# Validity and Reliability of the Arabic Version of the Copenhagen Neck Functional Disability Scale in Neck Pain Patients

Ahmed Mohammed Elbeltagy[1], Wadida Hassan El Sayed[2], Soheir Shehata Rezk Allah[2]

[1]Al Bank Al Ahly Hospital, Ministry of Health, Cairo, Egypt
[2]Department of Basic Science, Faculty of Physical Therapy, Cairo University, Giza, Egypt

**Study Design:** Prospective observational study.

**Purpose:** To test the face validity, content validity, feasibility, internal consistency, reliability, and test–retest reliability of the Arabic version of the Copenhagen Neck Functional Disability Scale (CNFDS).

**Overview of Literature:** CNFDS is a clinical evaluation tool that accurately reflects the patient's perception regarding his/her functionality with existing cervical pain. This study aimed to investigate the validity and reliability of the Arabic version of the CNFDS in measuring the disability level in chronic neck pain patients.

**Method:** Seventy-four patients with neck pain were recruited, and 135 sheets (test and retest sheets) were completed by patients; two expert panels (each comprising ten experts) participated in this study. Arabic translation (forward translation), development of the preliminary translated version, English translation (backward translation), development of the prefinal version, and testing of the prefinal version was performed by experts; thereafter, the final version was tested on patients. Index of clarity, expert proportion of clearance, index of content validity (CVI), expert proportion of relevance, descriptive statistics, missed item index, Cronbach's alpha, and Spearman's rank correlation coefficient were used for statistical analyses.

**Results:** The study revealed that the scale index of clarity was 86.84%, scale CVI was 99.33%, scale-level content validity index (universal agreement method) was 99.33%, 99.15% of the scale items were filled in all sheets, the scale was answered in less than 3 minutes in about 75% cases, Cronbach's alpha was 0.856 (0.796, 0.905), and all Spearman's correlations between the test and retest results were statistically significant.

**Conclusions:** The Arabic version of the CNFDS has adequate validity and reliability for the measurement of the disability level in chronic neck pain patients.

**Keywords:** Validity; Reliability; Arabic Copenhagen Neck Functional Disability Scale; Neck pain

## Introduction

Neck pain is one of the most common chronic pain prob-

lems, with a reported prevalence of 22%–30%. The types of disabilities resulting from neck pain have not been determined; however, it usually exerts a significant effect on daily

life, resulting in extensive use of healthcare resources [1].

Pain in the neck (cervical pain) may be due to different reasons. The term "cervical pain" does not refer to a disease itself but rather to a symptom or the manifestation of a muscular pain disorder. Cervical pain usually involves multi-factorial clinical aspects because it may involve individual risk factors, either physical or emotional characteristics, or be related to ergonomics and work activities [2].

Many scales provided in English-language that measures neck pain and functional disability have been developed as follows: Neck Disability Index (NDI), Pain Disability Index, Copenhagen Neck Functional Disability Scale (CNFDS), Northwick Park Neck Pain Questionnaire (NPQ), Neck Pain and Disability Scale (NPDS), Neck Bournemouth Questionnaire, and Functional Rating Index [3].

To assess neck pain, many scales have been developed. The CNFDS is one such scale that comprises 15 items that assess the impact of neck pain and disability level. Three items assess the pain severity, including the patient's acuity of the future impact of neck pain; eight items assess the disability during every-day activities; and four items focus on social interaction and recreation [4].

The CNFDS requires to be translated to make it easily understandable by the patients. In contrast to the NPDS, the CNFDS has no Visual Analog Scale that is perceived as difficult to apply by some patients. The amount of text is minimum compared with the 54 lines in the NPQ. Most importantly, the CNFDS uses qualitative items, and the clinical relevance of the changes in the item scores is readily perceived [5].

## Materials and Methods

### 1. Participants and design

This study was conducted in the outpatient clinics of the Al Bank Al Ahly Hospital to investigate the validity and reliability of the Arabic version of the CNFDS in patients with chronic neck pain. This study used previous reports that suggested strategies for translating, adapting, and validating the psychometric properties of instruments [6,7]. Seventy-four literate patients with chronic neck pain who were referred by a physician, did not have any malformations, and were not on any medications were enrolled. Two professional panels, each comprising 10 experts with at least 10-year experience or a master's degree were also involved in the study. The Board Council

of Higher Education of the School of Physical Therapy, the Institutional Review Board of Higher Education and Research of Cairo University, and the Supreme Council of Universities at Egypt approved our study (IRB approval no., /P.T.REC/012/001369).

### 2. Procedures

The following steps were as follows.

1) The translation of the CNFDS was kindly approved by the author.

2) Forward translation (targeted translation): translation of the primary scale from English to Arabic: two translators participated in forward translation. The scale in English was translated to Arabic to produce two Arabic (Ar1 and Ar2) versions of the scale.

3) Development of the Arabic version that is preliminary Both versions (Ar1 and Ar2) were compared and combined by the researchers: blind back translation (blind revision translation or blind double revision translation) of the preliminary Arabic version of the scale. Two translators performed the backward translation. The preliminary translated Arabic version of the scale was translated to English to generate two back-translated English revision versions (En1 and En2).

4) Comparison of the two back-translated English versions of the scale (En1 and En2): researchers of this work compared the two English back translation versions of the scale (En1 versus En2) and also compared both En1 and En2 with the original English scale in terms of the instructions, items, responses format, word meanings, sentence composition, significance, and relations. They found no significant differences between them; therefore, the preliminary translated Arabic version was considered the prefinal Arabic version of the scale.

5) Pilot study to test the intrarater reliability of the prefinal Arabic version of the scale: the first expert panel was requested to evaluate each word (instructions, items, and choices) of the tool for clarity using a dichotomous scale (yes/no). If the answer was no, they were required to give suggestions for improving its clarity. According to the suggestions of the first expert panel, slight changes

were made to improve the clarity index to the minimum acceptable value (80%) so that it could be filled by the patients. The second expert panel was requested to rate and evaluate all the items of the scale for content equivalence (relevance) using a scale comprising four grades: 1 means not relevant, 2 means unable to assess the relevance, 3 means relevant but needs minor alterations, and 4 means very relevant and succinct. They were then requested to provide suggestions to improve its relevance, and the results were interpreted as follows: first and second grades were considered not relevant, whereas the third and fourth grades were considered relevant. After the prefinal version passed the clarity and relevance test, it was called the final Arabic version of the CNFDS.

6) Pilot study to test the intrarater reliability of the final Arabic version of the scale was conducted on chronic neck pain patients as follows: patients filled the data collection sheet that was used to collect demographic data (name,

age, and sex), and the final Arabic version of the CNFDS.

7) Patients completed the sheet again after 2 days.

**3. Statistical analyses**

For data analyses, the IBM SPSS computer program ver. 23.0 (IBM Corp., Armonk, NY, USA) was used. Clarity index and expert proportion of the clearance were used to test face validity. Index of content validity (CVI) and expert proportion of relevance were used to test the content validity. Feasibility index was calculated using the missed item index and time taken to fill the questionnaire. Cronbach's coefficient alpha was used to measure the internal consistency reliability. Test–retest reliability was measured using mean scores and Spearman's rank correlation.

## Results

The scale index of clarity was 86.84%, as shown in Table 1. Expert proportion of clearance was calculated to be 81.58%, as presented in Table 2.

Item CVI of the final version was calculated; we found that the scale CVI (S-CVI) was 99.33%, as shown in Table 3. Expert proportion of relevance was calculated and is presented in Table 4. The study population comprised patients of both sexes (44 women and 30 men); further, 61 patients did and 13 did not perform the retest. Internal consistency was calculated for the final version; Cron-

**Table 1.** Item index of clarity of the prefinal Arabic version of the scale

| Item | No. of rater's agreements (clear responses) | Item index of clarity (%) |
|---|---|---|
| Instructions | 10 | 100 |
| (1) | 8 | 80 |
| (2) | 8 | 80 |
| (3) | 9 | 90 |
| (4) | 8 | 80 |
| (5) | 9 | 90 |
| (6) | 9 | 90 |
| (7) | 9 | 90 |
| (8) | 9 | 90 |
| (9) | 8 | 80 |
| (10) | 8 | 80 |
| (11) | 8 | 80 |
| (12) | 9 | 90 |
| (13) | 7 | 70 |
| (14) | 9 | 90 |
| (15) | 10 | 100 |
| (a) | 9 | 90 |
| (b) | 9 | 90 |
| (c) | 9 | 90 |
| Mean | 8.6 | 86.84 |

(1)→(15) items of the scale in order, (a) first response word, (b) second response word, and (c) third response word.

**Table 2.** Expert proportion of clearance of the prefinal Arabic version of the scale

| Item | No. of rater's agreements (clear responses) | Item index of clarity (%) |
|---|---|---|
| (1) | 18 | 95 |
| (2) | 18 | 95 |
| (3) | 18 | 95 |
| (4) | 18 | 95 |
| (5) | 16 | 84 |
| (6) | 13 | 68 |
| (7) | 17 | 89 |
| (8) | 10 | 53 |
| (9) | 18 | 95 |
| (10) | 9 | 47 |
| Mean | 8.15 | 81.58 |

(1)→(10) experts.

bach's alpha was 0.856 with lower bound 0.796 and upper bound 0.905 at 95% confidence interval (CI). Finally,

Spearman's correlation coefficients were calculated and are presented in Table 5.

## Discussion

The present study was designed to test the face validity, content validity, feasibility, internal consistency reliability, and test–retest reliability of the Arabic version of the CNFDS to measure the disability level in chronic neck pain

**Table 3.** Item index of content validity of the final version

| Item | No. of raters that agree (relevant responses) | I-CVI[a] (%) |
|---|---|---|
| (1) | 10 | 100 |
| (2) | 10 | 100 |
| (3) | 10 | 100 |
| (4) | 10 | 100 |
| (5) | 10 | 100 |
| (6) | 10 | 100 |
| (7) | 10 | 100 |
| (8) | 10 | 100 |
| (9) | 10 | 100 |
| (10) | 10 | 100 |
| (11) | 10 | 100 |
| (12) | 10 | 100 |
| (13) | 10 | 100 |
| (14) | 9 | 90 |
| (15) | 10 | 100 |
| Mean | 9.93 | 99.33 |

(1)→(15) items of the scale in order.
[a]Item index of content validity index.

**Table 4.** Expert proportion of relevance of the final version

| No. of expert | No. of agreements (relevant responses) | Proportion of relevance (%) |
|---|---|---|
| 1 | 15 | 100 |
| 2 | 15 | 100 |
| 3 | 15 | 100 |
| 4 | 15 | 100 |
| 5 | 15 | 100 |
| 6 | 15 | 100 |
| 7 | 14 | 93 |
| 8 | 15 | 100 |
| 9 | 15 | 100 |
| 10 | 15 | 100 |
| Mean | 14.9 | 99.33 |

**Table 5.** Spearman's rank correlations coefficients

| No. of item | $R$-value[a] | Correlation strength | Results of test regarding association between test and retest results |
|---|---|---|---|
| 1 | 0.872 | Very strong | Statistically significant |
| 2 | 0.813 | Very strong | Statistically significant |
| 3 | 0.837 | Very strong | Statistically significant |
| 4 | 0.910 | Very strong | Statistically significant |
| 5 | 0.889 | Very strong | Statistically significant |
| 6 | 0.891 | Very strong | Statistically significant |
| 7 | 0.946 | Very strong | Statistically significant |
| 8 | 0.817 | Very strong | Statistically significant |
| 9 | 0.851 | Very strong | Statistically significant |
| 10 | 0.748 | Strong | Statistically significant |
| 11 | 0.700 | Strong | Statistically significant |
| 12 | 0.924 | Very strong | Statistically significant |
| 13 | 0.863 | Very strong | Statistically significant |
| 14 | 0.871 | Very strong | Statistically significant |
| 15 | 0.939 | Very strong | Statistically significant |
| Total score | 0.958 | - | Statistically significant |

[a]Spearman's rank correlation.

patients. The Arabic version of the CNFDS exhibited good face validity, with scale index of clarity 86.84% and mean proportion of clearance (clear responses) 81.58%. The scale also demonstrated excellent content validity because the S-CVI was 99.33%, S-CVI/universal agreement was 93.33%, and the mean proportion of relevance (relevant responses) was 99.33%. The present results are in agreement with the reports of Polit and Beck [8] who stated that for a scale to be judged as having excellent content validity, it should be composed of items with item indices of content validity (I-CVI) that meet the following criteria (I-CVI of 1.00 with three to five experts and I-CVI ≥0.78 for 6–10 experts) and have S-CVI ≥0.90. The recommended standards may necessitate two rounds of expert review if the initial assessment suggests the need for substantial item improvements. This is consistent with the findings of Waltz et al. [9] who stated that S-CVI/average ≥0.90 is the minimum acceptable index and items that do not achieve the minimum acceptable indices are revised and reevaluated.

The Arabic version of the CNFDS demonstrated high feasibility as the scale items were filled in all the sheets by 99.15%, and ≤1.5 minute was required to complete the sheet in 64% of all sheets. Further, <5 minutes were required for almost all the sheets.

The present results agreed with the report by van Veen et al. [10], according to which the missing rate on the item level was considered acceptable if no single item had a missing rate >10%, and the completion time was considered acceptable if 95% of the sheets were completed in <15 minutes.

The Arabic version of the CNFDS exhibited good internal consistency and good test–retest reliability, with a Cronbach's alpha of 0.856 (0.796, 0.905) and all Spearman's rank correlation coefficients between test and retest results being statistically significant (item 1=0.872, item 2=0.813, item 3=0.837, item 4=0.910, item 5=0.889, item 6=0.891, item 7=0.946, item 8=0.817, item 9=0.851, item 10=0.748, item11=0.700, item 12=0.924, item 13=0.863, item 14=0.871, item 15=0.939, and total score=0.958). Therefore, according to Darren and Mallery [11], an α value between 0.7 and 0.9 indicates good internal consistency. Further, Spearman's rank correlation coefficient between 0.9 and 1 (as in items 4, 7, and 15) indicates excellent test–retest reliability, and Spearman's rank correlation coefficient between 0.7 and 0.9 (as in items 1, 2, 3, 5, 6, 8, 9, 10, 11, 12, 13, and 14) indicates acceptable test–retest reliability.

These results agree with a study that tested the out-

come measures of functionality, social interaction, and pain in patients with cervical spondylotic myelopathy. A validation study for the Iranian version of the CNFDS and a form including the individual characteristics of the patients were applied to 93 subjects with cervical spondylotic myelopathy. The Cronbach's alpha coefficient was 0.84, and intraclass correlation coefficient (ICC) was 0.95, certifying that the Iranian version of the CNFDS had acceptable reliability. Moreover, the correlation of each item with its hypothesized domain on the CNFDS was acceptable, suggesting that the items had a good relationship with their own domains. These results also indicate that the instrument was responsive to change ($p<0.0001$). Therefore, the Iranian version of the CNFDS was considered to have acceptable psychometric characteristics, based on the promising results of the ICCs for the scale scores and the good stability of the Iranian version of the CNFDS over time. Therefore, the CNFDS appears to provide internal consistency and reproducible results for patients with neck pain [12].

These results were similar to those reported by Yapali et al. [13] who conducted a study to adapt the CNFDS to the Turkish language and evaluate the psychometric properties of the Turkish version of the CNFDS in patients with chronic neck pain. The scale was translated from the source language to the target language, synthesized, back translated, revised, and pre-tested. Overall, 101 patients participated in the study. Test–retest reliability was tested using ICCs. The internal consistency was excellent, with an ICC value of 0.86 at 1 week, confirming the reliability of the Turkish version of CNFDS for patients with neck pain [13].

The present results are strengthened by the report by Misterska et al. [5] who conducted a study to assess the validity and the reliability of the NDI and CNFDS in patients with neck pain due to degenerative and discopathic disorders. The psychometric properties of the Neck Disability Index Polish versions (NDI-PL) and Copenhagen Disability Index Polish Version (CDI-PL), 60 patients were treated due to degenerative and discopathic disorders in the cervical spine filled out the completed patient-reported questionnaires (NDI-PL and CDS-PL). The Cronbach's alpha values were excellent for the NDI-PL in the test and retest (0.84 and 0.85, respectively) and for the CDS-PL (0.90 in the test and retest). ICCs were excellent for the CDS-PL and NDI-PL at 0.93 (95% CI, 0.89 to 0.95) and 0.87 (95% CI, 0.80 to 0.92), respectively. This study found that the NDL-PL and CDS-PL scales were valid and

reliable for patients with degenerative changes in the cervical spine.

The validity and reliability of the translated tools were assessed in two or three studies. The first study was designed to translate the tool to the targeted language; thereafter, the translated version was tested for face and content validity and reliability on a monolingual population. The second study was designed to test the full psychometrics of the translated tool in bilingual participants. The third study tested all the psychometric properties of the translated tool on a monolingual population, which indicated that there is no need for the second study to be made [6,7]. To our knowledge, the present study is the first trial to test the validity and reliability of the Arabic version of the CNFDS.

The final version is considered the base for future research that will establish all the psychometric properties of the Arabic version of the CNFDS.

## Conclusions

The present results and the corresponding discussion show that the Arabic version of the CNFDS possesses adequate face and content validity, feasibility, internal consistency, and test–retest reliability to enable the measurement of the disability level in chronic neck pain patients.

## Conflict of Interest

No potential conflict of interest relevant to this article was reported.

## Acknowledgments

## References

1. Wolsko PM, Eisenberg DM, Davis RB, Kessler R, Phillips RS. Patterns and perceptions of care for treatment of back and neck pain: results of a national survey. Spine (Phila Pa 1976) 2003;28:292-7.
2. Viana PB, Benini LV, Vasconcellos C. Programa de ginastica laboral versus desconforto laboral. Colecao Pesquisa em Educacao Fisica 2011;10:125-32.
3. Ferreira ML, Borges BM, Rezende IL, et al. Are neck pain scales and questionnaires compatible with the international classification of functioning, disability and health?: a systematic review. Disabil Rehabil 2010;32:1539-46.
4. Forestier R, Francon A, Arroman FS, Bertolino C. French version of the Copenhagen neck functional disability scale. Joint Bone Spine 2007;74:155-9.
5. 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.
6. Borsa JC, Damasio BF, Bandeira DR. Cross-cultural adaptation and validation of psychological instruments: some considerations. Paideia (Ribeirao Preto) 2012;22:423-32.
7. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. J Eval Clin Pract 2011;17:268-74.
8. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported?: critique and recommendations. Res Nurs Health 2006;29:489-97.
9. Waltz CF, Strickland OL, Lenz ER. Measurement in nursing and health research. 3rd ed. New York (NY): Springer Publishing Company; 2005.
10. Van Veen MJ, Birnie E, Poeran J, Torij HW, Steegers EA, Bonsel GJ. Feasibility and reliability of a newly developed antenatal risk score card in routine care. Midwifery 2015;31:147-54.
11. Darren G, Mallery P. SPSS for Windows step by step: a simple guide and reference 11.0 update. 4th ed. Boston (MA): Allyn and Bacon; 2003.
12. Aghaei HN, Azimi P, Shahzadi S, et al. Outcome measures of functionality, social interaction, and pain in patients with cervical spondylotic myelopathy: a validation study for the Iranian version of the Copenhagen Neck Functional Disability Scale. Asian Spine J 2015;9:901-8.
13. Yapali G, Gunel 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:E678-82.