	Yes	108	66.7
	No	54	33.3
Education	Elementary	10	6.2
	Mid school	18	11.1
	High school	68	42.0
	Graduate school	66	40.7
Work	No (pain)	0	0
	Student	2	1.2
	Employee	84	51.9
	Self employed	30	18.5
	Retired	23	14.2
	Unemployed	3	1.9
	Housewife	20	12.3
Smoke	Yes	33	20.4
	No	103	63.6
	Ex smoker	26	16.0
Duration symptoms	From 3 to 6 months	48	29.6
	> 6 months	114	70.4
Physical activity or workout	No	85	52.5
	Yes < 3 h/week	45	27.8
	Yes > 3 h/week	32	19.8
Mean (SD) NDI-I	%	162	28.8 (14.7)
Mean (SD) NBQ-I	0–70	162	28,4 (14.2)
Mean (SD) VAS	0–100	162	45.5 (20.9)
Mean (SD) CNFDS-I	0–30	162	10.5 (5.9)
Mean (SD) CNFDS-I test-retest, 2nd test	0–30	50	10.92 (5.9)

NDI Neck Disability Index, *NBQ* Neck Bournemouth Questionnaire, *CNFDS* Copenhagen Neck Functional Disability Scale, *VAS* Visual Analogue Scale

Test-retest reliability
The $ICC_{agreement}$ was equal to 0.997 with a 95% confidence interval ranging from 0.995 to 0.999. This result represents excellent test-retest reliability.

Measurement error
The $SEM_{agreement}$ equalled 3.0 scale points (10% of the scale range), while the SDC was 8.31 scale points (27% of the scale range).

Floor/ceiling effects
The CNFDS-I had no floor or ceiling effects, in fact no patients with a minimum and maximum score were identified in the test nor in the retest.

Discussion
Summary of the study
This study describes the cross-cultural adaptation and evaluation of the CNFDS-I in patients with chronic NP. The CNFDS-I is a unidimensional and internally consistent tool with excellent test-retest reliability and sufficient construct validity; the SDC was equal to 8.31 scale points (27% of the scale range). These results indicate that the CNFDS-I is ready to be used as a measuring method of neck-related disability in Italian patients with chronic NP.

This is the first study to perform a factor analysis of the CNFDS as the original study [8], while the previous ones [33–36] did not do that. The CNFDS-I resulted to be unidimensional and this represents a key finding, as it indicates that it is appropriate to use its total sum score [24]. Meanwhile, this finding does not support the use of the subscales proposed by the French developers [35]. In contrast with the CNFDS-I, other neck-related disability tools available in Italian (i.e. NDI-I and NPQ-I) were not shown to be unidimensional [14, 15, 37], questioning the suitability of using their total sum scores.

Strengths and limitations
The unidimensionality of the NDI has been questioned also by other studies in other language versions [38, 39]. A future study can aim at comparing directly the unidimensionality of all these tools, to check if the CNFDS is indeed the best performing method from a psychometric point of view. Item 9 was the only CNFDS-I item with a low factor loading; future studies should assess if this result is repeated in other languages.

The Cronbach's α of the entire questionnaire (α = 0.83) was major than the predefined threshold, and it was similar to the values obtained in the original version (α = 0.90) French (α = 0.83) Polish (α = 0.90) and Iranian (α = 0.84) [33–36]. The comparison is simplified in Table 3. This indicates a high interrelation of the items.

Construct validity was analyzed by comparing the CNFDS-I to the NDI-I and NBQ-I. The very high correlation between CNFDS-I and NDI-I (ρ = 0.846) suggests that the theoretical construct of these two instruments may be very similar [32]. We chose the NDI as a comparative questionnaire because it is the most validated and internationally used [15, 40]. The correlation between the total scores of CNFDS-I and NBQ-I was high too (ρ = 0.708). A second disability comparator questionnaire was included to compare the biopsychosocial aspects examined by the NBQ-I [13, 21]; the correlation between the two questionnaires outlines that their content is not totally different as described by Ferreira et al. [13], possibly because they may differ across cultures. Foremost, to fully compare the content validity of these different instruments, it is necessary to conduct a

Table 2 Descriptive Statistics, Factor Loadings, and item-total Correlations of the Items of the Questionnaire

Item	Description	Mean (SD)	Factor Loading	Item Total Correlation	Item rest Correlation
1	To sleep without pain	0.71 (0.77)	0.37	0.46	0.35
2	Daily activities levels	0.65 (0.70)	0.76	0.76	0.69
3	Daily activities without help from other	0.15 (0.49)	0.46	0.47	0.40
4	To manage clothes	0.21 (0.51)	0.49	0.52	0.45
5	To brush teeth	0.62 (0.75)	0.41	0.48	0.37
6	To spend time at home	0.63 (0.77)	0.73	0.73	0.66
7	To lift objects (2–4 kg)	0.92 (0.84)	0.46	0.52	0.41
8	To read	0.96 (0.83)	0.53	0.59	0.49
9	Headache	1.45 (0.71)	0.13	0.24	0.12
10	Ability to concentrate	0.97 (0.78)	0.52	0.58	0.48
11	Usual leisure/free time	0.88 (0.73)	0.61	0.62	0.53
12	To remain in bed	0.43 (0.68)	0.40	0.45	0.35
13	Emotional relationship	0.426 (0.694)	0.55	0.61	0.52
14	Social contacts	0.401 (0.682)	0.64	0.65	0.57
15	NP future influence	1.031 (0.814)	0.45	0.49	0.38

Fig. 1 Scree plot of eigenvalues after factor analysis. The scree plot displays the number of the factor versus its corresponding eigenvalue. It shows that the first factor explained 83% of the total variability of data, while the second factor explained only 13%
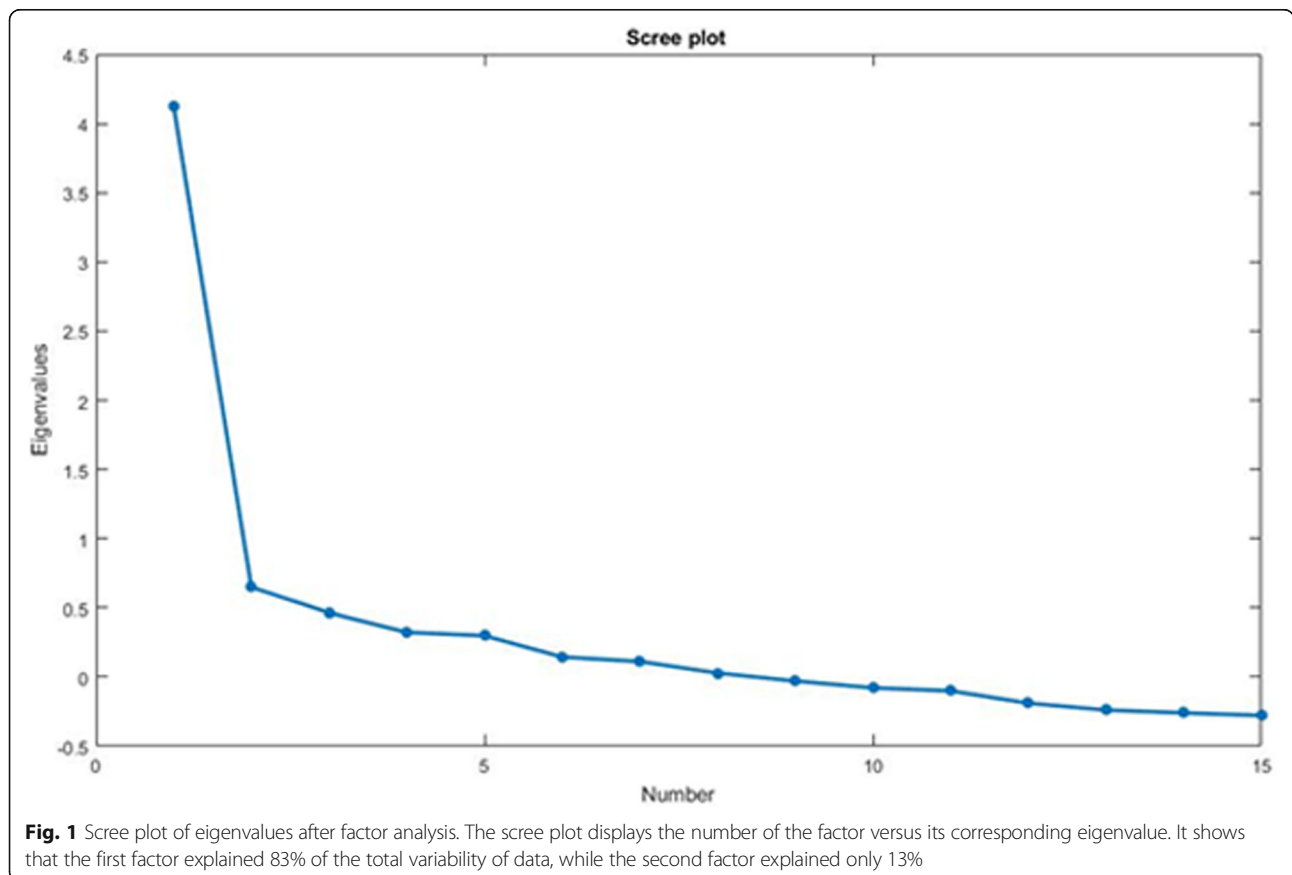
Table 3 Comparison Measurement Properties of CNFDS Adaptations and Construct Validity

	Internal Consistency (α)	Test-Retest Reliability (ICC)	SEM - SDC	Structural Validity (explained variance)	Costruct Validity (correlation)
CNFDS (Jordan A et al.)	0.90	0.99 (same day) - 0.90 (by mail 2 days later)	/ - /	/	r = 0.83 with pain score r = 0.89 with patient global assessment r = 0.56 doctor global assessment
CNFDS France (Forestier R et al.)	0.83	/	/ - /	/	r = 0.45 with VAS
CNFDS Poland (Misterska F et al.)	0.90	0.93 (24 h)	/ - /	/	r = 0.87 with NDI
CNFDS Turkey (Yapali G et al.)	/	0.86 (7 days)	/ - /	/	r = 0.78 with NPDS r = 0.73 with VAS
CNFDS Iran (Nayeb Aghaei H et al.)	0.84	0.95 (n.d.)	/ - /	/	r > 0.4 between each item and the three CNFDS subscales r = 0.80 with mJOA
CNFDS Italy	0.83	0.99 (3 days)	3 points - 8.31 points	6.36	r = 0.85 with NDI-I r = 0.71 with NBQ-I r = 0.57 with VAS

CNFDS Copenhagen Neck Functional Disability Scale, *SEM* Standard Error of Measurement, *SDC* Smallest Detectable Change, *VAS* Visual Analogue Scale, *NPDS* Neck Pain and Disability Scale, *mJOA* modified Japanese Orthopedic Association, *NDI* Neck Disability Index, *NBQ* Neck Bournemouth Questionnaire

head-to-head comparison study asking patients and clinicians if the most relevant neck-related disability aspects are included in each questionnaire [8]. Additionally, a consensus-based definition for neck-related disability should be established before further content assessment of the tools. These considerations are not totally new, as a recent systematic review clearly highlighted that the content validity of patient-reported outcome measures is understudied [41], and that a major effort should be made to fill this evidence gap.

The authors of French version [35], compared the CNFDS with the VAS scale, found a moderate correlation (0,45 Spearman's r).

Similarly, we have identified a correlation of ρ = 0.570. This result reflects those by Fejer [42], highlighting the moderate correlation between NP intensity and disability, which are strongly associated. Therefore, the CNFDS and the VAS evaluate different aspects, though maintaining a certain degree of correlation.

Moreover, after validating the CNFDS in Iran, Azhari et al. [43] investigated the most important aspects covered by the CNFDS, emphasizing the fact that its primary use is to gather and measure the patients' disability, not their perceived pain.

CNFDS-I test-retest reliability was excellent as highlighted in the original development study [8] and other studies [24, 34–36]. The ICC agreement represents an absolute measure of reliability and it is suggested that it should be interpreted in combination with a relative measure such as the SDC [44]. The SDC of the CNFDS-I was found to be 27% of the scale range; this may suggest that repeated measurements with this tool do not give similar results, but there is not a standard cut-off value to determine whether a SDS is indeed small enough.

Moreover, a measurement error larger than 20% of the scale range has been found for various broadly used questionnaires [45]. A comparison between SDC and minimal important change should be performed to have a better insight into the CNFDS measurement error. In light of these considerations, future studies on the CNFDS should calculate its minimal important change to be able to better interpret this instruments' changes in scores.

Unanswered questions and future research
This study did not assess responsiveness, defined as the ability of an instrument to detect changes in the construct to be measured [23]. It is an important measurement property to use an instrument as an outcome measurement instrument and it has not been evaluated in any language for the CNFDS. Overall, it remains unclear which neck-related disability instrument should be preferred in Italian patients with NP; thus, there is an urgent need of a head-to-head clinimetric study comparing content validity, structural validity and responsiveness of various instruments (including NDI-I and NBQ-I) in the same patients.

Conclusion
To sum up, the CNFDS-I was found to be a unidimensional, valid and reliable tool in patients with chronic NP. In absence of comparative evidence showing that the neck-related disability instrument is superior from a measurement point of view, the CNFDS-I can be used alongside other more widely used tools (e.g. NDI and NBQ), for research and routine clinical monitoring in patients with chronic NP.

Angilecchia *et al. BMC Musculoskeletal Disorders*   (2018) 19:409

Page 8 of 9

Additional files

Additional file 1: Table S1. CNFDS-I Copenhagen Neck Functional Disability Scale. The English questionnaire "Copenhagen Neck Functional Disability Scale" and the Italian version. It is reliable and valid evaluation instrument for disability in patients with neck pain. (DOCX 18 kb)

Abbreviations
CNFDS: Copenhagen neck functional disability scale; CNFDS-I: Copenhagen neck functional disability scale – Italian version; COSMIN: COnsensus-based standards for the selection of health measurement INstruments; ICC: Intraclass correlation coefficient; NBQ: Neck bournemouth questionnaire; NBQ-I: Neck Bournemouth questionnaire – Italian version; NDI: Neck disability index; NDI-I: Neck disability index – Italian version; NP: Neck pain; SDC: Smallest detectable change; SEM: Standard error of measurement; VAS: Visual analogue scale

Availability of data and materials
The datasets generated and analyzed during the current study is in part available in the Annex. Detailed data are not publicly available due to the respect of the confidentiality of data, but are achievable from the corresponding author on reasonable request.

Authors' contributions
DA designed the study, acted as project manager, led the development of the CNFDS-I, contributed substantially to the data collection and drafted the first version of the manuscript; MM performed the statistical analysis, discussed the interpretation of results and revised the manuscript for important intellectual content; AC contributed to the statistical analysis, discussed the interpretation of results and revised the manuscript for important intellectual content; AD contributed in designing the study, in the data collection and reviewed the manuscript for important intellectual content; GG contributed in designing the study, in the data collection and reviewed the manuscript for important intellectual content; FB contributed in designing the study, in organizing and conducting the data collection, in interpreting the results, and in providing important intellectual revision of the manuscript. All authors read and approved the submission of the manuscript.

Ethics approval and consent to participate
The research has been approved by the Ethics Committee ASL BR, Ref. N. 16,354 – March 9th, 2016. The Ethics Committee has analyzed the list of information to collect, written informed consent and information form. All data is collected anonymously and the manuscript not contains any individual person's data in any form. Authorization to adapt the CNFDS into Italian was obtained from the original developer by mail. Each participant signed the informed consent and provided consent to publish, as described in the "Information Form".

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Department of Physical Medicine and Rehabilitation OSMAIRM, via Cappuccini 9, 74014 Laterza, Taranto, Italy. [2]University "Tor Vergata", Faculty of Economics, Via Columbia, 2, 00133 Roma, Roma, Italy. [3]Department of Epidemiolgy and Biostatistics, Amsterdam Movement Sciences research institute, VU University Medical Center, Amsterdam UMC, Amsterdam, Netherlands. [4]Department of General Practice, Erasmus Medical Center, Rotterdam, Netherlands. [5]Centro Giovanni Paolo II - Istituto Neuromed, Viale Europa, 70017 Putignano, Bari, Italy. [6]Studio Professionale FTM, via Della Libertà 14, 73023 Lizzanello, Lecce, Italy. [7]Physioup, Via Novacella 19, 00142 Roma, Roma, Italy.

References
1. Hoy DG, Protani M, De R, Buchbinder R. The epidemiology of neck pain. Best Pract Res Clin Rheumatol. 2010;24(6):783–92.
2. Cohen SP. Epidemiology, diagnosis, and treatment of neck pain. Mayo Clin Proc. 2015;90(2):284–99.
3. Guzman J, Hurwitz EL, Carroll LJ, Haldeman S, Côté P, Carragee EJ, Peloso PM, van der Velde G, Holm LW, Hogg-Johnson S, Nordin M, Cassidy JD. A new conceptual model of neck pain: linking onset, course, and care: the bone and joint decade 2000-2010 task force on neck pain and its associated disorders. Spine (Phila Pa 1976). 2008;33(4 Suppl):S14–23.
4. Lidgren L. Preface: neck pain and the decade of the bone and joint 2000-2010. J Manip Physiol Ther. 2009;32(2 Suppl):S2–3.
5. Hogg-Johnson S, van der Velde G, Carroll LJ, Holm LW, Cassidy JD, Guzman J, Côté P, Haldeman S, Ammendolia C, Carragee E, Hurwitz E, Nordin M, Peloso P. The burden and determinants of neck pain in the general population: results of the bone and joint decade 2000–2010 task force on neck pain and its associated disorders. Spine (Phila Pa 1976). 2008;33(4 Suppl):S39–51.
6. Croft PR, Lewis M, Papageorgiou AC, Thomas E, Jayson MI, Macfarlane GJ, Silman AJ. Risk factors for neck pain:longitudinal study in the general population. Pain. 2001;93(3):317–25.
7. World Health Organization. The international classification of functioning, Disability and health (ICF). 2001; Available at: http://www.who.int/classifications/icf/en/
8. Jordan A, Manniche C, Mosdal C, Hindsberger C. The Copenhagen neck functional disability scale: a study of reliability and validity. J Manip Physiol Ther. 1998;21(8):520–7.
9. Vernon H, Mior S. The Neck Disability Index: a study of reliability and validity. J Manip Physiol Ther. 1991;14(7):409–15 Erratum in: J Manipulative Physiol Ther 1992 Jan; 15(1):followi.
10. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. Physiotherapy. 1980;66(8):271–3.
11. Waddell G, Main CJ. Assessment of severity in low back disorders. Spine. 1984;9:204–8.
12. Turk DC, Rudy TE. Toward an empirically derived taxonomy of chronic pain patients: integration of psychological assessment data. J Consult ClinPsychol. 1988;56:233–8.
13. Ferreira ML, Borges BM, Rezende IL, Carvalho LP, Soares LP, Dabes RA, Carvalho G, Drummond AS, Machado GC, Ferreira PH. Are neck pain scales and questionnaires compatible with the international classification of functioning, disability and health? A systematic review. Disabil Rehabil. 2010; 32(19):1539–46.
14. Monticone M, Ferrante S, Vernon H, Rocca B, Dal Farra F, Foti C. Development of the Italian version of the neck disability index: cross-cultural adaptation, factor analysis, reliability, validity, and sensitivity to change. Spine (PhilaPa 1976). 2012;37(17):E1038–44.
15. Geri T, Signori A, Gianola S, Rossettini G, Grenat G, Checchia G, Testa M. Cross-cultural adaptation and validation of the neck Bournemouth questionnaire in the Italian population. Qual Life Res. 2015;24(3):735–45.
16. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of selfreport measures. Spine (PhilaPa 1976). 2000;25(24):3186–91.
17. Classification of chronic pain, Second Edition (Revised), International Association for the Study of Pain (IASP),1994. Available at: http://www.iasp-pain.org/PublicationsNews/Content.aspx?ItemNumber=1673 ; Part II, Section C.
18. Thornbury JR, Fryback DG, Turski PA, Javid MJ, McDonald JV, Beinlich BR, Gentry LR, Sackett JF, Dasbach EJ, Martin PA. Disk-caused nerve compression in patients with acute low-back pain: diagnosis with MR, CT myelography, and plain CT. Radiology. 1993;186(3):731–8.
19. Huskisson EC. Measurement of pain. Lancet. 1974;2(7889):1127–31.

20. Schellingerhout JM, Verhagen AP, Heymans MW, Koes BW, de Vet HC, Terwee CB. Measurement properties of disease-specific questionnaires in patients with neck pain: a systematic review. Qual Life Res. 2012; 21(4):659–70.

21. Bolton JE, Humphreys. The Bournemouth questionnaire: a short-form comprehensive outcome measure. II. Psychometric properties in neck pain patients. J Manip Physiol Ther. 2002;25(3):141–8 BK.

22. Hjermstad MJ, Fayers PM, Haugen DF, Caraceni A, Hanks GW, Loge JH, Fainsinger R, Aass N, Kaasa S. Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults: a systematic literature review. J Pain Symptom Manag. 2011;41(6):1073–93.

23. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for healthrelated patient-reported outcomes. J Clin Epidemiol. 2010; 63(7):737–45.

24. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, Thissen D, Revicki DA, Weiss DJ, Hambleton RK, Liu H, Gershon R, Reise SP, Lai JS, Cella D. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the patient-reported outcomes measurement information system (PROMIS). Med Care. 2007;45(5 Suppl 1):S22–31.

25. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine. USA: Cambridge University Press, New York; 2011.

26. Andresen EM. Criteria for assessing the tools of disability outcomes research. Arch Phys Med Rehab. 2000;81(12 Suppl 2):S15–20.

27. Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, Schwartz C, Revicki DA, Moinpour CM, McLeod LD, Lyons JC, Lenderking WR, Hinds PS, Hays RD, Greenhalgh J, Gershon R, Feeny D, Fayers PM, Cella D, Brundage M, Ahmed S, Aaronson NK, Butt Z. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Qual Life Res. 2013;22(8):1889–905.

28. Chiarotto A, Falla D, Polli A, Monticone M. Validity and responsiveness of the pain self-efficacy questionnaire in patients with neck pain disorders. J Orthop Sports Phys Ther. 2017;19:1–36. https://doi.org/10.2519/jospt.2018.7605.

29. Prinsen CA, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, Williamson PR, Terwee CB. How to select outcome measurement instruments for outcomes included in a "Core outcome set" - a practical guideline. Trials. 2016;17(1):449.

30. Holt AE, Shaw NJ, Shetty A, Greenough CG. The reliability of the low Back outcome score for back pain. Spine. 2002;27:206–10.

31. Fleiss JL. The measurement of interrater agreement. In: Statistical methods for rates and proportions, 2nd edition. NewYork: John Wiley; 1981. p. 212–36.

32. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42.

33. Yapali G, Günel MK, Karahan S. The cross-cultural adaptation, reliability, and validity of the Copenhagen neck functional disability scale in patients with chronic neck pain: Turkish version study. Spine (Phila Pa 1976). 2012;37(11):E678–82.

34. Nayeb Aghaei H, Azimi P, Shahzadi S, Azhari S, Mohammadi HR, Alizadeh P, Montazeri A. Outcome measures of functionality, social interaction, and pain in patients with cervical Spondylotic myelopathy: a validation study for the Iranian version of the Copenhagen NeckFunctional disability scale. Asian Spine J. 2015;9(6):901–8.

35. Forestier R, Françon A, Arroman FS, Bertolino C. French version of the Copenhagen neck functional disability scale. Joint Bone Spine. 2007;74(2):155–9 Vi.

36. 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 Musculoskelet Disord. 2011;12:84.

37. Geri T, Piscitelli D, Meroni R, Bonetti F, Giovannico G, Traversi R, Testa M. Rasch analysis of the neck Bournemouth questionnaire to measure disability related to chronic neck pain. J Rehabil Med. 2015;47(9):836–43.

38. Ailliet L, Knol DL, Rubinstein SM, de Vet HC, van Tulder MW, Terwee CB. Definition of the construct to be measured is a prerequisite for the assessment of validity. The neck disability index as an example. J Clin Epidemiol. 2013;66(7):775–82 quiz 782.e1–2.

39. Bakhtadze MA, Vernon H, Zakharova OB, Kuzminov KO, Bolotov DA. The neck disability index-Russian language version (NDI-RU): a study of validity and reliability. Spine (Phila Pa 1976). 2015;40(14):1115–21.

40. Pellicciari L, Bonetti F, Di Foggia D, Monesi M, Vercelli S. Patient-reported outcome measures for non-specific neck pain validated in the Italian-language: a systematic review. Arch Physiother. 2016;6:9.

41. Chiarotto A, Ostelo RW, Boers M, Terwee CB. A systematic review highlights the need to investigate the content validity of patient-reported outcome measures for physical functioning in low back pain. J Clin Epidemiol. 2018; 95:73-93.

42. Fejer R, Hartvigsen. Neck pain and disability due to neck pain: what is the relation? Eur Spine J. 2008;17(1):80–8.

43. Azhari S, Azimi P, Shazadi S, Khayat Kashany H, Nayeb Aghaei H, Mohammadi HR. Surgical Outcomes and Correlation of the Copenhagen neck functional disability scale and modified Japanese orthopedic association assessment scales in patients with cervical Spondylotic myelopathy. Asian Spine J. 2016;10(3):488–94.

44. de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. J Clin Epidemiol. 2006;59(10):1033–9.

45. Chiarotto A, Terwee CB, Ostelo RW. Choosing the right outcome measurement instruments for patients with low back pain. Best Pract Res Clin Rheumatol. 2016;30(6):1003–20.