Spine

HEALTH SERVICES RESEARCH

OPEN

Cross-cultural Adaptation, Reliability, Validity, and Responsiveness of the Simplified-Chinese Version of Neck Disability Index

Hanniel Han Rong Lim, M ClinPhys,[a] Zhi Yin Tang, BSc (Hons),[a]
Masayu Afiqah Binte Masagoes Hashim, BSc (Hons),[a] Mingxing Yang, MMuscsklSportPhysio,[a]
Eileen Yi Ling Koh, BSc,[b] and Kim Hwee Koh, FCFPS, FAMS[c]

Study Design. Cross-cultural adaptation and psychometric evaluation.
Objective. The aim of this study was to translate the Neck Disability Index (NDI) into the simplified-Chinese language and to evaluate the reliability, validity, and responsiveness of the new questionnaire.
Summary of Background Data. Neck pain is a major health problem resulting in major disability. NDI is the most frequently used scale for self-rating of disability due to neck pain. At present, there is no simplified-Chinese version of the NDI. The aims of this study were to culturally adapt and translate the NDI into the simplified-Chinese language (NDI-SC) and to evaluate its psychometric properties in patients with neck pain.
Methods. The NDI was translated into simplified-Chinese version based on established guidelines. A total of 70 patients participated in this study. Patients were asked to complete a set of questionnaires comprising of their demographic information, the NDI-SC, and a visual analog scale (VAS) of pain. Fifty-six patients returned after 1 to 2 weeks to complete the same set of questionnaires and the global rating of change (GROC) scale.

Then, the NDI-SC was evaluated for content validity, construct validity, internal consistency, test-retest reliability, and responsiveness.
Results. The NDI-SC demonstrated excellent internal consistency (Cronbach $\alpha = 0.92$) and good test-retest reliability (ICC$_{2,1}$ = 0.85). Content validity was confirmed as no floor or ceiling effects were detected for the NDI-SC total score. Construct validity was established with factor analysis revealing two-factor subscales explaining 66% of the variance. The NDI-SC showed a strong correlation with VAS (Rp = 0.61, $P < 0.001$) and a moderate correlation with GROC (Rs = 0.46, $P < 0.001$). The correlation between NDI-SC change scores and VAS change scores was also moderate (R$_p$ = 0.59, $P < 0.001$).
Conclusion. The results showed that the NDI-SC is a reliable, valid, and responsive instrument to measure functional limitations in patients with neck pain.
Key words: neck disability index, reliability, responsiveness, simplified-Chinese, translation, validity.
Level of Evidence: 3
Spine 2020;45:541–548

Neck pain is a major health problem worldwide, affecting more than 30% of the general population annually.[1–3] It is a highly prevalent musculoskeletal condition and poses significant economic and health burden.[4] In Singapore, neck pain remains the highest reported musculoskeletal disorder in office workers.[5] Although many definitions of neck pain exist, the International Association for the Study of Pain has defined neck pain as pain perceived anywhere in the posterior region of the cervical spine, from the superior nuchal line to the first thoracic spinous process.[6]

Neck pain can have numerous negative effects on a person's functional ability, work activities, and quality of life.[7–9] The intensity of symptoms can also vary largely, causing a similarly large variance in self-reported disability as a result of neck pain.[10] Quantification of neck pain is, therefore, necessary to determine and understand how it

may impact patients' perception of disability and assessment of clinical outcomes.[11] Consequently, this would empower clinicians' decision-making in the management of these patients.[12]

The Neck Disability Index (NDI) is the most commonly used patient self-reported measure of neck pain symptoms and its effect on functions and activities.[13,14] It has been shown to be a valid and reliable tool[15] that has been translated and validated in various languages.[16–28] To our knowledge, there is currently no simplified-Chinese language for neck pain and disability measures. Translating the NDI into the simplified-Chinese language instead of development of a new comprehensive instrument would allow for comparisons of different populations, and permits clinicians and researchers to exchange information across cultural and linguistic barriers. Therefore, the aims of this study are to translate and culturally adapt the NDI into the simplified-Chinese version and to evaluate the reliability, validity, and responsiveness of the new questionnaire in patients with neck pain, http://links.lww.com/BRS/B496.

METHODS

Instruments

NDI
The NDI is a condition-specific instrument for self-report of disability adapted from the Oswestry Low Back Pain Questionnaire.[29] It is a valid and reliable questionnaire[30] and comprises of 10 items relating to pain intensity, personal care, lifting, reading, concentration, work, driving, sleeping, recreation, and headache. Each question has six possible responses that score between 0 (no pain and no functional limitation) and 5 (worst pain and maximal limitation). Patients are required to choose an answer that best reflects their condition at the present time. The total score is then presented as a percentage with higher scores representing greater disability.

Translation and Cultural Adaptation
The linguistic validation procedure was initiated after contacting the developer and copyright holder of the instrument with permission obtained for the purpose of this study. This procedure was based on previous guidelines established by Beaton *et al*.[31] Two independent physiotherapists translated the questionnaire into simplified-Chinese (forward translation). The mother tongue of both translators is the simplified-Chinese language and they are also proficient in the English language. Both forward translations were then compared and discussed by the translators and authors to obtain consensus. The consensus version was then backward translated by two other independent English-speaking translators. Both were unaware of the questionnaire concept and had not seen the original English questionnaire. The expert review committee consisted of all the translators, the authors, and two other experienced physiotherapists

reviewed all translations for semantic, idiomatic, experiential, and conceptual equivalence. After reaching consensus, a preliminary final simplified-Chinese version (NDI-SC) was determined. The preliminary final version was then tested on a small sample of 10 patients with neck pain to determine whether all questions are clear and understandable. No modifications were needed following the preliminary test.

Visual Analogue Scale for Pain
The Visual Analogue Scale (VAS) consists of an 11-point scale on a 100-mm horizontal line with 0 representing the words "no pain" and 10 being the "worst possible pain" at each ends.[32] Patients were asked to quantify their current neck pain by drawing a vertical mark on the area of the horizontal line that best represented their current perception of pain level. The VAS has been shown to be a reliable and valid tool to measure pain intensity.[33,34]

Global Rating of Change
The Global Rating of Change (GROC) assesses self-perception of change in the patient's condition between sessions.[35] Participants are asked to rate the change in their condition on a 15-point transitional scale from −7 (a very great deal worse) to 7 (a very great deal better).

Participants
The NDI-SC questionnaire was administered to patients with neck pain, seeking physiotherapy care within a primary healthcare polyclinic in Singapore. Patients eligible for the study were consecutively recruited between January and June 2019. Screening of the participants was carried out by physiotherapists with more than 9 years of experience. Eligibility criteria were: the presence of neck pain, age between 21 and 70 years, the ability to read simplified-Chinese, and absence of symptoms below the elbows related to specific neck disorders. Patients were excluded if they had any of the following co-morbid diagnoses: inflammatory diseases, current infection, cancer or suspected tumors, history of fracture and surgery on the cervical spine, cervical myelopathy or radiculopathy, or clinically recognizable cognitive impairments. All participants provided written informed consent before the study. This study was approved by the SingHealth Centralized Institutional Review Board (CIRB reference 2018/3054), Singapore.

Procedures
During the first visit, all participants completed the NDI-SC and VAS questionnaires. Demographic information such as age, sex, current educational level, and duration of pain were recorded. Participants then returned at 1 to 2 weeks later to complete the VAS, GROC, and NDI-SC with changed reordered items. The selection of this interval period to assess reproducibility by retest was to allow more realistic estimates of the variability to be observed among control subjects in a longitudinal study.[36]

Statistical Analyses

All statistical procedures were conducted using IBM SPSS 25 (IBM Corp., Armonk, NY). The critical values for significance were set at $P < 0.05$. All data were assessed for completeness. Descriptive statistics were used to describe the demographic and clinical characteristics of all participants. Mean scores and standard deviations (SD) were calculated at item-level and total scores for both administrations of the NDI-SC.

Reliability

Reliability was assessed through internal consistency, test-retest reliability, and measurement errors. Cronbach α was calculated to determine the internal consistency of the NDI-SC.[37] α values >0.8 are considered to be good-excellent.[38] Test-retest reliability was calculated using Intraclass Correlation Coefficient (ICC) and Bland and Altman method. The primary reliability measure was based on an ICC (2,1) model with two-way random effects model of variance, and absolute agreement definition reporting single measures, as participants completed the NDI-SC only once in each session.[39] ICC values >0.75 are indicative of good-excellent reliability.[39] A Bland and Altman plot illustrates the spread of the difference in scores between the test and retest scores for each individual. It is expected that 95% of the differences to be less than two SD.[40] The size of the retest sample was estimated based on a method developed to calculate the required number of subjects in a reliability study.[41] Parameters regarding the probability of error type I and type II were $\alpha = 0.05$ and $\beta = 0.20$, respectively. Following these assumptions, at least 46 participants will be necessary for the test-retest analysis. Participants who scored between −3 and +3 on the GROC were included in the test-retest analysis and are assumed that they did not demonstrate any clinically relevant changes during this interval period. Measurement errors were determined by calculating the standard error of measurement (SEM) and the minimal detectable change (MDC).

Content Validity

Content validity was assessed by the completeness of the item responses in NDI-SC and the size of floor and ceiling effects. Floor and ceiling effects were considered to be present if >15% of the participants achieved the lowest or highest possible total score, respectively.[41]

Construct Validity

Construct validity was assessed through exploratory factor analysis using varimax rotation and confirmatory factor analysis.[42] Pearson correlation was assessed to examine the correlation between NDI-SC and VAS. In accordance to COSMIN guidelines,[43] 70 participants will be necessary as minimally seven samples are required for each item of the NDI-SC. To account for an estimated 10% drop-out rate, at least 80 participants will be recruited.

Responsiveness

Responsiveness was analyzed by analyzing the Spearman and Pearson correlation coefficients to quantify the relationship between NDI-SC change scores to the GROC, and also the change scores of the NDI-SC with the change scores of the VAS, respectively. Both Pearson and Spearman correlation coefficient values >0.6 are considered to be strong.[44]

RESULTS

Participants

Eighty-five patients with neck pain were enrolled in this study. Of these, six did not meet the inclusion criteria and were excluded from the study. Of the remaining subjects who were eligible, 70 provided informed consent and participated in the study. Of those, 14 (20%) did not return or complete the questionnaires for a second time. The remaining questionnaires were completed and none had more than two items of the NDI-SC missing. The final sample consisted of 70 participants with valid NDI-SC scores for statistical analysis (Figure 1).

All descriptive statistics were reported using mean ± SD. For the study, there were 70 participants (37 male, 33 female) with a mean age of 44 ± 12 years. Demographic and clinical characteristics are summarized in Table 1.

NDI-SC Instrument

Of the 70 NDI-SC scores included in the analyses, 35 (50%) had no missing NDI-SC item, 34 (48.6%) had one missing item and one (1.4%) had two missing items. Where missing items were present, the total score was presented as a percentage by aggregating the total item scores and dividing by the maximum score possible derived from the remaining items. Notably, most of the missing values were from the item related to driving as most of the participants do not drive.

Test-Retest Reliability

Fifty-six participants returned at the second session to complete the questionnaires. The mean duration interval between the first and second session was 13 ± 5 days. 50 participants who scored between -3 and +3 on the GROC were included in the test-retest analysis. ICC values were 0.85, indicating good reliability. The Bland and Altman analysis showed that the mean of the difference was −1.6 ± 11.9 (Figure 2). SEM and MDC for the NDI-SC scores were 4.2 and 11.7, respectively.

Internal Consistency

Cronbach α for the NDI-SC was 0.92 indicating excellent internal consistency. The item-scale correlations between single items and total scores of the NDI-SC were fair to strong with correlation coefficients ranging from 0.27 to 0.72, confirming internal consistency of the NDI-SC (Table 2).

Content Validity

No floor and ceiling effects were observed for the total scores as only three participants (0.5%) had the lowest score and none had the highest score. However, all the items with
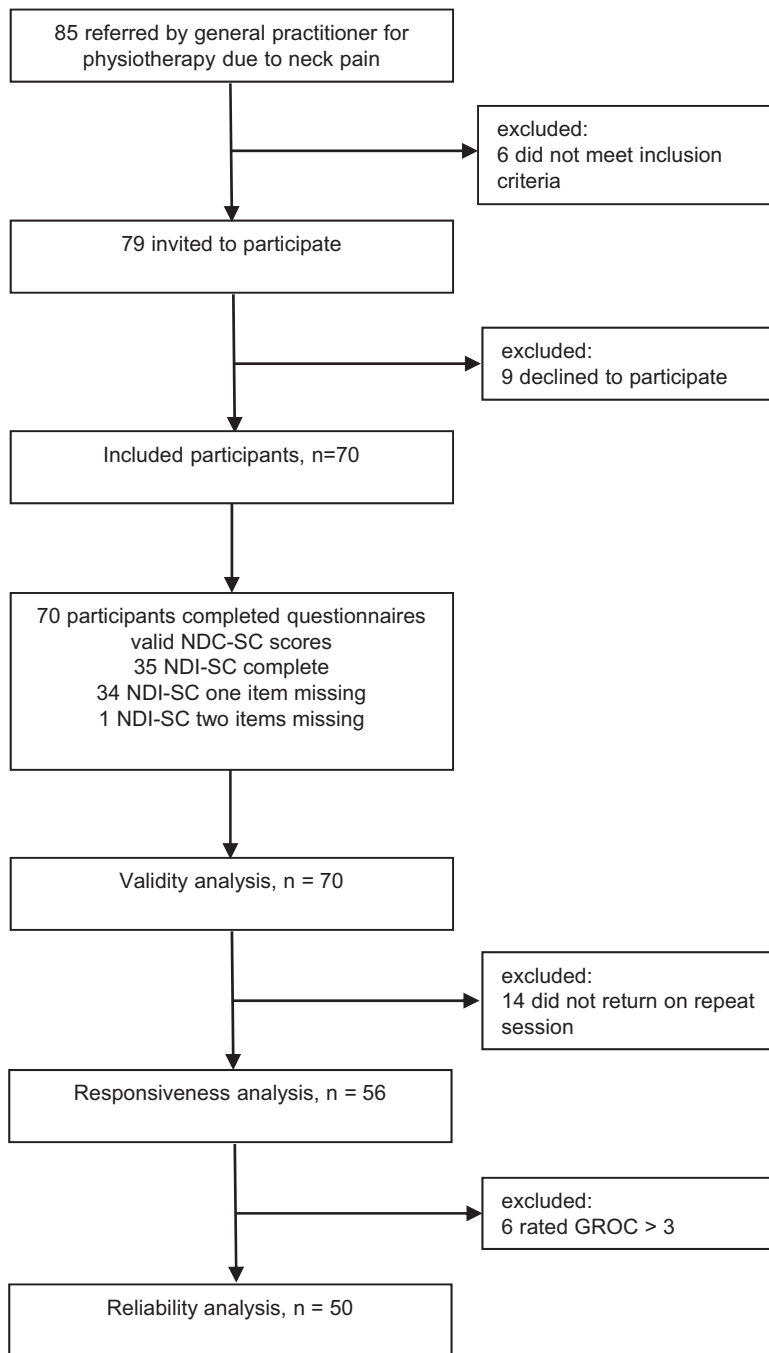
Figure 1. Flowchart of participants.

the exception of "Pain Intensity" had floor effects, with 37.7% of all item entries scoring the lowest possible value. There were no ceiling effects for all of the individual items.

Construct Validity
A strong correlation was found between NDI-SC and VAS ($R_p = 0.61$, $P < 0.001$).

The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy (0.81), and Bartlett test of sphericity (194.25, $P < 0.001$) showed sufficient sample size to conduct a satisfactory factor analysis. A two-factor structure with Eigen values >1 was extracted by factor analysis, which explained

over a total of 66% of the reliable variance of the item scores. The Scree test also showed two factors further confirming internal construct validity (Figure 3). Factor loadings of all the items were similar, showed that rotated component loadings ranged from 0.664 to 0.832 for factor one and 0.507 to 0.905 for factor two, with Item 2 (Personal Care) and Item 5 (Headaches) receiving the highest value in factor one and factor two, respectively (Table 3).

Responsiveness
The NDI-SC showed moderate responsiveness with GROC ($R_s = 0.46$, $P < 0.001$). Table 4 represents the descriptive

| TABLE 1. Participant Characteristics (n = 70)				
	N	N%	Mean	SD
Age, y			44	12
Sex				
Female	37	52.9		
Male	33	47.1		
Education				
Primary	7	10.0		
Secondary	12	17.1		
Tertiary (Junior Colleges/Polytechnics)	22	31.4		
University	20	28.6		
Post-graduate	9	12.9		
Pain duration				
<8 days	10	14.3		
≥8 days to ≤12 wk	19	27.1		
>12 wk	41	58.6		
Previous no. of episodes				
None	24	34.3		
1–10	25	35.7		
>10	21	30.0		
Trauma				
Yes	4	5.7		
No	66	94.3		
VAS (0–10)			4.2	2.4
NDI-SC (0%–100%)			19.1	9.7
NDI-SC indicates Neck Disability Index-Simplified Chinese; VAS, Visual Analogue Scale.				
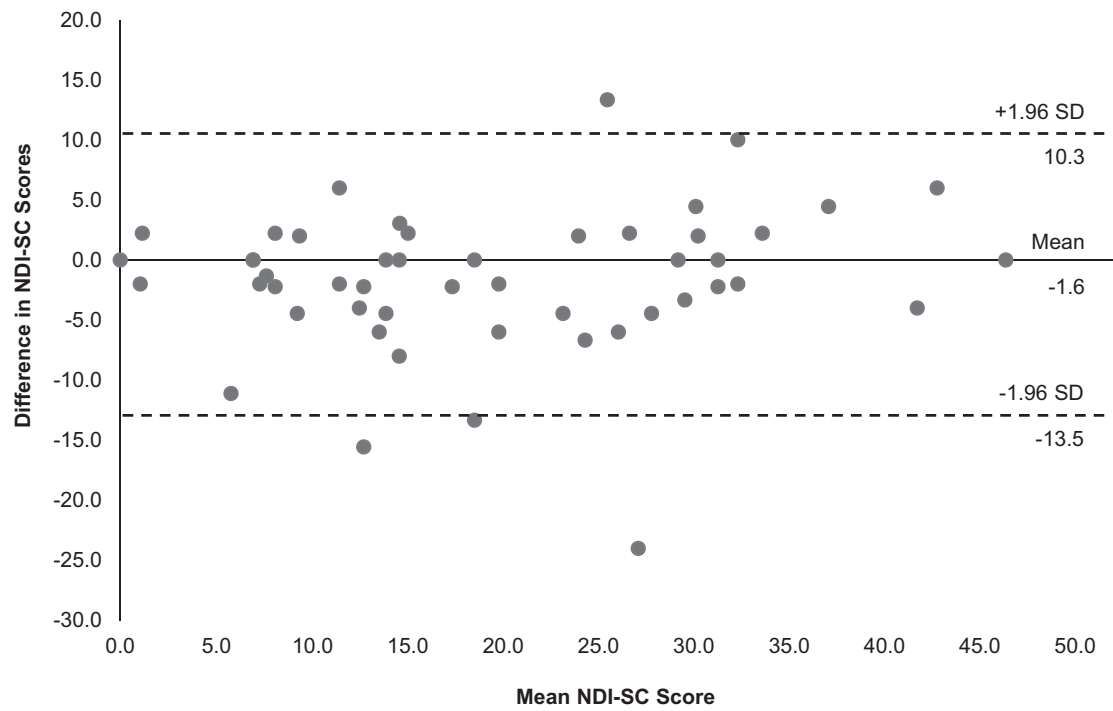
Figure 2. Bland-Altman plot illustrating the test-retest reliability of the NDI-SC. The central line represents the mean difference between test and retest scores, and the outer reference lines represent the 95% limits of agreement. NDI-SC indicates Neck Disability Index into the simplified-Chinese language.

TABLE 2. Test-Retest Reliability, Measurement Errors, and Item-scale Correlations, n = 50

NDI-SC Score (0%–100%)	1st Measurement, mean ± SD	2nd Measurement, Mean ± SD	$ICC_{2,1}$ (95% CI)	SEM	MDC_{95}	Item-scale Correlation
1. Pain Intensity	2.8 ± 1.5	2.7 ± 1.6	0.79 (0.66–0.88)	0.35	1.0	0.44
2. Personal Care	1.1 ± 1.2	1.0 ± 1.4	0.51 (0.27–0.69)	0.43	1.2	0.29
3. Lifting	2.2 ± 2.5	2.1 ± 1.9	0.44 (0.18–0.64)	0.90	2.5	0.52
4. Reading	1.6 ± 1.3	1.7 ± 1.4	0.55 (0.32–0.72)	0.45	1.2	0.59
5. Headaches	2.4 ± 2.6	1.8 ± 2.0	0.67 (0.48–0.80)	0.74	2.0	0.28
6. Concentration	1.2 ± 1.2	1.2 ± 1.2	0.64 (0.44–0.78)	0.36	1.0	0.63
7. Work	1.5 ± 1.5	1.5 ± 1.6	0.73 (0.57–0.84)	0.39	1.1	0.72
8. Driving	0.9 ± 1.0	1.1 ± 1.4	0.61 (0.27–0.82)	0.32	0.9	0.27
9. Sleeping	2.9 ± 2.7	2.5 ± 2.7	0.73 (0.57–0.84)	0.80	2.2	0.66
10. Recreation	2.2 ± 1.8	1.7 ± 1.8	0.65 (0.45–0.79)	0.35	1.0	0.58
Total	19.4 ± 10.9	17.8 ± 11.8	0.85 (0.74–0.91)	4.2	11.7	

CI indicates confidence interval; ICC, intra-class correlation coefficient; MDC, minimal detectable change; NDI-SC, Neck Disability Index-Simplified Chinese; SEM, standard error of measurement.

statistics of the mean change NDI-SC scores according to each GROC grading.

The correlation between NDI-SC change scores and VAS change scores was also moderate ($R_p = 0.59$, $P < 0.001$).

DISCUSSION

The aims of this study were to translate the NDI into the simplified-Chinese version and to evaluate its psychometric properties. The results showed that the NDI-SC had good reliability, validity, and responsiveness in patients with neck pain.

In our present study, the participants had a mean age of 44 ± 12 years which was as expected. Although the variation in age was large, similar studies conducted by Shaheen *et al*,[21] Wu *et al*,[26] and Cramer *et al*[27] also reported having age ranges between 41 ± 10 years, 43 ± 13 years, and 49 ± 16 years, respectively.

Notably, in our present study, the mean NDI-SC score was lower than that of the previous studies,[22,26,27] indicating that participants only had mild disability. This may be attributed to the nature of the primary healthcare setting where participants received care at an earlier onset of neck pain.

The NDI-SC showed good internal consistency with Cronbach α value of 0.92, which is comparable to values reported in earlier studies (0.74–0.97).[16–28] Testing of reliability was also done within a short time interval to minimize changes in participants' condition. The $ICC_{2,1}$ values were found to be 0.85, indicating good reliability and is in line with other studies (0.81–0.92).[16–28] For
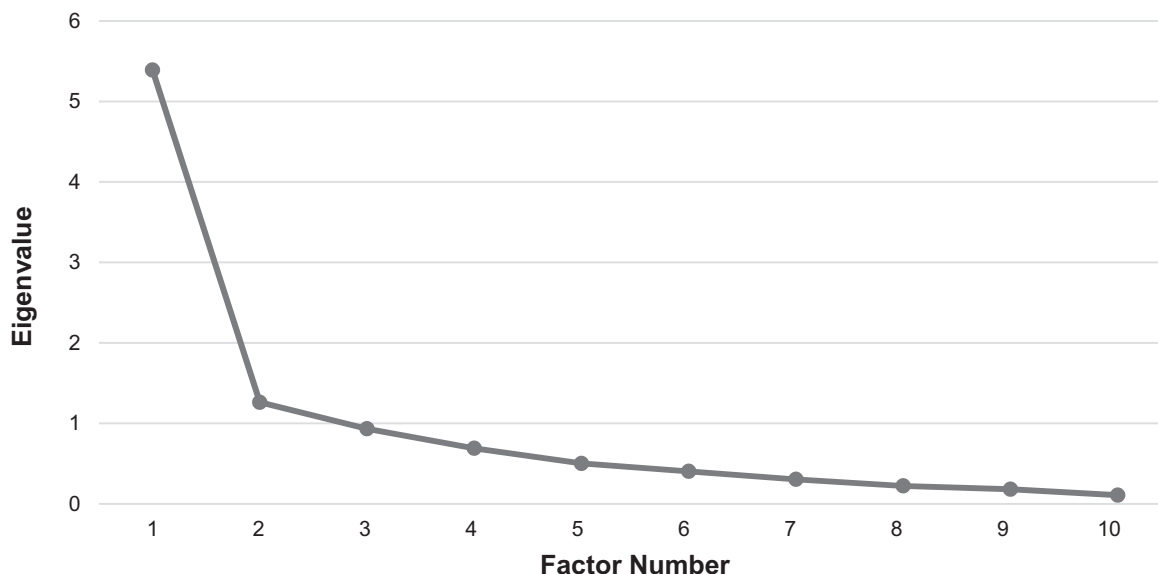
Scree Plot

Figure 3. Scree plot for the 10 items of the simplified-Chinese Neck Disability Index.

TABLE 3. Exploratory Factor Analysis for the NDI-SC*

NDI-SC Items	Factor 1	Factor 2
1. Pain Intensity	0.774	
2. Personal Care	0.832	
3. Lifting	0.664	
4. Reading	0.760	
5. Headaches		0.905
6. Concentration	0.724	
7. Work	0.710	
8. Driving	0.736	0.507
9. Sleeping		0.653
10. Recreation	0.776	

NDI-SC indicates Neck Disability Index-Simplified Chinese.
*Factor loadings of <0.5 were suppressed.

previous studies that reported higher ICC values, we were unable to ascertain the ICC modeling that were used.[23,24] The results of SEM and MDC were also similar to those of other NDI versions.[20,24] The MDC value of the NDI-SC in this study is higher than that reported by Bakhtadze et al,[16] but comparable to values reported by other studies.[24,30]

Given the embrace of a car-lite society in Singapore, there were considerably more missing "driving" responses. This finding was not a translation issue and was in line with previous studies.[23,25,26] Similar to previous studies, the present study did not find any floor or ceiling effects for the NDI-SC total scores.[23,28] However, with the only exception of "Pain Intensity", all other individual components of the NDI-SC were observed to have floor effects. Both of these phenomena were not unique and were also consistent with previous studies reported.[20,22,23,28]

Factor analysis revealed a two-factor structure related to "Pain and Disability" and "Brain Processing and Function" that explained 66% of the total variance. The percentage of the variance was comparable with previous studies that also reported having a two-factor structure.[19,20–22,25] Some controversies exist in relation to the factorial structure of the NDI as some studies revealed a one-factor structure for

TABLE 4. Mean Score Changes for the NDI-SC, According to GROC Grading, n = 56

GROC Score	N (N%)	NDI-SC Mean Change ± SD
≥−3 and <0	2 (1.0%)	−1.0 ± 1.4
0	16 (28.6%)	−0.9 ± 4.2
+1	16 (28.6%)	1.7 ± 8.0
+2	8 (14.3%)	6.4 ± 5.4
+3	8 (14.3%)	2.5 ± 3.7
>3	6 (13.2%)	9.6 ± 13.0

GROC indicates Global Rating of Change; NDI-SC, Neck Disability Index-Simplified Chinese.

the NDI instead.[17,23,24,27] However, the discrepancies found in the factor structure of the current study as compared to other studies may be attributed to the influence of cultural differences.[45]

The correlation between the NDI-SC and VAS was significantly strong ($R_p = 0.61$). This is in accordance to previous studies that found similar values (0.48–0.75) between the NDI and VAS.[22,24–26] Regarding responsiveness, a significant and strong correlation was also observed between NDI-SC change scores and GROC values ($R_s = 0.46$). The NDI-SC also showed a strong significant correlation between its change scores and VAS change scores ($R_p = 0.59$), which agreed with the results of earlier studies.[26–28] Where a longer test-retest interval may reflect a greater responsiveness in the NDI-SC over time, our choice interval duration was consistent with those of previous studies.[21,46]

LIMITATIONS

Criterion validity was not assessed in this study as we did not use an alternative criterion standard for the health-related questionnaire. Furthermore, only the VAS and GROC were used to compare with the NDI-SC. Although the GROC is commonly used due to it being quick and simple to use, the degree of severity of the patient's condition at the time of scoring may influence the scoring. Patients with lower symptom severity are likely to score a greater positive change of the GROC on their retest session.[35] Another limitation of the study was the relatively smaller sample size. However, the KMO value of 0.81 suggested that the study had adequate statistical power for the psychometric evaluation of NDI-SC.

CONCLUSION

The NDI has been successfully cross-culturally adapted and translated into the simplified-Chinese version. The NDI-SC is shown to be a reliable, valid, and responsive measurement tool of pain and functional limitation in simplified-Chinese speaking patients with neck pain.

➤ Key Points

❑ The NDI was translated into the simplified-Chinese language and culturally adapted for Chinese-speaking patients with neck pain.

❑ The NDI-SC demonstrated an excellent level of internal consistency, strong test-retest reliability, content and construct validity, two-factor subscales, and responsiveness.

❑ The NDI-SC is a reliable, valid, and responsive instrument to measure functional limitations in patients with neck pain.

References

1. Fejer R, Kyvik KO, Hartvigsen J. The prevalence of neck pain in the world population: a systematic critical review of the literature. *Eur Spine J* 2006;15:834–48.
2. Hoy DG, Protani M, De R, et al. The epidemiology of neck pain. *Best Pract Res Clin Rheumatol* 2010;24:783–92.
3. Cohen SP. Epidemiology, diagnosis, and treatment of neck pain.- *Mayo Clinic Proceedings*. Elsevier; 2015; 284-299 In.
4. Ferrari R, Russell AS. Neck pain. *Best Pract Res Clin Rheumatol* 2003;17:57–70.
5. Praveen JJ, Lim TJ, O'Brien A. Neck pain in Changi General Hospital: an observational study. *Proc Singapore Healthc* 2014; 23:209–17.
6. Harvey AM. Classification of chronic pain—descriptions of chronic pain syndromes and definitions of pain terms. *Clin J Pain* 1995;11:163.
7. Côté P, van der Velde G, Cassidy JD, et al. The burden and determinants of neck pain in workers: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. *J Manipulative Physiol Ther* 2009;32:S70–86.
8. Lee H, Hübscher M, Moseley GL, et al. How does pain lead to disability? A systematic review and meta-analysis of mediation studies in people with back and neck pain. *Pain* 2015;156:988–97.
9. van Randeraad-van der Zee CH, Beurskens AJHM, Swinkels RAHM, et al. The burden of neck pain: its meaning for persons with neck pain and healthcare providers, explored by concept mapping. *Qual Life Res* 2016;25:1219–25.
10. Walton DM, Balsor B, Etruw E. Exploring the causes of neck pain and disability as perceived by those who experience the condition: a mixed-methods study. *ISRN Rehabil* 2012.
11. Jackowski D, Guyatt G. A guide to health measurement. *Clin Orthop Relat Res* 2003;413:80–9.
12. Bussieres AE, Stewart G, Al-Zoubi F, et al. The treatment of neck pain-associated disorders and whiplash-associated disorders: a clinical practice guideline. *J Manipulative Physiol Ther* 2016;39: 523–64.
13. Vernon H. The Neck Disability Index: state-of-the-art. *J Manipulative Physiol Ther* 2008;31:491–502.
14. MacDermid JC, Walton DM, Avery S, et al. Measurement properties of the neck disability index: a systematic review. *J Orthop Sport Phys Ther* 2009;39:400–17.
15. Vernon H, Mior S. The Neck Disability Index: a study of reliability and validity. *J Manipulative Physiol Ther* 1991;14:409–15.
16. Bakhtadze MA, Vernon H, Zakharova OB, et al. The Neck Disability Index-Russian Language Version (NDI-RU). *Spine (Phila Pa 1976)* 2015;40:1115–21.
17. Cook C, Richardson JK, Braga L, et al. Cross-cultural adaptation and validation of the Brazilian Portuguese version of the Neck Disability Index and Neck Pain and Disability Scale. *Spine (Phila Pa 1976)* 2006;31:1621–7.
18. Joseph SD, Bellare B, Vernon H. Cultural adaptation, reliability, and validity of Neck Disability Index in Indian rural population: a Marathi version study. *Spine (Phila Pa 1976)* 2015;40:E68–76.
19. Kaka BO, Ogwumike O, Vernon H, et al. Cross-cultural adaptation, validity and reliability of the Hausa version of the Neck Disability Index questionnaire. *Int J Ther Rehabil* 2016;23:380–5.
20. Nakamaru K, Vernon H, Aizawa J, et al. Crosscultural adaptation, reliability, and validity of the Japanese version of the neck disability index. *Spine (Phila Pa 1976)* 2012;37:E1343–7.
21. Shaheen AAM, Omar MTA, Vernon H. Cross-cultural adaptation, reliability, and validity of the Arabic version of neck disability index in patients with neck pain. *Spine (Phila Pa 1976)* 2013; 38:E609–15.
22. Shashua A, Geva Y, Levran I. Translation, Validation, and Cross-cultural adaptation of the Hebrew version of the neck disability index. *Spine (Phila Pa 1976)* 2016;41:1036–40.
23. Trouli MN, Vernon HT, Kakavelakis KN, et al. Translation of the Neck Disability Index and validation of the Greek version in a sample of neck pain patients. *BMC Musculoskelet Disord* 2008; 9:106.
24. Uthaikhup S, Paungmali A, Pirunsan U. Validation of Thai versions of the Neck Disability Index and Neck Pain and Disability Scale in patients with neck pain. *Spine (Phila Pa 1976)* 2011; 36:E1415–21.
25. Wlodyka-Demaille S, Poiraudeau S, Catanzariti J-F, et al. French translation and validation of 3 functional disability scales for neck pain. *Arch Phys Med Rehabil* 2002;83:376–82.
26. Wu S, Ma C, Mai M, et al. Translation and validation study of Chinese versions of the neck disability index and the neck pain and disability scale. *Spine (Phila Pa 1976)* 2010;35:1575–9.
27. Cramer H, Lauche R, Langhorst J, et al. Validation of the German version of the Neck Disability Index (NDI). *BMC Musculoskelet Disord* 2014;15:91.
28. Farooq MN, Mohseni-Bandpei MA, Gilani SA, et al. Urdu version of the neck disability index: a reliability and validity study. *BMC Musculoskelet Disord* 2017;18:149.
29. Fairbank JC, Couper J, Davies JB, et al. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980;66:271–3.
30. Cleland JA, Childs JD, Whitman JM. Psychometric properties of the Neck Disability Index and Numeric Pain Rating Scale in patients with mechanical neck pain. *Arch Phys Med Rehabil* 2008;89:69–74.
31. Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)* 2000;25:3186–91.
32. Huskisson EC. Measurement of pain. *Lancet* 1974;304:1127–31.
33. Bijur PE, Silver W, Gallagher EJ. Reliability of the visual analog scale for measurement of acute pain. *Acad Emerg Med* 2001;8:1153–7.
34. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. *J Clin Nurs* 2005;14:798–804.
35. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther* 2009;17:163–70.
36. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures statistics and strategies for evaluation. *Control Clin Trials* 1991;12:S142–58.
37. Schellingerhout JM, Heymans MW, Verhagen AP, et al. Measurement properties of translated versions of neck-specific questionnaires: a systematic review. *BMC Med Res Methodol* 2011;11:87.
38. DeVellis RF. *Scale Development: Theory and Applications*. Sage publications; 2016.
39. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86:420.
40. Bland JM, Altman DG. Precision of test methods 1: guide for the determination and reproducibility for a standard test method (BS 597, part 1). London, United Kingdom. *Br Stand Institutions*.
41. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med* 1998;17:101–10.
42. Bowling A. *Research Methods in Health: Investigating Health and Health Services*. UK: McGraw-hill education; 2014.
43. Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21:651–7.
44. Portney LG, Watkins MP. *Foundations of clinical research: applications to practice*. Upper Saddle River, NJ: Pearson/Prentice Hall; 2009.
45. Pickering PM, Osmotherly PG, Attia JR, et al. An examination of outcome measures for pain and dysfunction in the cervical spine: a factor analysis. *Spine (Phila Pa 1976)* 2011;36:581–8.
46. Young BA, Walker MJ, Strunce JB, et al. Responsiveness of the Neck Disability Index in patients with mechanical neck disorders. *Spine J* 2009;9:802–8.