Original Article / Özgün Araştırma

Northwick Park Neck Pain Questionnaire: Reliability and Validity of a Hindi Adaptation for Patients with Non-inflammatory Neck Pain

Northwick Park Boyun Ağrısı Anketi: Enflamatuvar Olmayan Boyun Ağrısı Olan Hastalarda Hintçe Adaptasyonun Güvenilirliği ve Geçerliliği

 Swati Kumari, Irshad Ahmad, Ankita Sharma, Moattar Raza Rizvi

Manav Rachna International Institute and Studies (MRIIRS), School of Allied Health Sciences, Department of Physiotherapy, Faridabad, India

Abstract

Objective: To examine the psychometric properties of the Northwick Park neck pain questionnaire (NPQ) by translating it from English to Hindi and evaluating its content validity (CV) and internal consistency to assess its feasibility for assessing neck pain.

Materials and Methods: This cross-sectional study included 60 patients who had been experiencing non-inflammatory neck pain for at least 4 months. The psychometric properties, internal consistency, and test–retest reliability of the back-translated version were assessed and compared with the original English version. A panel of six experts validated the instrument for CV.

Results: The factor analysis revealed a three-factor structure, which accounted for 76.7% of variance in the translated Hindi version of the NPQ. Internal consistency was indicated by a Cronbach's alpha of 0.936, and test–retest reliability was assessed using an intraclass correlation coefficient of 0.935 (95% confidence interval: 0.906–0.958). Significant correlations were found between the individual items of the translated Hindi and original English NPQs.

Conclusion: The English-to-Hindi-translated NPQ demonstrated a good factor structure and exhibited similar psychometric properties to the original English version. Therefore, it can be recommended for use in neck pain research in India.

Keywords: Northwick Park neck pain questionnaire, Hindi, neck pain, translation, validity

Öz

Amaç: Çalışmanın amacı, Northwick Park ağrı anketi (NPQ) olarak bilinen anketin İngilizce'den Hintçe diline çevrilerek psikometrik özelliklerini incelemek, içerik geçerliliğini ve iç tutarlılığını değerlendirmek ve boyun ağrısını değerlendirmede kullanılabilirliğini araştırmaktır.

Gereç ve Yöntem: Bu kesitsel çalışma, en az 4 ay boyunca non-enflamatuvar boyun ağrısı yaşayan 60 hastayı içermiştir. Geri çevrilen çevirinin psikometrik özellikleri, iç tutarlılık ve test-tekrar test güvenilirliği değerlendirilerek orijinal İngilizce versiyonuyla karşılaştırılmıştır. İçerik geçerliliği açısından altı kişilik uzman panel tarafından onaylanmıştır.

Bulgular: Faktör analizi, çevrilen Hintçe NPQ varyansının %76,7'sini açıklayan üç faktörlü bir yapıyı ortaya çıkardı. İç tutarlılık, 0,936 Cronbach alfa değeri ile gösterildi ve test-tekrar test güvenilirliği, 0,935 (%95 güven aralığı: 0,906–0,958) aralığında sınıf içi korelasyon katsayısı kullanılarak değerlendirildi. Hintçe'ye çevrilen NPQ'nun bireysel maddeleri ile orijinal İngilizce NPQ arasında anlamlı korelasyonlar bulundu.

Sonuç: Hintçeye çevrilen İngilizce NPQ, iyi bir faktör yapısı sergiledi ve orijinal NPQ ile benzer psikometrik özelliklere sahipti. Bu nedenle, Hindistan'daki boyun ağrısı araştırmalarında kullanılması önerilebilir.

Anahtar Kelimeler: Northwick Park boyun ağrısı anketi, Hintçe, boyun ağrısı, çeviri, validasyon

Address for Correspondence/Yazışma Adresi: Prof. Moattar Raza Rizvi MD, Manav Rachna International Institute and Studies (MRIIRS),
School of Allied Health Sciences, Department of Physiotherapy, Faridabad, India
Phone: +09312048470 E-mail: rajrizvi@gmail.com ORCID: orcid.org/0000-0001-8424-7163

Introduction

The most common joint affected in adults is the neck, with global prevalence rates ranging from 16.7%–75.1% (1,2). The growing use of cell phones and visual display terminal units and their improper handling can cause or exacerbate neck pain (3). Approximately 48%–78% of undergraduate students have neck and upper extremity complaints (4). In many nations, neck pain is the leading cause of morbidity and disability in daily life activities and activities in the workplace. However, it is still unclear how exactly neck pain causes disability, despite its profound impact on daily activities and healthcare services. In fact, healthcare professionals can occasionally experience neck pain throughout their employment. According to a systematic review of work-related musculoskeletal disorders, 35%–45% of medical professionals, including nurses, midwives, and doctors, experience neck, shoulder, and upper back pain (5). The frequency among dental professionals ranges from 50% to 93%, with shoulders being the most commonly affected area (6). Neck pain in this population can have severe consequences, including reduced productivity, limitations in everyday activities, disrupted sleep, and even job loss (5,6,7). More than half of those who suffer from chronic neck pain also experience neuropathic or neuropathic-nociceptive symptoms (8). There are two types of neck discomfort, the first of which is non-specific discomfort, which is characterized by a reduced range of motion of the cervical spine, neck muscle spasms, and discomfort caused by pathogenic factors such as stress. Pathological factors or the environment (9) may cause this type of neck pain. The second type of neck discomfort is known as specific pain, which is caused by damage to specific soft neck tissue. There is evidence to suggest a gradual increase in the incidence of neck pain among adults aged 18 and above over time (10).

The evaluation of neck discomfort is essential in clinical practice, as it contributes to improving the quality of life and functional status of patients. According to the available data, neck pain does not discriminate based on age or gender and can affect individuals of any age. Similar to lower back pain, neck pain often occurs intermittently (11). A recent analysis suggests that neck pain follows a "chronic-episodic course" (12). Chronic neck pain is widely recognized as a complex phenomenon, encompassing various interconnected dimensions; however, the understanding of these specific dimensions and their interrelationships remains incomplete. For example, since pain and disability are believed to be intertwined in many circumstances, it may be advantageous to view pain intensity and disability as a single concept of global pain intensity (13).

An expanded perspective on the state-of-the-art in the medical field may be obtained by evaluating the psychometric characteristics of patient-reported outcome measurements through systematic studies. The emergence of systematic reviews in this field has been facilitated by the development of critical evaluation tools, particularly for psychometric investigations. Researchers and clinicians can benefit from evidence-based recommendations on selecting patient-reported outcome measures when making decisions regarding tool selection (14,15). Information concerning the indication of disease that is not made evident by clinical and laboratory data may only be gathered with the help of validated instruments. Some experts suggest using region-specific instruments, designed specifically for neck pain or disability, which offer higher responsiveness and content validity (CV), while others recommend using generic instruments that encompass a broader range of health status attributes (16,17). With more than 260 million people in over 100 countries using Hindi as their primary language (18), we aimed to test and assess the validity and reliability of the Hindi translation of the Northwick Park neck pain questionnaire (NPQ) for neck discomfort.

Materials and Methods

Study Design

Patients were selected using a convenience sampling method, whereby participants were chosen based on their accessibility and willingness to participate in the study. Prior to completing the questionnaire, all selected participants were required to provide written informed consent for inclusion in the study.

Participants

The study population comprised 60 patients who had been experiencing neck pain for a duration of 4 months. The selection of participants was based on predefined inclusion and exclusion criteria. The sample size was determined based on previous studies recommending a minimum of 50 participants for calculating the inter-rater reliability of the assessment tool (19,20). The detailed procedure was explained to the participants over the phone, and online consent was obtained. The study utilized an online cross-sectional survey conducted through "Google Forms". The participants provided demographic information and details on posture, pain, occupation, extra weight, and muscle strain. They also completed the NPQ. Questionnaires were excluded if portions were completed incorrectly, or more than one portion was omitted.

Ethical Issues

This study adhered to the ethical guidelines for medical research involving human participants following the World Medical Association's Declaration of Helsinki (2013). Approval was obtained from the Manav Rachna International Institute of Research and Studies, Faculty of Allied Health Sciences Research Ethics Committee (reference number: MRIIRS/FAHS/DEC/2021-BPT71). The participants were provided with an information sheet and consent form, which they were asked to read and sign. The participants' identities were kept anonymous and data confidentiality was maintained.

Inclusion/Exclusion Criteria

The inclusion criteria for the study were patients aged 18–20 years who subjectively reported neck pain and had proficiency in reading and interpreting the Hindi language. Patients with inflammatory types of arthritis were excluded. Furthermore, patients with a history of cervical spine injury or surgery, active infection in the cervical spine, other musculoskeletal problems, neck pain caused by another illness (such as a tumor, neurological disease, or vascular disease), radiculopathy with neurological deficit, infection in the cervical spine, or previous diagnosis of a mental disorder were also excluded. These exclusion criteria were confirmed via a thorough medical history review, as well as physical and X-ray examinations.

Questionnaire

The NPQ is a measurement tool used to assess the intensity of neck pain and any resulting disability in performing daily living activities. It evaluates how the neck pain limits the patient's ability to carry out various activities, such as carrying heavy objects, reading, watching television, working, socializing, and driving (optional). The NPQ includes items related to pain intensity, symptom duration, nighttime pins-and-needles or numbness, and sleep disturbance, as well as the impact on social life, carrying, reading/watching TV, working/housework, and driving. Each question is scored on a scale of 0–4, with 0 indicating no disability and 4 indicating the highest level of disability. The scores are summed to calculate a percentage representing the overall disability. Section 10 of the questionnaire assesses changes in pain during follow-up and is not included in the final NPQ score (21,22). The NPQ was adapted from the Oswestry questionnaire, which is commonly used to assess lower back pain.

Translation

The process of translating the grading scales into Hindi involved a three-way round of forward and backward translation. Two separate multilingual translation experts independently adapted the original NPQ instrument, including the instructions, items, and answer choices, into Hindi. The translated versions were then reviewed by experts to assess any differences in meaning, both linguistically and conceptually. Through consensus, a single unified version of the Hindi NPQ was agreed upon. This rigorous translation process ensured that the Hindi version accurately reflected the intended meaning of the original instrument.

To ensure the accuracy of the translated version, two additional translators independently performed back-translation from Hindi to English. A committee member who was fluent in both English and Hindi reviewed the original NPQ and the back-translations. The translators were not familiar with the questionnaire they were adapting, nor were they made aware of the goals of the study. Two separate groups, each comprising three individuals, evaluated the back-translated versions. They used a four-point Likert scale ranging from 1 (significantly altered) to 4 (not significantly altered) to assess whether there were any significant differences in the meaning of the instructions, items, and answer choices compared with the original instrument (which remained unaltered). Based on the evaluations of the back-translations, all sections of the scale were revised and modified as necessary to ensure the originality of the instrument. The final Hindi version incorporated the feedback obtained from the evaluation of the back-translations to enhance its accuracy and fidelity.

The face validity of the instrument was assessed to determine whether the items, instructions, and response scale were easily understandable to the targeted demographic. The Hindi version of the NPQ was made available online to a sample of 15 participants. These participants provided feedback on the overall readability of the survey, as well as specific comments and suggestions for improvement on each topic. To cater to the subsequent study participants, the Hindi version of the NPQ underwent meticulous refinement, integrating feedback from sample participants. This ensured its clarity and suitability for the intended target population.

Assessment of Validity

This study employed empirical techniques to assess the CV, including the calculation of the CV index (CVI), the CV ratio (CVR), and semi-structured cognitive evaluations (23). The following empirical techniques were employed to validate and assess the reliability of this tool.

Content Validity

The item-CVI (I-CVI) and scale-level-CVI (S-CVI) were used to assess the CV during the development of the instrument (24). The I-CVI calculated the proportion among the total number of experts who rated each item as "very relevant". The I-CVI scores ranged from 0–1, where values above 0.79 indicated item relevance, 0.70 suggested the need for item modifications, and values below 0.70 suggested item deletion. Conversely, the S-CVI evaluated the proportion of items in the tool that were rated as "extremely relevant". Two approaches were used to calculate the S-CVI: universal agreement (S-CVI/UA) and average CVI (S-CVI/Ave). The S-CVI/UA reflected the level of agreement among the experts, while the S-CVI/Ave served as a less conservative method of calculation (23).

The S-CVI/UA was calculated by dividing the sum of all items with an I-CVI equal to 1 by the total number of items in the scale, while the S-CVI/Ave was computed by dividing the sum of all I-CVI scores by the total number of items in the scale (16). Excellent CV was demonstrated when the S-CVI/UA was equal to or greater than 0.8 and the S-CVI/Ave was equal to or greater than 0.9 (23). These thresholds indicated a high level of agreement among experts and overall item relevance, respectively.

Content Validity Ratio

The CVR was another empirical analysis technique used to assess the importance of each item in the NPQ. The CVR values ranged from -1 to +1, with a higher value indicating a higher level of consensus among the panel of experts. The formula for calculating the CVR was $(Ne - N/2)/(N/2)$, where Ne is the number of respondents who marked a response option as "important", and N is the total number of respondents (23).

Factor Analysis

An evaluation of the NPQ's factor structure was conducted using principal component analysis. The adequacy of the sample size and the validity of the factor analysis were assessed using the Kaiser–Meyer–Olkin (KMO) measure of sample adequacy and the Bartlett test of sphericity. Principal axis factoring with Varimax rotation was determined to be the most effective approach for extracting the desired information from the data. To determine the number of loaded factors, criteria such as eigenvalues ≥1 (Kaiser's criterion) and a scree plot were utilized. These methods helped to identify the appropriate number of factors in the NPQ (25,26).

Assessment of Reliability

Internal Consistency and Test–retest Reliability

In the reliability investigations, two methods were employed to assess the reliability of the NPQ: internal consistency and test–retest reliability.

Internal Consistency

Cronbach's alpha (α) coefficient and the total item correlation coefficients were examined to assess the internal consistency. A Cronbach's α of 0.7–0.9 is generally considered indicative of good internal consistency (27). If the overall correlation coefficient of an item is >0.2, the item is considered acceptable (28).

Test–retest Reliability

To measure test–retest reliability, all participants completed the translated Hindi version of the NPQ twice (initially and after 48 hours), without any therapy to minimize the retention influence on the responses (29). The intraclass correlation coefficient (ICC) and corresponding 95% confidence interval were used to assess reproducibility (test–retest reliability), with the reliability deemed adequate if the values were 0.50–0.75, good if the values were 0.75–0.90, and excellent if the values were >0.90 (30). For the test–retest reliability, the Two-Way random ICC model (21) was estimated.

For the Hindi translation of the NPQ, the following equation was used to determine the standard error of measurement (SEM) as an indicator of absolute reliability:

$$SEM = SD \times \sqrt{(1-ICC)}$$

Furthermore, minimum detectable change (MDC$_{95}$) is a statistical estimate of the minimal detectable change that may be indicated by a measure that refers to a noticeable change in ability, which was determined using the following expression (31):

$$MDC_{95} = SEM \times 1.96\sqrt{2}$$

Statistical Analysis

Data analysis was conducted using Microsoft Excel 2016 for Windows to manage the subject information, which was then imported into SPSS Statistics software (SPSS v.23; IBM Corp, Armonk, NY, USA) on a Windows platform for statistical analysis. Quantitative data were provided for all scales. Prior to the analysis, all responses were reviewed for consistency, accuracy, and completeness, and for the possible presence/absence of multiple or inconsistent responses. The normal distribution of data was assessed using the Kolmogorov–Smirnov (K–S) test. Paired t-tests were used to compare the results of the test and retest. All statistical tests were two-tailed, and a P value of <0.05 was considered to reflect statistical significance.

Results

Cultural Adaptation

The outcomes of the two forward translations were relatively comparable, requiring only minor adjustments to reach a consensus on the synthesized version. The back-translations were found to be identical to the original instrument documents in terms of their respective meanings. When evaluating the back translation synthesis in comparison with the original instrument, no necessary modifications were identified by the authors of the present study. In the pre-test, all respondents rated the questionnaire as "easy to understand", and there were only a few minor comments regarding word order and the substitution of a phrase with a synonym.

Assessment of Validity

All CV (CVI and CVR) calculations were performed for the translated Hindi version of the NPQ, which comprised 10 items.

The Item-content Validity Index Results (Relevancy of Individual Items)

Table 1 presents the calculated I-CVI weights for each item, indicating their relevancy. Out of 10 items, five were marked as excellent with I-CVI values ranging from 0.76 to 1.00. Five items were marked as fair, all having an I-CVI value of 0.42. One item (Q10) was marked as poor with an I-CVI of -0.14 and was subsequently removed.

Table 1. Content validity of the back-translated version of the Northwick Park neck pain questionnaire

	Content validity index							Content validity ratio		
	Experts	Agree	I-CVI	UA	pc	k	Interpretation	Experts	Agree	CVR
Q1	5	5	1.00	1	0.03	1.00	Excellent	7	6	0.71
Q2	5	3	0.60	0	0.31	0.42	Fair	7	5	0.43
Q3	5	5	1.00	1	0.03	1.00	Excellent	7	6	0.71
Q4	5	3	0.60	0	0.31	0.42	Fair	7	5	0.43
Q5	5	5	1.00	1	0.03	1.00	Excellent	7	6	0.71
Q6	5	5	1.00	1	0.03	1.00	Excellent	7	6	0.71
Q7	5	3	0.60	0	0.31	0.42	Fair	7	5	0.43
Q8	5	3	0.60	0	0.31	0.42	Fair	7	5	0.43
Q9	5	3	0.60	0	0.31	0.42	Fair	7	5	0.43
Q10	5	1	0.20	0	0.16	0.05	Poor	7	3	-0.14
Proportion relevance			0.72	0.40		0.61			Average	0.49
			S-CVI/Ave	S-CVI/UA						

CVI: Content validity index, I-CVI: Index of CVI, S-CVI: Scale level CVI, UA: Universal agreement, pc: Percent; k: Kappa value, Ave: Average

Scale-level Content Validity Index Results (Relevancy of the Overall Questionnaire)

The S-CVI/UA was 0.72, while the S-CVI/average was 0.40. The UA was obtained by summing all I-CVIs equal to 1.00 (five items) and dividing the result by 10, while the average was calculated by summing all I-CVIs and again dividing the result by 10. Overall, the UA approach demonstrated intermediate CV for the NPQ, whereas the average approach exhibited excellent CV.

Kappa

Although CVI is often used to measure the CV, Shi et al. (24) argued that this index does not account for the risk of chance agreement and recommended the use of a kappa statistic in addition to CVI. The kappa statistic provides a level of agreement that would be expected by chance only. The kappa values for the Hindi-translated version of the NPQ ranged from 0.42 to 1.0, indicating good agreement.

Content Validity Ratio Results

The CVR was calculated for each item of the questionnaire. The CVR for non-essential items was <0.99 (based on the total number of experts, N = 6). The elimination of non-essential items was possible; however, in this case, they were not eliminated. Four items had a CVR of 0.71, five items had a score of 0.43, and one item had a score of -0.14. The average CVR value was 0.49.

Factor Structure and Reliability of the Northwick Park Neck Pain Questionnaire

One K–S test sample indicated the presence of normality (P = 0.07). The KMO measure of sample adequacy for the Hindi version of the NPQ was 0.847, exceeding the cutoff value of >0.60. Based on the chi-squared statistical distribution, Bartlett's sphericity test for the translated Hindi version yielded a result of 452 with a significance level below 0.001, indicating that the data were suitable for factor analysis.

Principal axis factoring with a Varimax rotation solution was performed, resulting in a three-factor solution for the Hindi version of the NPQ. A scree plot was also used to support this three-factor solution. These three factors accounted for 76.7% of the variance corresponding to the originally defined dimensions (Table 2).

Assessment of Reliability

Homogeneity

For the overall reliability of the total scale, an excellent Cronbach's α of 0.936 was observed for the translated Hindi version of the NPQ (Table 3). Cronbach's α (if an item was deleted) ranged from 0.92 to 0.94, indicating excellent internal consistency.

All of the correlations between the items were positive and statistically different from 0, indicating that a scale may be constructed using these items since they all measured the same attribute. For the translated Hindi version of the NPQ, correlations between item total scores were found to be significant ($P < 0.001$), with Spearman coefficients ranging from 0.686 to 0.909. The data revealed that the back-translated NPQ had excellent homogeneity and internal consistency.

Reproducibility

The time interval between the test and retest was 48 hours. The ICC of test–retest reliability was fair for the translated Hindi version of the NPQ and reached 0.83 (Table 4). The SEM and MDC_{95} values for the NPQ are summarized in Table 4.

Table 5 shows the relationship between the test and retest scores of the translated Hindi version of the NPQ. The paired t-test verified that there was no significant difference between the tests or the retest scores for Q1, Q2, Q3, Q5, and Q6; however, there was a significant difference between the tests and retest scores for Q4, Q7, Q8, and Q9.

Table 6 illustrates the validity of the translated Hindi version as determined by Spearman's correlation coefficient between the individual items of the original English and the translated Hindi version of the NPQ. Each item in the original English NPQ was found to be strongly correlated with the respective item in the back-translated NPQ ($P < 0.001$).

Discussion

The results of this study demonstrated that the Hindi translation of the NPQ is a valid and reliable tool for assessing the severity of pain in individuals with mechanical neck discomfort. Neck pain significantly impacts a person's daily functioning and overall well-being (15). Nearly all respondents (98%) completed the questionnaire, and no items were reported as unclear, indicating good acceptance of the Hindi translation. Moreover, no floor or ceiling effects were observed in the distribution of NPQ scores in the Hindi version. Based on our findings, the Hindi NPQ was well understood by the participants, was completed efficiently, and was proven to be highly valuable in therapeutic settings. The literature suggests that various translations of the NPQ exhibited high levels of internal consistency (Cronbach's α range: 0.83–0.88). Cronbach's α for the Hindi version was 0.94, which was higher than the values found in the other translations (Chinese: 0.87; Korean: 0.88; modified Chinese version: 0.88) (17,22). Notably, the Hindi translation of the NPQ maintained a range of values of 0.79–0.90, indicating good internal consistency, even after item deletion.

The individual item mean scores were found to correspond with pain severity in a sample of patients with neck discomfort at a UK rheumatology clinic using the NPQ (21). Validity studies have demonstrated the NPQ's sensitivity to change and test–retest reliability (21,32). When examining test–retest reliability, it is important to find a balance between the time it takes patients to forget the test and how quickly they forget the results. Studies assessing the test–retest reliability of NPQ scales reported varying time intervals, ranging from 20–30 min to 7 days. In the present study, the Hindi version of the NPQ was administered to 55 patients with mechanical neck pain. The initial form was completed within one day, and the patients were then asked to complete the form again after a period of 48 hours, without following any exercise protocols. The analysis of the test–retest reliability for the Hindi version of the NPQ revealed an ICC ranging from 0.55 to 0.83, indicating good-to-excellent reproducibility. The value of ICC in the original version showed test–retest values of 0.09–0.72. This indicated moderate reliability and an average consistency level. The ICC value for the translated version ranged from 0.31 to 0.90, indicating good reliability and a satisfactory level of consistency.

Table 2. Psychometric properties of the back-translated version of the Northwick Park neck pain questionnaire				
Bartlett's test of sphericity				
χ²		df		*P*
452		36		<0.001
KMO measure of sampling adequacy				
Overall KMO		**0.847**		
Q1, intensity of neck discomfort		0.847		
Q2, sleeping		0.746		
Q3, numbness		0.805		
Q4, duration		0.866		
Q5, carrying		0.892		
Q6, reading/television		0.893		
Q7, working		0.889		
Q8, social		0.83		
Q9, driving		0.823		

Component loading

	Component			Uniqueness
	1	2	3	
Q1, intensity of neck discomfort	-	0.673	0.4	0.333
Q2, sleeping	0.413	0.835	-	0.103
Q3, numbness	0.379	0.69	0.311	0.284
Q4, duration	0.385	0.368	0.73	0.183
Q5, carrying	0.712	-	-	0.386
Q6, reading/television	0.834	0.363	0.303	0.080
Q7, working	0.549	0.605	0.474	0.109
Q8, social	0.7	0.435	0.442	0.126
Q9, driving	0.53	0.374	0.495	0.495

Note: The principal axis factoring extraction method was used in combination with a varimax rotation

Variance

Component	SS loadings	% of variance	Cumulative %
1	2.79	31	31
2	2.61	29	60
3	1.5	16.7	76.7

Scree plot

χ²: Chi-squre, df: Degree of freedom, *P*: Level of significance, KMO: Kaiser–Meyer–Olkin, SS loadings: Stands for "Sum of Squares Loadings"

Table 3. Cronbach's alpha of individual items of back translated version of Northwick Park neck pain questionnaire

Scale reliability statistics	Mean	SD		Cronbach's α
	1.37	0.597		0.936
Item reliability statistics				
	Mean	SD	Item-rest correlation	If item dropped Cronbach's α
Q1, intensity of neck discomfort	1.37	0.653	0.697	0.932
Q2, sleeping	1.35	0.588	0.799	0.928
Q3, numbness	1.33	0.549	0.767	0.93
Q4, duration	1.33	0.644	0.761	0.929
Q5, carrying	1.5	0.986	0.686	0.938
Q6, reading/television	1.52	0.885	0.871	0.922
Q7, working	1.31	0.668	0.909	0.921
Q8, social	1.24	0.612	0.898	0.923
Q9, driving	1.41	0.88	0.673	0.936

SD: Standard deviation, Q1: Intensity of neck discomfort, Q2: Sleeping, Q3: Numbness, Q4: Duration, Q5: Carrying, Q6: Reading/television, Q7: Working, Q8: Social, Q9: Driving,

Table 4. Two-Way random intraclass correlation coefficient for the Hindi version of the Northwick Park neck pain questionnaire taken at baseline and after 48 hours (test–retest)

Q	Mean (SD)	Mean (SD)	ICC (2.1)	95% CI	α	MDC_{95}	SEM	t	*P*
Q1	1.37 ± 0.65	1.52 ± 0.89	0.66	0.41–0.8	0.93	1.06	0.38	-1.38	0.17
Q2	1.35 ± 0.59	1.54 ± 0.84	0.72	0.52–0.84	0.93	0.86	0.31	-2.02	0.05
Q3	1.33 ± 0.55	1.37 ± 0.65	0.55	0.23¬–0.74	0.93	1.02	0.37	-0.41	0.69
Q4	1.33 ± 0.64	1.59 ± 1.06	0.77	0.60–0.87	0.93	0.86	0.31	-2.52	0.02
Q5	1.50 ± 0.99	1.54 ± 0.97	0.58	0.27–0.76	0.94	1.78	0.64	-0.26	0.80
Q6	1.52 ± 0.89	1.65 ± 0.87	0.83	0.70–0.9	0.92	1.02	0.37	-1.41	0.16
Q7	1.31 ± 0.67	1.50 ± 0.72	0.75	0.58–0.86	0.92	0.92	0.33	-2.21	0.03
Q8	1.24 ± 0.61	1.46 ± 0.77	0.74	0.55–0.85	0.92	0.87	0.31	-2.57	0.01
Q9	1.41 ± 0.88	1.65 ± 1.08	0.66	0.41–0.80	0.94	1.43	0.52	-2.21	0.03

ICC: Interclass correlation, CI: Confidence interval, SEM: Standard error mean, MDC_{95}: Minimum detectable change, Q1: Intensity of neck discomfort, Q2: Sleeping, Q3: Numbness, Q4: Duration, Q5: Carrying, Q6: Reading/television, Q7: Working, Q8: Social, Q9: Driving, SD: Standard deviation

Table 5. Comparison between the scores for individual items in the original and back-translated versions of the Northwick Park neck pain questionnaire

NPQ	Mean difference	SE difference	95% Confidence interval		t	df	*P*
			Lower	Upper			
Q1	-0.06	0.06	-0.17	0.06	-1.00	53	0.32
Q2	-0.13	0.08	-0.29	0.03	-1.63	53	0.11
Q3	-0.04	0.10	-0.23	0.15	-0.39	53	0.70
Q4	-0.04	0.07	-0.19	0.11	-0.50	53	0.62
Q5	-0.19	0.15	-0.49	0.12	-1.24	53	0.22
Q6	-0.04	0.08	-0.20	0.13	-0.44	53	0.66
Q7	0.00	0.08	-0.17	0.17	0.00	53	1.00
Q8	-0.02	0.04	-0.10	0.07	-0.44	53	0.66
Q9	0.11	0.11	-0.11	0.33	1.00	53	0.32

df: Degree of freedom, Q1: Intensity of neck discomfort, Q2: Sleeping, Q3: Numbness, Q4: Duration, Q5: Carrying, Q6: Reading/television, Q7: Working, Q8: Social, Q9: Driving, SE: Standard error, NPQ: Neck pain questionnaire

Table 6. Spearman correlation coefficient between nine items of the original and back-translated version of the Northwick Park neck pain questionnaire

	OV_Q1	OV_Q2	OV_Q3	OV_Q4	OV_Q5	OV_Q6	OV_Q7	OV_Q8	OV_Q9
HV_Q1	0.807***								
HV_Q2		0.686***							
HV_Q3			0.496***						
HV_Q4				0.585***					
HV_Q5					0.51***				
HV_Q6						0.691***			
HV_Q7							0.587***		
HV_Q8								0.826***	
HV_Q9									0.517***

$*P < 0.05$, $**P < 0.01$, $***P < 0.001$, OV: Original version, HV: Hindi version, Q1: Intensity of neck discomfort, Q2: Sleeping, Q3: Numbness, Q4: Duration, Q5: Carrying, Q6: Reading/television, Q7: Working; Q8: Social, Q9: Driving

Finally, the Hindi version revealed ICC values for test–retest of 0.55–0.83, demonstrating good reliability and good consistency.

The present findings also revealed a strong to very strong positive correlation between the original and the Hindi version of the NPQ, with the highest correlation observed in Q8 ($r = 0.826$) and the lowest in Q3 ($r = 0.496$). González et al. (33) conducted a study to validate a Spanish version of the NPQ and assess its usefulness in clinical practice. They found a good intraclass correlation in the test–retest, indicating good agreement; the correlation with the visual analogue scale (VAS) was also strong. To assess pain in Spanish-speaking individuals with persistent neck pain, the NPQ was translated and validated (33). Lee et al. (22) conducted a study to verify the accuracy and validity of the NPQ after translating it into Korean for individuals with neck discomfort. The ICC and Cronbach's α were used to measure the reliability of the tests. The reliability of the Korean NPQ was assessed by correlating it with the VAS reliability. The translated version of the NPQ showed high levels of test–retest reliability as indicated by the ICC. The researchers concluded that the Korean version of the NPQ was a reliable and valid tool for measuring neck pain.

Individual item mean scores were found to correspond with pain severity in a UK rheumatology clinic sample of patients who presented with complaints of neck discomfort using the NPQ. Validity studies have demonstrated the questionnaire's sensitivity to modification and test–retest reliability. Yeung et al. (34) conducted a validation study on the use of a modified version of the NPQ in patients who had experienced impairment after radiation treatment to the neck. The purpose of this research was to analyze the feasibility of using the Chinese Northwick Park Pain Questionnaire for assessing neck impairment in patients who had undergone irradiation. To ensure the accuracy of the information, eight physiotherapists and five patients were surveyed. Construct validity was evaluated by measuring the levels of neck pain, range of motion, maximal isometric neck muscle strength, and health assessment score using the short form 36 (SF-36). In addition to strong test–retest reliability, internal consistency, and concept validity, the questionnaire also exhibited strong CV. Most of the SF-36's component scores and the numeric rating scale were strongly associated with the NPQ. The modified Chinese version of the NPQ was determined to be a valid and reliable tool for measuring neck impairment after irradiation. The purpose of this research was to compare the results of magnetic resonance imaging (MRI) of the cervical spine with self-reported pain and disability levels using the NPQ. Researchers examined the causes of neck pain in 251 randomly selected patients. Patients who had been hospitalized for cervical spine trauma or who had a history of discitis, surgery, or tumors were not included. All patients completed the NPQ and underwent a series of MRI techniques, including sagittal gradient-echo T1 and turbo spin-echo T2, axial gradient-echo T2*, and highly T2-weighted MR myelographic weighted imaging. Two disc levels were found to be the most affected, and an MR imaging score of 0–30 was assigned based on this analysis. In this study, no evidence of a link between NPQ and MRI scores was found. Sleeplessness and numbness were the only NPQ items associated with a higher MR imaging score (34).

A previous study utilized the Chinese version of the NPQ to evaluate the reliability, validity, and responsiveness of the instrument in a sample of Chinese patients in Hong Kong who were experiencing neck pain (35). The study involved 532 consecutive adult patients with neck pain from seven physiotherapy outpatient departments in Hong Kong who completed the NPQ at the start of physiotherapy, 7 days later, 3 weeks later, and 6 weeks later. The Chinese translation of the NPQ demonstrated very good CV, test–retest reliability, and internal consistency.

Study Limitations

This study has some limitations, including the use of a cross-sectional design, largely due to time and recruiting constraints. Additionally, due to the chronic and slow-developing nature of neck fibrosis and the resulting impairments, a significant duration of monitoring may be required to detect changes.

Conclusion

The Hindi-translated version of the NPQ was determined to be a valid tool with adequate reliability and validity for assessing neck pain in the Indian population. The use of this questionnaire in Hindi can assist the general population in understanding the impact of pain on their symptoms and functional activities.

Furthermore, on comparing the translated version with other versions, it was observed that the Hindi version exhibited greater reliability in our population.

Ethics

Ethics Committee Approval: Approval was obtained from the Manav Rachna International Institute of Research and Studies, Faculty of Allied Health Sciences Research Ethics Committee (reference number: MRIIRS/FAHS/DEC/2021-BPT71).

Informed Consent: Consent was obtained.

Peer-review: Internally peer-reviewed.

Authorship Contributions

Concept: S.K., I.A., A.S., M.R.R., Design: S.K., I.A., A.S., M.R.R., Data Collection or Processing: S.K., A.S., Analysis or Interpretation: M.R.R., Literature Search: I.A., Writing: S.K., I.A., A.S., M.R.R.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Vingård E. Chapter 5.6: Major public health problems-musculoskeletal disorders. Scand J Public Health Suppl 2006;67:104-112.

2. 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;15:834-848.

3. Dockrell S, Bennett K, Culleton-Quinn E. Computer use and musculoskeletal symptoms among undergraduate university students. Comput Educ 2015;85:102-109.

4. Waersted M, Hanvold TN, Veiersted KB. Computer work and musculoskeletal disorders of the neck and upper extremity: a systematic review. BMC Musculoskelet Disord 2010;11:79.

5. Long MH, Bogossian FE, Johnston V. The prevalence of work-related neck, shoulder, and upper back musculoskeletal disorders among midwives, nurses, and physicians: a systematic review. Workplace Health Saf 2013;61:223-229.

6. Hayes M, Cockrell D, Smith DR. A systematic review of musculoskeletal disorders among dental professionals. Int J Dent Hyg 2009;7:159-165.

7. Cagnie B, Danneels L, Van Tiggelen D, De Loose V, Cambier D. Individual and work related risk factors for neck pain among office workers: a cross sectional study. Eur Spine J 2007;16:679-686.

8. Liu R, Kurihara C, Tsai HT, et al. Classification and treatment of chronic neck pain: a longitudinal cohort study. Reg Anesth Pain Med 2017;42:52-61.

9. Murphy S, Buckle P, Stubbs D. A cross-sectional study of self-reported back and neck pain among English schoolchildren and associated physical and psychological risk factors. Appl Ergon 2007;38:797-804.

10. G Genebra CVDS, Maciel NM, Bento TPF, Simeão SFAP, Vitta A. Prevalence and factors associated with neck pain: a population-based study. Braz J Phys Ther 2017;21:274-280.

11. Guzman J, Hurwitz EL, Carroll LJ, et al. 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. J Manipulative Physiol Ther 2009;32(2 Suppl):17-28.

12. Hoy DG, Protani M, De R, Buchbinder R. The epidemiology of neck pain. Best Pract Res Clin Rheumatol. The epidemiology of neck pain. Best Pract Res Clin Rheumatol 2010;24:783-792.

13. Saavedra-Hernández M, Castro-Sánchez AM, Cuesta-Vargas AI, et al. The contribution of previous episodes of pain, pain intensity, physical impairment, and pain-related fear to disability in patients with chronic mechanical neck pain. Am J Phys Med Rehabil 2012;91:1070-1076.

14. Bobos P, MacDermid JC, Walton DM, Gross A, Santaguida PL. Patient-Reported Outcome Measures Used for Neck Disorders: An Overview of Systematic Reviews. J Orthop Sports Phys Ther 2018;48:775-788.

15. Bobos P, MacDermid J, Nazari G, Furtado R; CATWAD. Psychometric properties of the global rating of change scales in patients with neck disorders: a systematic review with meta-analysis and meta-regression. BMJ Open 2019;9:e033909.

16. Gondivkar SM, Gadbail AR, Sarode SC, et al. Measurement properties of oral health related patient reported outcome measures in patients with oral cancer: A systematic review using COSMIN checklist. PloS One 2019;14:e0218833.

17. Schellingerhout JM, Verhagen AP, Heymans MW, et al. Measurement properties of disease-specific questionnaires in patients with neck pain: a systematic review. Qual Life Res 2012;21:659-670.

18. Kjellén E. Language recognition and voter turnout: Insights from India. In:202. https://www.diva-portal.org/smash/get/diva2:1623628/FULLTEXT01.pdf

19. Streiner DL, Kottner J. Recommendations for reporting the results of studies of instrument and scale development and testing. J Adv Nurs 2014;70:1970-1979.

20. Johanson GA, Brooks GP. Initial scale development: sample size for pilot studies. Educ Psychol Meas 2010;70:394-400.

21. Leak AM, Cooper J, Dyer S, et al. The Northwick Park Neck Pain Questionnaire, devised to measure neck pain and disability. Br Rheumatol 1994;33:469-474.

22. Lee KW, Seo HD, Jung KS, Kim SH, Chung Y. Reliability and validity of Korean version Northwick Park Neck Pain Questionnaire in neck pain patients. Phys Ther Korea 2010;17:68-76.

23. Rodrigues IB, Adachi JD, Beattie KA, MacDermid JC. Development and validation of a new tool to measure the facilitators, barriers and preferences to exercise in people with osteoporosis. BMC Musculoskelet Disord 2017;18:540.

24. Shi J, Mo X, Sun Z. [Content validity index in scale development]. Zhong Nan Da Xue Xue Bao Yi Xue Ban 2012;37:152-155.

25. Almeida MQG, Dibai-Filho AV, Guirro RRJ, et al. Psychometric properties of the Brazilian short-version of the Northwick Park Neck Pain Questionnaire. Clin Rehabil 2022;36:980-992.

26. Ahmad AA, Akindele MO, Umar AM, et al. The Hausa Northwick Park Neck Pain Questionnaire: translation, cross-cultural adaptation and psychometric assessment in patients with non-specific neck pain. Disabil Rehabil 2023:1-10.

27. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ 2018;48:1273-1296.

28. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi Med J 2012;24:69-71.

29. No authors listed. Health measurement scales: a practical guide to their development and use (5th edition). Aust N Z J Public Health 2016;40:294-295.

30. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155-163.

31. Donoghue D; Physiotherapy Research and Older People (PROP) group; Stokes EK. How much change is true change? The minimum detectable change of the Berg Balance Scale in elderly people. J Rehabil Med 2009;41:343-346.

32. Wlodyka-Demaille S, Poiraudeau S, Catanzariti JF, et al. French translation and validation of 3 functional disability scales for neck pain. Arch Phys Med Rehabil 2002;83:376-382.

33. González T, Balsa A, Sáinz de Murieta J, et al. Spanish version of the Northwick Park Neck Pain Questionnaire: reliability and validity. Clin Exp Rheumatol 2001;19:41-46.

34. Yeung PL, Chiu TT, Leung AS. Use of modified Northwick Park Neck Pain Questionnaire in patients with postirradiation neck disability: validation study. Head Neck 2004;26:1031-1037.

35. Chiu TT, Lam TH, Hedley AJ. Subjective health measure used on Chinese patients with neck pain in Hong Kong. Spine (Phila Pa 1976) 2001;26:1884-1889.