






# Translation, reliability, and validity of the Turkish version of the Neck Bournemouth Questionnaire

Onur Yılmaz<sup>1</sup> , Ümit Gafuroğlu<sup>1</sup> , Selcen Yüksel<sup>2</sup> 

<sup>1</sup>Department of Physical Medicine and Rehabilitation, Ankara Numune Training and Research Hospital, Ankara, Turkey

<sup>2</sup>Department of Statistic, Yıldırım Beyazıt University, Ankara, Turkey

Received: January 31, 2018 Accepted: May 22, 2018 Published online: December 24, 2018

## ABSTRACT

**Objectives:** The aim of this study was to translate the Neck Bournemouth Questionnaire into Turkish and to test the reliability and validity of the Turkish version of the Neck Bournemouth Questionnaire (BQc-t).

**Patients and methods:** Between June 2014 and July 2015, a total of 97 patients with neck pain (27 males, 70 females; mean age 46.6±10.6 years; range, 18 to 65 years) were included in the study. The patients underwent a physical therapy and rehabilitation program. For translation, the American Association of Orthopedic Surgeons guideline was used. The reliability was measured with internal consistency and test-retest by calculation of the Cronbach alpha and intraclass correlation coefficient (ICC) respectively. Internal construct validity of the BQc-t was analyzed with confirmatory factor analysis. For external construct validity, the correlations between the BQc-t results and the Neck Pain and Disability Scale (NPAD), Modified Neck Disability Index (MNDI), and Short Form-36 (SF-36) were analyzed before and after treatment. Responsiveness was calculated as the effect size (ES) and standardized response mean (SRM). Minimal detectable change (MDC) score was calculated to evaluate interpretability.

**Results:** The ICC value for test-retest of total score was 0.945. Pre- and post-treatment Cronbach alpha coefficients were 0.877 and 0.907, respectively, showing that the reliability of the BQc-t was considerably high. Confirmatory factor analysis showed that questions were found to cluster in a single dimension. In terms of the external construct validity, there was a positive statistically significant correlation between the BQc-t questions, except for Question 7, and relevant subscales of the NPAD and MNDI. There was also a negative statistically significant correlation between the BQc-t questions and SF-36 subgroups. The ES and SRM were 1.23 and 1.48, respectively. The MDC was 20.31.

**Conclusion:** Our study results show that the BQc-t is reliable, valid, and sensitive to clinical changes.

**Keywords:** Neck Bournemouth Questionnaire; neck pain; questionnaire; translation, reliability; validity.

In the pathologies of the musculoskeletal system, one of the most common problems is neck pain with a prevalence of 22 to 30% in the general population.<sup>[1]</sup> For the treatment, the effects of pain-related disability should be evaluated. The functional scales used should be valid, reliable, and sensitive to clinical changes. Currently, biopsychosocial scales developed for the assessment of pain are preferred over biological models.<sup>[2]</sup> The Bournemouth questionnaires were developed by Bolton and Breen<sup>[3]</sup> and Bolton and Humphreys<sup>[4]</sup> to test treatment efficacy for back pain in 1999 and for neck pain in 2002. The Neck Bournemouth Questionnaire

(BQc) is a multidimensional scale evaluating pain, physical function, social activity, anxiety, depression, work-related fear avoidance and pain control with seven items.<sup>[4]</sup> The maximum score is 70 points, obtained by totaling the scores of each of the seven items. It is sensitive to clinical changes and patients can fill it out themselves in a short time in clinical practice.<sup>[5]</sup> The French, Dutch, German, and Italian versions of the BQc were studied previously.<sup>[5-8]</sup> In 2016, the reliability and validity of the Turkish version of the Neck Bournemouth Questionnaire (BQc-t) for low back pain was studied by Gunaydin et al.<sup>[9]</sup>

**Corresponding author:** Onur Yılmaz, MD. Dr. Nafiz Körez Sincan Devlet Hastanesi, Fizik Tedavi ve Rehabilitasyon Kliniği, 06940 Sincan, Ankara, Turkey.

e-mail: onuryilmaz11@gmail.com

Cite this article as:

Yılmaz O, Gafuroğlu Ü, Yüksel S. Translation, reliability, and validity of the Turkish version of the Neck Bournemouth Questionnaire. Turk J Phys Med Rehab 2019;65(x):i-viii.

In the present study, we aimed to translate the BQc into Turkish and to test the reliability and validity of the Turkish version of the BQc-t (Appendix 1) based on the need for a comprehensive scale to evaluate patients with chronic neck pain in Turkish from a biopsychosocial perspective.

## PATIENTS AND METHODS

This article is produced from Dr. Onur Yilmaz's 2016 thesis study. Before the study, permission was obtained to initiate the study from Jennifer Bolton<sup>[4]</sup> who developed the BQc. A written informed consent was obtained from each patient. The study protocol was approved by the Ankara Numune Training and Research Hospital, Ethics Committee (220/2014). The study conducted in accordance with the principles of the Declaration of Helsinki.

### Translation and cultural adaptation

For translation and cultural adaptation, the American Association of Orthopedic Surgeons (AAOS) guideline was used.<sup>[10,11]</sup>

- *Stage 1 (Translation)*: The BQc was independently translated into Turkish by three translators: by a medical doctor fluent in English, by an English teacher who was working at a high school for 12 years, and by a professional translator. Only the medical doctor was informed about the questionnaire and the objective of translating it.

- *Stage 2 (Synthesis)*: A joint text was prepared with the participation of all translators.

- *Stage 3 (Back translation)*: The joint text was translated back into English by a professional translator and an English teacher who was living in Turkey for the last 10 years and who spoke Turkish, neither of whom were informed about the original text.

- *Stage 4 (Clinician's committee)*: The back translation was compared with the original text by a committee consisting of three psychiatrists. They were asked to comment on points which could be related to cultural differences and might result in difficulties in daily life. The questionnaire was also checked by a Professor of Turkish language and literature and corrections were made in terms of the sentence structures.

- *Stage 5 (Face validity, pretest)*: The pre-final version of the questionnaire was used on 20 patients. The patients were asked to mention the points they did not understand.

- *Stage 6 (Committee evaluation and test)*: The results of the pretest were evaluated by a committee

consisting of three psychiatrists and two physiotherapists. The committee members were also asked to provide their opinions about the questions. The final version was prepared.

### Study protocol

This study included a total of 97 patients (27 males, 70 females; mean age  $46.6 \pm 10.6$  years; range, 18 to 65 years) who were admitted to a physical therapy and rehabilitation outpatient unit with neck pain (for at least the previous 12 weeks) between June 2014 and July 2015. Exclusion criteria were as follows: having an inflammatory or infectious joint disease, shoulder pathology and fibromyalgia, diagnosis of psychiatric or neurological disease, history of malignancy, cervical trauma, and physical therapy to the neck within the previous six months.

All patients were evaluated with the BQc-t. In addition, for pain and disability, the Modified Neck Disability Index (MNDI) and the Visual Analog Scale (VAS) for the Neck Pain and Disability Scale (NPAD); for measuring the quality of life, the Short Form-36 (SF-36) were applied.<sup>[12-14]</sup> The patients were, then, recommended a physical therapy and rehabilitation program, and 79 patients accepted. They received hot pack therapy (20 min), continuous ultrasound ( $0.5 \text{ watt/cm}^2$ , 10 min), conventional transcutaneous electrical nerve stimulation (TENS) (30 min), and a neck exercise program at each therapy session for a period of two weeks.<sup>[15]</sup> After the patients completed the program, they were reevaluated with the existing scales. Eighteen patients who refused the therapy program underwent only a single evaluation.

### Statistical analysis

Statistical analysis was performed using the IBM SPSS version 21.0 software (IBM Corp., Armonk, NY, USA) with the R package "polycor", and MS-Excel 2007. Descriptive statistics were expressed in mean and standard deviation (SD) for normally distributed data and in median (min-max) number (n) and percentage (%) for nominal data. Type-1 error rate was considered  $\alpha=0.05$  as statistically significant. Depending on the types of variables, the Spearman rho or polyserial correlation coefficients were calculated. A two-way mixed model was used to calculate intraclass correlation coefficients (ICCs).

- *Reliability*: The reliability was measured with the internal consistency and test-retest by calculation of the Cronbach alpha and ICC, respectively.<sup>[16]</sup> For this purpose, 30 patients were called back 72 h after the first evaluation and before the therapy. They were

**Table 1.** Test-retest reliability of BQc-t

Item	ICC	95% CI		p
		Upper limit	Lower limit	
1	0.998	0.996	0.999	<0.001
2	0.995	0.990	0.998	<0.001
3	0.988	0.975	0.994	<0.001
4	0.998	0.996	0.999	<0.001
5	0.997	0.994	0.999	<0.001
6	0.997	0.994	0.999	<0.001
7	0.998	0.997	0.999	<0.001
Total	0.945	0.912	0.970	<0.001

BQc-t: Turkish version of the Neck Bournemouth Questionnaire; CI: Confidence Interval; ICC: Intraclass correlation coefficient.

asked to answer the questions in the BQc-t again. The reason for deciding on 72 h was that it was a short enough time period for the clinical findings not to change and also a long enough one for the patients who were unable to remember their first answers.

- *Validity:* The internal construct validity of the scale was analyzed with the confirmatory factor analysis. The confirmatory factor analysis (CFA) is a multivariate statistical procedure used to test how well the measured variables represent the number of constructs. For testing the external construct validity of the scale, the correlation between the BQc-t results and NPAD, MNDI, and SF-36 was analyzed before and after treatment.<sup>[12-14]</sup> Correlation coefficients of 1.00-0.91 were regarded as perfect, 0.90-0.71 as good, 0.70-0.51 as moderate, 0.50-0.31 as acceptable, and <0.30 as weak.<sup>[16,17]</sup> In the comparison between the scales, the questions matched for meaning during the previous studies were used.<sup>[4,6]</sup> For testing the external longitudinal construct validity, the differences of every item on the BQc-t and of the total scores in the corresponding subscales and the total scores in the other scales for pre- and post-treatment evaluations were calculated, and the significance of each difference and its level of correlation was interpreted.

- *Responsiveness:* To understand whether the BQc-t could measure clinical change before and after treatment, sensitivity to change was calculated with

the Kazis and Cohen methods as the effect size (ES) and standardized response mean (SRM). The common estimate of ES is the change in total score divided by the standard deviation of the pre-treatment value for the scale. The SRM is a value calculated by dividing the mean change score by the standard deviation of the change. When ES and SMR values were interpreted, values from 0.20-0.50 were regarded as weak, 0.50-0.80 as moderate, and above 0.80 as strong effects.<sup>[18,19]</sup> Interpretability of scale was evaluated by minimal detectable change (MDC). The MDC is the minimal amount of change score outside of measurement error which may reflect true change. Based on total scores of the BQc-t before and after treatment, MDC was calculated.<sup>[20]</sup> The MDC was estimated to analyze whether treatment resulted in a real change beyond measurement error. The MDC is calculated by multiplying the standard error of measurement by the Z score associated with the desired confidence level and the square root of 2, adjusted for sampling from two different measures. The standard error of measurement is estimated as the  $SD_{pooled}$  of pre- and post-treatment assessments multiplied by the square root of  $(1-r)$ , where  $r$  is the ICC. The standard error of measurement quantifies within-subject variability and reflects the amount of measurement error.

## RESULTS

A total of 97 patients with chronic neck pain were included in the study. The mean duration of neck pain was 48.0 (range, 6.0 to 300.0) months.

### Translation and cultural adaptation

All translators easily came to a consensus at the translation stage. Although there was no need for cultural adaptation, minor changes were made in some words. The translators reported that the term to help used in Question 7 might mean to cope or to aid, resulting in a difficulty in understanding by the patients, and they agreed on using the term to alleviate.

*Reliability-test-retest:* The ICC values of the questions of the BQc-t were found to be considerably high. Total ICC value was calculated as 0.945 (Table 1).

**Table 2.** Internal consistency reliability of BQc-t

Item	Item-corrected total correlation							Cronbach alpha (Total score)	n
	1	2	3	4	5	6	7		
Before treatment	0.594	0.731	0.744	0.719	0.651	0.719	0.477	0.877	97
After treatment	0.658	0.743	0.76	0.732	0.729	0.821	0.618	0.907	79

BQc-t: Turkish version of the Neck Bournemouth Questionnaire.

**Table 3.** Factor loadings and Cronbach alpha results based on eliminated questions

Item	Factor loadings	Cronbach alpha when question is deleted
1	0.706	0.868
2	0.824	0.851
3	0.833	0.848
4	0.797	0.852
5	0.751	0.862
6	0.811	0.852
7	0.588	0.882

- *Internal consistency:* The Cronbach alpha coefficients were calculated as 0.877 and 0.907 for pre- and post-treatment, respectively. For all subscales, the reliability was found to be considerably high in

the context of internal consistency (Table 2). The factor loading of Question 3 was found to be the highest (0.833), whereas that of Question 7 was the lowest (0.588). With the elimination of Question 3, the ICC would be calculated as 0.848, while with the elimination of Question 7, it would be calculated as 0.882 (Table 3).

### Validity

- *Internal construct validity:* The scale questions that were analyzed with factor analysis were found to cluster in a single dimension. This single dimension was found to explain 58.23% of the total variance.

- *External construct validity:* The Question 7 of the NPAD and Part 8 of the MNDI were found to be left blank by the patients who did not drive as anticipated. As these questions might lead to mistakes in the

**Table 4.** External construct validity of BQc-t items

Item	NPAD		MNDI		SF-36	
	Relevant dimension		Relevant dimension		Relevant dimension	
	Before treatment	After treatment	Before treatment	After treatment	Before treatment	After treatment
Pain	0.552***,†	-0.515***,†	0.489***,†	0.621***,†	-	-
Physical function	0.583***	0.540***	0.557***	0.578***	-0.290**	-0.223*
Social activity	0.745***	0.662***	0.414***,†	0.613***,†	-0.565***	-0.546***
Anxiety	0.527***	0.485***	0.388***,†	0.495***,†	-	-
Depression	0.565***,†	0.563***,†	-	-	-0.384***	-0.477***
Work related fear avoidance	0.710***,†	0.636***,†	-	-	-0.384***	-0.477***
Pain control	-0.052†‡	0.136†‡	-	-	-	-

BQc-t: Turkish version of the Neck Bournemouth Questionnaire; † Polyserial correlation coefficient, others Spearman rho coefficient; \* p<0.05; \*\* p<0.01; \*\*\* p≤0.001; ‡ p>0.05; NPAD: Neck Pain and Disability Scale; MNDI: Modified Neck Disability Index; SF-36: Short Form-36.

**Table 5.** External longitudinal construct validity of BQc-t items

Item	n	Correlation coefficient	NPAD	MNDI	SF-36
Pain	79	r	0.489	0.418	
		p	<0.01	<0.001	
Physical function	79	r	0.431	0.469	-0.69*
		p	<0.001	<0.001	0.001
Social activity	79	r	0.609	0.333	-0.638†
		p	<0.001	0.003	<0.001
Anxiety	79	r	0.525	0.397	
		p	<0.001	<0.001	
Depression	79	r	0.443		-0.316‡
		p	<0.001		0.005
Work related fear avoidance	79	r	0.545		
		p	<0.001		
Pain control	79	r	0.130		
		p	0.253		

BQc-t: Turkish version of the Neck Bournemouth Questionnaire; NPAD: Neck Pain and Disability Scale; MNDI: Modified Neck Disability Index; SF-36: Short Form-36; \* SF-36 Physical Functioning; † SF-36 Social Functioning; ‡ SF-36 Mental Health.

**Table 6.** Internal longitudinal construct validity of BQc-t items

Item	Item-corrected total correlation							Cronbach alpha (Total score)	n
	1	2	3	4	5	6	7		
	0.766	0.581	0.831	0.694	0.647	0.775	0.656	0.898	79

BQc-t: Turkish version of the Neck Bournemouth Questionnaire.

**Table 7.** Effect sizes of BQc-t, NPAD and MNDI and standardized response means

	$\Delta$	SD_bt	SD_at	SD*	SD**	ES†	ES‡	SRM
BQc-t	16.24	13.24	10.35	11.95	10.95	1.23	1.36	1.48
NPAD	17.39	16.59	13.73	15.31	10.86	1.05	1.14	1.14
MNDI	7.54	7.73	5.80	6.86	5.36	0.98	1.10	1.41

BQc-t: Turkish version of the Neck Bournemouth Questionnaire; NPAD: Neck Pain and Disability Scale; MNDI: Modified Neck Disability Index;  $\Delta$ : Mean change score; SD\_bt: Standard deviation of values before treatment; SD\_at: Standard deviation of values after treatment; SD\*: Joint standard deviation; SD\*\*: Standard deviation of change scores; ES†: Effect size calculated by Kazis method; ES‡: Effect size calculated by Cohen method; SRM: Standardized response mean.

statistical analysis, they were eliminated from the analysis as in previous studies.<sup>[21,22]</sup> By matching the related subscales of the NPAD and MNDI with the BQc-t before and after treatment, except for Question 20 of the NPAD which matched with Question 7 of the BQc-t, a moderate to strong, positive and statistically significant correlation was found. In addition, by matching the items on the BQc-t with the relevant subscales of the SF-36 a weak, acceptable, moderately negative and statistically significant correlation was found (Table 4). Furthermore, by comparing the total scores of the BQc-t with those of the NPAD and MNDI before and after treatment, a strong, positive and statistically significant correlation was found ( $r=0.763, 0.721$  and  $0.698, 0.750$ , respectively) ( $p \leq 0.001$ ).

- *External longitudinal construct validity:* Based on the data obtained by matching the relevant subscales of the NPAD and MNDI with the items on the BQc-t before and after treatment, except for Question 7, a weak, moderately positive, statistically significant correlation was found. Likewise, by comparing the items on the BQc-t with the relevant subscales of the SF-36 before and after treatment, a weak and moderately negative, statistically significant correlation was found ( $p < 0.001$ ) (Table 5). In addition, comparing the total scores of the BQc-t with those of the NPAD and MNDI before and after treatment was also found to be of statistical significance ( $r=0.833$  and  $r=0.727$ , respectively) ( $p < 0.001$ ).

### Responsiveness

The internal longitudinal construct validity of the BQc-t items was found to be considerably high and the Cronbach alpha coefficient was calculated as 0.898

(Table 6). When the sensitivities of the scales to change were analyzed, the ES of the BQc-t was 1.23 based on the Kazis method and 1.36 based on the Cohen method, and the SRM was found as 1.48. The ES and SRM values of the BQc-t were higher than those of the NPAD and MNDI (Table 7). The MDC score of the BQc-t was calculated as 20.31.

## DISCUSSION

When examining the prevalence of neck pain, it is important to use an objective measurement tool for optimal clinical follow-up. Therefore, we attempted to adapt the BQc, which evaluates treatment efficacy in patients from a biopsychosocial perspective in seven items, into Turkish. The present study included 97 patients with neck pain. Tabachnick and Fidell<sup>[23]</sup> reported that, for factor analysis, a ratio of 10 participants per item was sufficient for the sample size. Thus, the required sample size was sufficient for the seven items. At the stage of adapting the BQc into Turkish, the AAOS guideline was used. In the literature, the number of subjects suggested for participation in the pretest stage ranges from 3 to 50.<sup>[11]</sup> Soklic et al.<sup>[6]</sup> and Gunaydin et al.<sup>[9]</sup> conducted a pretest on 30 patients in their studies. In the present study, a pretest on 20 patients was performed. At this stage, the patients reported that the questionnaire was understandable. This supports the face validity of the BQc-t.<sup>[24]</sup>

For a scale to be valid, its reliability is a prerequisite. If the ICC value of a test is higher than 0.7, it is reliable in the context of test-retest.<sup>[25,26]</sup> The ICC values of the French and the German versions of the BQc

were found to be 0.99 and 0.97, respectively.<sup>[5,6]</sup> In the present study, the ICC was identified as 0.945. This result shows that, in repetitive measurements, the BQc-t is a scale with a low margin of error, generating consistent results from one implementation to another. In the French study, the time between test-retest was 24 h, while it was two h in the German study, and in this study it was 72 h.<sup>[5,6]</sup> The fact that ICC values in previous studies were higher can be explained by their having shorter time periods between test and retest.

The internal consistency evaluates whether the questions in a scale are correlated amongst themselves.<sup>[24,25]</sup> In general, having a Cronbach alpha coefficient of 0.7 or higher is regarded as satisfactory.<sup>[26]</sup> The Cronbach alpha coefficient measured before and after treatment was 0.79 and 0.80, respectively in the study by Soklic et al.<sup>[6]</sup> and as 0.87 and 0.90, respectively in the study by Bolton and Humphreys<sup>[4]</sup> In the present study, the Cronbach alpha coefficients for pre- and post-treatment were calculated as 0.877 and 0.907, respectively. These values indicate that the reliability of the BQc-t is at an acceptable level. In the present study, similarly to the studies conducted in the United Kingdom and Germany, it was found that Question 7 had the lowest factor loading (0.588).<sup>[4,6]</sup> Despite the fact that the Cronbach alpha coefficient was higher than 0.4, compared to previous questions, it was significantly lower. When approached from the perspective of integrity of meaning, it is seen that the first six questions examine the patient's symptoms and their effects on daily life, while Question 7 examines the patient's individual efforts to eliminate the pain. It is considered that this difference in meaning distances Question 7 from the other questions in terms of measuring the same feature.

For testing external construct validity, Bolton and Humphreys<sup>[4]</sup> compared the BQc with the Copenhagen Functional Disability Scale, Neck Disability index (NDI), and SF-36 subscales before and after treatment in their study and, except for Question 7, they observed a statistically significant correlation among all subscales. The authors explained this weak correlation by asserting that the general health scale of SF-36 could not sufficiently reflect the subscale examined in Question 7, which is the ability to control the pain by oneself. In the aforementioned study, the social function scale of the SF-36 and Question 3 were not correlated for the external longitudinal construct validity. In the present study, before and after treatment, based on the external construct validity and external longitudinal construct validity results of the items on the BQc-t matched with the relevant subscales

of the NPAD and MNDI, except for Question 20 of the NPAD which matched with Question 7 of the BQc-t, there was a positive, statistically significant correlation. In the study by Soklic et al.,<sup>[6]</sup> correlation between these two questions was statistically significant, yet it was low compared to the others. It is considered that the reason for this result in the present study was related to the difficulty in comprehending the question by the patients. Question 20 of the NPAD questioned the effectiveness of painkillers; however, Question 7 of the BQc-t asked the patients to analyze their individual efforts to control their pain as well as the treatment methods they employed to this end. The reason for the difficulty in meaning is also related to the different question techniques. The increase in the VAS scores is numerically inversely related to the success in taking the pain under control; thus, the increase in scores for controlling the pain means a reduction in pain. Likewise, before and after treatment, based on the external construct validity and external longitudinal construct validity results of the relevant subscales of the SF-36 and the corresponding items of the BQc-t, there was a negative, statistically significant relationship. Bolton and Humphreys<sup>[4]</sup> and Martel et al.<sup>[5]</sup> compared the total scale scores of the BQc and NDI in their studies. They found the pre- and post-treatment external construct validity, and external longitudinal construct validity of the BQc as statistically significant ( $r=0.51, 0.71, 0.50$  and  $0.67, 0.61, 0.42$ , respectively). In the present study, when the total scores of the BQc-t with the MNDI and NPAD were compared, the pre- and post-treatment external construct validity, and external longitudinal construct validity of the BQc-t were found to be 0.698, 0.750, 0.727, and 0.763, 0.721, 0.833, respectively. All these results prove the external construct validity of the BQc-t.

In the study by Bolton and Humphreys<sup>[4]</sup> which evaluated whether the BQc measured clinical changes before and after treatment, the internal longitudinal construct validity was calculated and a statistically significant correlation was found. In the present study, the Cronbach alpha coefficient for internal longitudinal construct validity of the BQc-t was 0.898. Furthermore, for all the subscales, the internal longitudinal construct validity was considerably high.

Sensitivity to clinical changes of the scale is an important, but often overlooked subject. If the scores of the scale remain unchanged with clinical changes, the scale is considered not being sufficiently sensitive. Also, having a sensitive scale makes it possible to evaluate the treatment success. Bolton and Humphreys<sup>[4]</sup> calculated the ES and SRM values of the BQc with reference to

Kazis as 1.67 and 1.43, respectively, while these were calculated as 0.80 and 0.83, respectively for the NDI. Martel et al.<sup>[5]</sup> calculated the ES and SRM values for the BQc as 0.56 and 0.61 and for the NDI as 0.51 and 0.58, respectively. In both these studies, the ES and SRM values of the BQc were higher than those of the NDI. In the present study, the sensitivity of the BQc-t to change after physical therapy by the Kazis and Cohen methods was found to be 1.23 and 1.36, respectively. In addition, the SRM value for the BQc-t was calculated as 1.48. In the present study, the ES and SRM values of the BQc-t were higher than the values of the NPAD and MNDI. These results demonstrate that the BQc-t is a scale which is sensitive to changes and that it can be used to follow patients. Furthermore, Bolton<sup>[27]</sup> indicated that an improvement of 13 points on the total score was associated with clinically significant improvement. In the present study, the MDC score of the BQc-t was calculated as 20.31. In other words, to identify whether there was a clinically significant change before and after treatment, a difference of 20.31 points between the total scores of the BQc-t was accepted as the threshold value.

In conclusion, based on our study results, the BQc-t is reliable, valid, and sensitive to clinical changes. All these results demonstrate that the use of the BQc-t in clinical follow-up is helpful in evaluating treatment results in an objective manner.

#### Declaration of conflicting interests

The authors declared no conflicts of interest with respect to the authorship and/or publication of this article.

#### Funding

The authors received no financial support for the research and/or authorship of this article.

## REFERENCES

1. Bovim G, Schrader H, Sand T. Neck pain in the general population. *Spine (Phila Pa 1976)* 1994;19:1307-9.
2. Guzman J, Hurwitz EL, Carroll LJ, Haldeman S, Côté P, Carragee EJ, 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. *Spine (Phila Pa 1976)* 2008;33:14-23.
3. Bolton JE, Breen AC. The Bournemouth Questionnaire: a short-form comprehensive outcome measure. I. Psychometric properties in back pain patients. *J Manipulative Physiol Ther* 1999;22:503-10.
4. Bolton JE, Humphreys BK. The Bournemouth Questionnaire: a short-form comprehensive outcome measure. II. Psychometric properties in neck pain patients. *J Manipulative Physiol Ther* 2002;25:141-8.
5. Martel J, Dugas C, Lafond D, Descarreaux M. Validation of the French version of the Bournemouth Questionnaire. *J Can Chiropr Assoc* 2009;53:102-20.
6. Soklic M, Peterson C, Humphreys BK. Translation and validation of the German version of the Bournemouth Questionnaire for Neck Pain. *Chiropr Man Therap* 2012;20:2.
7. Schmitt MA, de Wijer A, van Genderen FR, van der Graaf Y, Helder PJ, van Meeteren NL. The Neck Bournemouth Questionnaire cross-cultural adaptation into Dutch and evaluation of its psychometric properties in a population with subacute and chronic whiplash associated disorders. *Spine (Phila Pa 1976)* 2009;34:2551-61.
8. Geri T, Signori A, Gianola S, Rossetini G, Grenat G, Checchia G, et al. Cross-cultural adaptation and validation of the Neck Bournemouth Questionnaire in the Italian population. *Qual Life Res* 2015;24:735-45.
9. Gunaydin G, Citaker S, Meray J, Cobanoglu G, Gunaydin OE, Hazar Kanik Z. Reliability, Validity, and Cross-Cultural Adaptation of the Turkish Version of the Bournemouth Questionnaire. *Spine (Phila Pa 1976)* 2016;41:1292-7.
10. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993;46:1417-32.
11. Acquadro C, Conway K, Hareendran A, Aaronson N. Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value Health* 2008;11:509-21.
12. Kesiktas N, Ozcan E, Vernon H. Clinimetric properties of the Turkish translation of a modified neck disability index. *BMC Musculoskelet Disord* 2012;13:25.
13. Bicer A, Yazici A, Camdeviren H, Erdogan C. Assessment of pain and disability in patients with chronic neck pain: reliability and construct validity of the Turkish version of the neck pain and disability scale. *Disabil Rehabil* 2004;26:959-62.
14. Koçyiğit H, Aydemir Ö, Ölmez N, Memiş A. Kısa Form-36 (KF-36)'nın Türkçe versiyonunun güvenilirliği ve geçerliliği. *İlaç ve Tedavi Dergisi* 1999;12:102-6.
15. Swenson RS. Therapeutic modalities in the management of nonspecific neck pain. *Phys Med Rehabil Clin N Am* 2003;14:605-27.
16. Alpar R. Uygulamalı İstatistik ve Geçerlilik ve Güvenilirlik. 4. Baskı. Ankara: Detay Yayıncılık; 2012.
17. Donner A, Eliasziw M. Sample size requirements for reliability studies. *Stat Med* 1987;6:441-8.
18. McDowell I. Measuring health: a guide to rating scales and questionnaires. 3rd ed. Oxford: Oxford University Press; 2006.
19. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. New York: Lawrence Erlbaum Associates; 1998.
20. Schwenk M, Gogulla S, Englert S, Czempik A, Hauer K. Test-retest reliability and minimal detectable change of repeated sit-to-stand analysis using one body fixed sensor in geriatric patients. *Physiol Meas* 2012;33:1931-46.
21. Wlodyka-Demaille S, Poiraudou S, Catanzariti JF, Rannou F, Fermanian J, Revel M. French translation and validation of 3 functional disability scales for neck pain. *Arch Phys Med Rehabil* 2002;83:376-82.
22. González T, Balsa A, Sáinz de Murieta J, Zamorano E, González I, Martín-Mola E. Spanish version of the Northwick Park Neck Pain Questionnaire: reliability and validity. *Clin Exp Rheumatol* 2001;19:41-6.

23. Tabachnick BG, Fidell LS. Using multivariate statistics. 5th ed. Needham Heights: Allyn&Bacon Inc; 2006.
24. DeVon HA, Block ME, Moyle-Wright P, Ernst DM, Hayden SJ, Lazzara DJ, et al. A psychometric toolbox for testing validity and reliability. J Nurs Scholarsh 2007;39:155-64.
25. Gözüm S, Aksayan S. Kültürlerarası Ölçek Uyarlaması için Rehber 2: Psikometrik Özellikler ve Kültürlerarası Karşılaştırma. Hacettepe Araştırma ve Geliştirme Dergisi 2003;1:3-14.
26. DeVellis RF. A consumer's guide to finding, evaluating, and reporting on measurement instruments. Arthritis Care Res 1996;9:239-45.
27. Bolton JE. Sensitivity and specificity of outcome measures in patients with neck pain: detecting clinically significant improvement. Spine (Phila Pa 1976) 2004;29:2410-7.

### Appendix 1. Turkish version of the Neck Bournemouth Questionnaire

#### BOURNEMOUTH BOYUN ANKETİ

Hasta adı:

Tarih:

**Yönerge:** Aşağıdaki ölçekler boyun ağrınızı ve bunun sizi nasıl etkilediğini ortaya çıkarmak için tasarlanmıştır. Lütfen TÜM ölçekleri, her bir ölçek üzerinde nasıl hissettiğinizi en iyi tanımlayan BİR numarayı işaretleyerek cevaplayınız.

1. Geçtiğimiz hafta boyunca, ortalama olarak, boyun ağrınızı nasıl derecelendirirsiniz?

Ağrı yok

Olabilecek en kötü ağrı

0 1 2 3 4 5 6 7 8 9 10

2. Geçtiğimiz hafta boyunca, boyun ağrınız günlük aktivitelerinizi (ev işi, yıkama, giyinme, kaldırma, okuma, araba kullanma) ne kadar engelledi?

Engellemedi

Aktiviteleri gerçekleştirmemi engelledi

0 1 2 3 4 5 6 7 8 9 10

3. Geçtiğimiz hafta boyunca, boyun ağrınız eğlence (hobi), sosyal ve aile aktivitelerinde yer alma becerinizi ne kadar engelledi?

Engellemedi

Aktiviteleri gerçekleştirmemi engelledi

0 1 2 3 4 5 6 7 8 9 10

4. Geçtiğimiz hafta boyunca, ne kadar endişeli (gergin, sinirli, asabi, konsantre olmakta zorluk/gevşemede zorluk) hissettiniz?

Hiç endişeli değil

Son derece endişeli

0 1 2 3 4 5 6 7 8 9 10

5. Geçtiğimiz hafta boyunca, ne kadar depresif (canı sıkın, üzgün, keyifsiz, kötümser, mutsuz) hissettiniz?

Hiç depresif değil

Son derece depresif

0 1 2 3 4 5 6 7 8 9 10

6. Geçtiğimiz hafta boyunca, işinizin (ev içinde ve dışında) boyun ağrınızı nasıl etkilediğini (ya da etkileyeceğini) hissettiniz?

Kötüleştirmedim

Çok kötüleştirdi

0 1 2 3 4 5 6 7 8 9 10

7. Geçtiğimiz hafta boyunca, boyun ağrınızı kendi kendinize ne kadar kontrol (azaltmak/hafifletmek) edebildiniz?

Tamamen kontrol ettim

Hiç kontrol edemedim

0 1 2 3 4 5 6 7 8 9 10