Taylor & Francis
Taylor & Francis Group

Disability and Rehabilitation

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/idre20

The validity and reliability of the Turkish version of the 12-item Örebro musculoskeletal screening questionnaire (ÖMSQ-12-TR)

Emel Tasvuran Horata, Pervin Demir, Gözde Yağcı, Suat Erel, Fatma Eken & Charles Philip Gabel

To cite this article: Emel Tasvuran Horata, Pervin Demir, Gözde Yağcı, Suat Erel, Fatma Eken & Charles Philip Gabel (2022): The validity and reliability of the Turkish version of the 12-item Örebro musculoskeletal screening questionnaire (ÖMSQ-12-TR), Disability and Rehabilitation, DOI: 10.1080/09638288.2022.2089918

To link to this article: https://doi.org/10.1080/09638288.2022.2089918

DISABILITY AND REHABILITATION
https://doi.org/10.1080/09638288.2022.2089918

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

The validity and reliability of the Turkish version of the 12-item Örebro musculoskeletal screening questionnaire (ÖMSQ-12-TR)

Emel Tasvuran Horata[a], Pervin Demir[b], Gözde Yağcı[c], Suat Erel[d], Fatma Eken[a] and Charles Philip Gabel[e]

[a]Department of Physiotherapy and Rehabilitation, Faculty of Health Sciences, Afyonkarahisar Health Science University, Afyonkarahisar, Turkey; [b]Department of Biostatistics and Medical Informatics, Faculty of Medicine, Ankara Yıldırım Beyazıt University, Ankara, Turkey; [c]Department of Physiotherapy and Rehabilitation, Faculty of Physical Therapy and Rehabilitation, Hacettepe University, Ankara, Turkey; [d]School of Physiotherapy and Rehabilitation, Pamukkale University, Denizli, Turkey; [e]Access Physiotherapy, Coolum Beach, Qld Australia

ABSTRACT

Purpose: The 12-item Örebro Musculoskeletal Screening Questionnaire (ÖMSQ-12) is a multidimensional questionnaire assessing general musculoskeletal problems. This study aimed to investigate its construct validity and reliability.

Materials and methods: Confirmatory factor analysis (CFA) was performed for construct validity. The Tampa Scale for Kinesiophobia (TSK) and the SF-12 and Pain Numerical Rating Scale (P-NRS) were used for convergent validity. Reliability (ICC), internal consistency (Cronbach's alpha), reproducibility, and known-group validity were assessed. The cut-off value was measured.

Results: A total of $n = 378$ individuals (aged 35.7 ± 12.4 years, female = 73.3%) with a musculoskeletal problem participated in the study. P-NRS score of the individuals was 5. Results showed that a 3-factor model did fit well under CFA ($\chi^2/df = 2.76 \leq 3$). The questionnaire had good reliability (ICC = 0.865) and internal consistency ($\alpha = 0.810$). There were no floor or ceiling effects (<%15). Total ÖMSQ-12-TR scores had a correlation with the TSK, SF-12 and P-NRS ($r = 0.303$–0.609). The AUC for the risk of absenteeism from work was obtained as 0.738 ($p < 0.001$). The risk of absenteeism was high in individuals with an ÖMSQ-12-TR score of ≥57.5.

Conclusions: The ÖMSQ-12-TR is a valid and reliable questionnaire that can be used in determining the risk of absenteeism in musculoskeletal disorders and is convenient for online use.

Clinical Trial Number: NCT04723615

ARTICLE HISTORY
Received 19 January 2022
Revised 8 June 2022
Accepted 10 June 2022

KEYWORDS
Musculoskeletal diseases; musculoskeletal pain; absenteeism; reliability and validity; risk assessment; patient reported outcome measures

Introduction

Musculoskeletal disorders are very common in developed and developing countries [1]. The data of the Global Burden of Diseases in 2019 revealed that approximately 1.7 billion people around the world had musculoskeletal disorders [2]. Although its prevalence varies according to age and diagnosis, globally people of all ages are affected [1,3]. Lumbar pain is the most common musculoskeletal disorder and accounts for about 36.8% of musculoskeletal disorders while neck pain accounts for about 18.4% [3].

Musculoskeletal disorders increase the need for rehabilitation and disability as well as being the main cause of absenteeism from work [2]. The rate of absenteeism due to musculoskeletal disorders is about 23.7% [4]. A decrease in production and the treatment methods developed for individuals with a high risk of disability [5] cause economic burdens for the present health systems and individuals [6]. Early diagnosis of musculoskeletal disorders and early interventions can prevent the disability, decrease healthcare costs [7] and reduce absenteeism [8]. Moreover, using high-cost assessments and treatment interventions only for the population at high risk can improve the ratio between the cost and benefit obtained from the treatment of musculoskeletal disorders. Consequently, cost-effective, low administrative, and patient burden screening questionnaires are needed to shed light on physiotherapy, rehabilitation assessments, and treatment approaches used to determine the high-risk individuals with musculoskeletal disorders.

There are a limited number of scales developed to screen musculoskeletal disorders. The Nordic Musculoskeletal Questionnaire was developed to screen for musculoskeletal symptoms and analyze the activity restrictions caused by them [9]. The STarT Back Screening Tool is used in identifying the modifiable physical and psychological risk factors in individuals with low back pain (LBP) [10]. The Health 2000 Survey is a screening tool assessing the risk of disability retirement for the workers [11]. While most of these scales assess the symptoms of musculoskeletal problems unidimensionally, some of them focus on a single region. The Örebro Musculoskeletal Screening Questionnaire-12 (ÖMSQ-12) is a multidimensional and practical questionnaire assessing general musculoskeletal problems [12,13] and determining the risk of absenteeism in individuals with a musculoskeletal problems [14]. The (ÖMSQ-12) was adapted from the ÖMSQ-21 [12,15] which was developed by a multi-stage reduction process from the

CONTACT Emel Tasvuran Horata ethorata@gmail.com Department of Physiotherapy and Rehabilitation, Afyonkarahisar Health Sciences University, Zafer Health Campus, A Block, 2078 Street No: 3, Afyonkarahisar, Turkey

Supplemental data for this article can be accessed online at https://doi.org/10.1080/09638288.2022.2089918

original Acute LBP screening questionnaire (ALPBSQ) [16]. The ALBPSQ was subsequently adapted through simple language modification to become the Örebro Musculoskeletal Screening Pain Questionnaire (ÖMPSQ) [17] however this was validated for workers, with pain, in the low back region; where as the ÖMSQ and ÖMSQ-12 were developed and validated for all individuals workers or not, with any problem, not just pain, and for any region, not just the low back. This modification and features enabled application to a more general population [12,14] which provides an advantage to the evaluator. Therefore, this study aimed to perform the translation and cross-cultural adaptation of the ÖMSQ-12 to Turkish and investigate its clinimetric properties including factor structure, construct validity and reliability.

Methods

Study design

This study was methodologically designed and approved by the Clinical Research Ethics Committee of Afyonkarahisar Health Science University (2020/469). The study was conducted in conformity with the principles of the Declaration of Helsinki. The clinical Trial Number is NCT04723615.

Translation and adaptation process

Permission to investigate the clinimetric properties of ÖMSQ-12-TR was obtained from the original author in 2018. The ÖMSQ-12 was originally developed in the English language through modifications to the Örebro Musculoskeletal Pain Screening Questionnaire (ÖMPSQ) by removing the emphasis from pain to problem, from low back to all musculoskeletal conditions, and including all populations, not just workers. The study aimed to translate the questionnaire into Turkish (the target language). Three physiotherapists, each with >16 years of experience, and with a strong command of the English language, independently and individually translated the English-language version of the ÖMSQ-12 into Turkish (forward translation). The study team examined the three versions and created a reconciled consensus translation. Then 11 academics, experts in the field of physiotherapy and rehabilitation, were asked whether the forward consensus translation was appropriate for the content and cultural relevance of the translation. Items 4, 6, 7, 11, and 12 required a revision and were adapted in line with the suggestions of the expert opinions. Subsequently, a native English specialist who understands and speaks Turkish fluently blindly translated the draft Turkish version back into English (backward translation). The text translated back was compared with the original questionnaire and sent to the researcher publishing the original scale to assess the compatibility of the translation and his/her approval was obtained [18].

The pre-final version of the questionnaire was pre-tested for any ambiguity in meaning during the questionnaire administration by 10 participants (5 males, 5 females; age 36–57 years) with musculoskeletal conditions. There were no ambiguities in meaning reported and the final draft was retained without further modification.

Participants

The snowball sampling method was used for the study [19]. The online questionnaire form created over Google forms was sent to the individuals *via* the internet. The online questionnaire form included an informed explaining the objective and content of the study before the questions of the questionnaire were shown on the screen. The participants who consented to participate in the study could then answer the questions. The questionnaire of those who did not consent to participate in the study was automatically concluded. Data were collected between December 2020 and August 2021.

The *inclusion criteria* were: having an acute/subacute or chronic musculoskeletal complaint/problem and a diagnosis in the spine, upper extremity and/or lower extremity, and being within the age group of 18–65.

The *exclusion criteria* were: pregnancy, presence of red flag signs including high fever, sudden weight loss, unbearable pain, and so forth, having an additional acute or chronic orthopedic, neurologic or cognitive disorder apart from musculoskeletal disorders, and having undergone surgery for the upper extremity, lower extremity or spine within the last six months.

Data collection

Data collection tools

Considering the data obtained from the previous studies, a participant information form assessing the participants' personal demographic characteristics and musculoskeletal disorders (age, gender, body mass index, occupational status, medical certificate, etc.) was prepared for this study. The convergent validity of the questionnaire was assessed using the Tampa Scale [20] and, SF-12 Quality of Life Questionnaire [21] and a Pain Numerical Rating Scale (P-NRS) [22]. Each of these criteria questionnaires has Turkish versions and measure the same concept within the ÖMSQ-12-TR's subdimensions. The ÖMSQ-12-TR is expected to be negatively correlated with SF12 and positively correlated with the TAMPA and P-NRS scales. The ÖMSQ-12-TR was re-completed by the participants to assess the reproducibility or test-retest reliability three days after the first study data were obtained.

ÖMSQ-12

The ÖMSQ-12 is a 12-item self-report questionnaire aiming to screen for the risk of chronicity or delayed recovery and to predict a variety of outcomes including problem severity, functional impairment, the status of receiving a medical certificate, cost, and time of recovery. It is designated for individuals who have had an acute or subacute work injury and had presented with musculoskeletal pain in the regions of the spine, upper and lower extremities [23]. The questionnaire has regional patient-reported outcome measures (PROM) for function and an 11-point numerical rating scale for perceived problem or pain severity, except for the first item. Each item is scored from the response on an 11-point Likert scale (0–10) according to the question asked. Items 8, 11, and 12 are reverse-scored items. The highest score that can be obtained from the questionnaire is 120. A high score means a high risk of absenteeism, high cost, chronicity or delayed recovery, and problem severity due to the problematic musculoskeletal disorder of the individual. Takasaki and Gabel assessed a population with all musculoskeletal disorders (subacute and chronic) not caused by work injuries and revealed in their study that the questionnaire could be used for all musculoskeletal disorders [24]. The questionnaire has 3 subdimensions: 1-Physical, fear avoidance and satisfaction (Items 8, 9, 10, 11, and 12), 2- Psyche and other (Items 2, 5, 6, and 7) and 3-Problem (Items 1, 3 and, 4). The original questionnaire was developed by Gabel et al. [14]. The Cronbach's alpha coefficient of the original questionnaire is 0.75.

Tampa scale for kinesiophobia (TSK)

The TSK was first developed by Kori et al. to assess the kinesiophobia in adult individuals with chronic musculoskeletal pain [25].

It is a 17-item self-report scale with a four-point Likert-type rating. The highest score that can be obtained from the scale is 68. A high score obtained from the scale is associated with higher disability and fear of activity [26]. The validity and reliability of the Turkish version were performed by Tunca Yilmaz et al. [20].

SF-12

The SF-12 has similar clinimetric properties to the SF-36 and is the preferred version as it is more practical and easy to use. It has two subdimensions providing general information about physical and mental health. The items with two, three and four answer alternatives are used, and the time frame considered is the last four weeks for the assessment. The subscales assess the health condition between 0 and 100 where 0='Poor health', 100='Good health'. Turkish validity and reliability have been performed [21].

Pain numerical rating scale (P-NRS)

The P-NRS was provided to participants to rate the present pain severity they perceived due to the musculoskeletal problem (including spine, lower extremity and, upper extremity) on an 11-point Likert scale (0–10), 0 = "No pain", 10 = Worst/severe pain possible' [27].

Sample size

The minimum sample size was determined for each statistical hypothesis as follows.

Construct validity [CFA]

"Sample sizes are given in factor analysis of 100 as poor, 200 as fair, 300 as good, 500 as very good and 1000 as excellent. It is comforting to have at least 300 cases for factor analysis" [28]. As a rule of thumb, the minimum is to have at least five or ten times as many observations as the number of variables to be analyzed (subject-to-variables ratio: $n = 120$ for this study, being at least 10 times the number of items). Considering recommendations in the literature, the sample size for factor analysis was determined as $n > 300$ participants.

External construct validity [convergent validity]

The required sample size was accepted at $n > 107$ participants to detect a correlation coefficient of at least 0.70 (>0.50), $\alpha = 0.05$ and power of 90.0% [29].

Reproducibility [test-retest reliability]

The minimum sample size requirement for estimating at least 0.95 (>0.90) intraclass correlation coefficient (ICC) with 90% power and $\alpha = 0.05$ $n = 68$ participants [30].

Statistical analysis

The statistical analyses and calculations were performed using the SPSS Version 21.0 and R software [31]. Using R packages: the "psych" [32], "lavaan" [33], and "semPlot" [34] packages. The statistical significance level was accepted as $p < 0.05$. The quantitative and qualitative variables were summarized as mean ± standard deviation (SD, median, quartile 1, quartile 3) and frequency (percentage), respectively.

The clinimetric properties of the ÖMSQ-12-TR were examined as follows:

The **language and content validity** of the scale was examined by using the content validation form based on the Davis method [35]. Eleven experts were requested to provide a score on each item independently. The item-level content validity index (I-CVI), scale-level content validity index based on the average method (S-CVI/Ave), and based on the universal agreement method (S-CVI/UA) were calculated to evaluate the content validity of the measurement tool [36]. The McNemar-Bowker test was used to evaluate whether the answers given to the two evaluations with more than two categories on understandability were consistent.

The **internal construct validity** was defined by CFA based on a polychoric correlation matrix. A second-order CFA was employed by using the R software to confirm whether the data could fit the 3-factor model of the ÖMSQ-12 as determined by Gabel et al. [14]. The overall model fit was assessed based on fit indices [37]. The goodness-of-fit indices were calculated and their cut-off values for adequate fit are as follows: Chi-square/degree of freedom (χ^2/df $\leq$3), Root Mean Square Error of Approximation (RMSEA $\leq$ 0.08), Standardized Root Mean Square Residuals (SRMR $\leq$ 0.08), Comparative Fit Index (CFI $\geq$ 0.95), Tucker-Lewis Index (TLI $\geq$ 0.95), Goodness of Fit Index (GFI $\geq$ 0.95), Normed Fit Index (NFI $\geq$ 0.95).

The **internal consistency reliability** of the ÖMSQ-12-TR instrument was determined by Cronbach's α value. The rule of thumb for assessing the Cronbach α was: >0.90:excellent, >0.80:good, >0.70:acceptable, >0.60:questionable, >0.50:poor internal consistency [38]. If more than 15% of respondents achieved the lowest or highest possible ÖMSQ-12-TR total score, floor and ceiling effects were present [39].

The **reproducibility (test-retest reliability)** was assessed by the intraclass correlation coefficient (ICC, two-way mixed-effects model, single measurement, absolute agreement) and its 95% confidence interval (CI). The ICC score >0.75 indicated good reliability [40]. The standard error of measurement (SEM = SD$\sqrt{}$(1-ICC), with SD representing the standard deviation of the measure) was calculated.

The spearman rho correlation coefficients between ÖMSQ-12-TR scores and P-NRS, TAMPA, SF12-PCS, and SF12-MCS scores were calculated for **convergent validity**. The coefficients <0.30:negligible, <0.50:low, <0.70:moderate, <0.90:high, and >0.90:very high correlation [41].

Known-group validity to determine the degree to which the ÖMSQ-12-TR scale could discriminate between the participants' risk of not returning to work was tested using the Wilcoxon rank-sum test and the receiver operating characteristic curve (ROC) to determine cut-off points using the Youden Index. The area under the ROC (AUC), sensitivity, specificity, and accuracy values were calculated.

After the evaluation of the validity and reliability, a comparison of the participant's ÖMSQ-12-TR score according to the demographic properties was tested by the Wilcoxon rank-sum test, Kruskal-Wallis test, or Spearman rho correlation coefficient.

Results

The questionnaire was completed by a total of 416 individuals with musculoskeletal problems. Two individuals did not accept to participate in the study and 36 individuals who did not meet the inclusion criteria were excluded. The rate of female was 73.3%. The mean age and BMI of the participants were 35.7 ± 12.4 years and 24.9 ± 4.3 kg/m^2 respectively (Supplementary Table 2). The median P-NRS score of the participants was 5 (Q1:4, Q3:6) (Supplementary Table 3). There were missing observations in some of the demographic information, but no missing value in item responses.

At the translation and adaptation of the questionnaire stage, the experts suggested the revisions for 4, 6, 7, 11, and 12 items. The items were adapted in line with the suggestions of the expert opinions. The adaptations and reasons are summarized in Supplementary Table 1. After the pre-test stage of the questionnaire, there were no ambiguities in meaning reported and the final draft was retained without further modification.

Clinimetric properties of the ÖMSQ-12 Turkish version (ÖMSQ-12-TR)

The validity index results were similar for language and content validity forms. The I-CVI ranged from 0. 91 to 1.00, the S-CVI/Ave was 0.96 and the S-CVI/UA was 0.58 as a result of the initial assessments. At the second evaluation after the suggesting adjustments, the I-CVI was 1.00 for each item, the S-CVI/Ave and the S-CVI/UA were obtained as 1.00. In the first forms, 82.3% of the participants stated their status of understanding the questions in the questionnaire as good or very good. The answers given about the understandability of questionnaire items in the second form were similar to the answers in the first forms ($p = 0.618$).

There was a high correlation between Item 5 and Item 6 (rp = 0.78) and Item 11 and Item 12 (rp = 0.85). Considering the covariance between the related items, the results of CFA model created for the determining factors of ÖMSQ-12-TR were summarized in Figure 1. When the factor loadings were assessed it was observed that the coefficient obtained for Item 1 was quite low. According to the answers given to that item, the rate of participants with pain or problem present for more than 1 year was 62%. When the related item is removed from the model fit it then indicates a poor fit; therefore, the model was established without removing the item. As a result of the CFA, the 3-factor model did fit well ($\chi^2/df = 2.76 \leq 3$). When the other fit indices were assessed the model fit was found to be satisfactory (RMSEA = $0.07 \leq 0.08$, SRMR = $0.07 \leq 0.08$, CFI = $0.97 \geq 0.95$, TLI $\geq$ 0.96, GFI = $0.998 \geq 0.95$, NFI = $0.95 \geq 0.95$). The CFI, TLI, GFI, and NFI were >0.95 to be considered as an acceptable fit.

The internal consistency coefficient of the ÖMSQ-12-TR was obtained $\alpha = 0.810$ (good). The Cronbach α ranged between 0.546 and 0.719 for the subscales of the ÖMSQ-12-TR (Table 1). The median ÖMSQ-12-TR score was 52 (Q1:39, Q3:64, min: 9, max: 104). There were no floor and ceiling effects (<%15) for any scales.

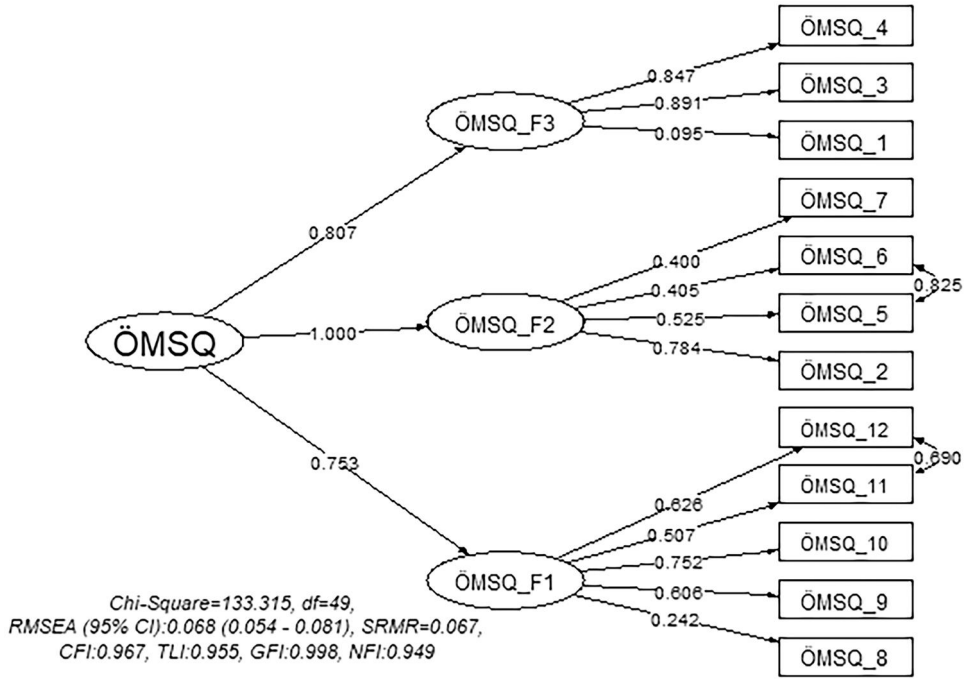

Figure 1. Confirmatory factor analysis results of the second-order three-factor model.

Table 1. Internal consistency coefficients and descriptive statistics for scales.

Scales	Number of samples	Number of items	Cronbach's alpha (lower–upper)	Min-max	Median (Q1–Q3)	Floor-ceiling effect %
ÖMSQ	378	12	0.810	9–104	52 (39–64)	0–0
ÖMSQ F1	378	5	0.719	0–41	14 (8–20)	1.9–0
ÖMSQ F2	378	4	0.695	0–38	20 (14–26)	0.3–0
ÖMSQ F3	378	3	0.546	3–30	18 (14–21)	0–0.8
TAMPA	352	17	0.764	20–58	38 (33–42)	0–0
SF12- PCS	346	6	0.760	0–100	63 (42–80)	0.6–0.3
SF12-MCS	346	6	0.755	0–100	42 (33–71)	0.3–0.6
RT ÖMSQ	90	12	0.848	5–95	55 (43–67)	0–0
RT ÖMSQ F1	90	5	0.706	0–40	16 (9–24)	2.2–0
RT ÖMSQ F2	90	4	0.697	1–35	20 (15–26)	0–0
RT ÖMSQ F3	90	3	0.622	4–30	17 (13–22)	0–1.1

The Cronbach alpha value of >0.90:excellent, >0.80:good, >0.70:acceptable, >0.60:questionable, >0.50:poor [38]. Scores were summarized as the min-max: Minimum- maximum, Median (Q1 – Q3):the median (Quartile 1 – Quartile 3).
Floor-ceiling effect: The percentage of those with the lowest and highest total scores.

Table 2. Test-retest reliability and convergent validity results.

Scales	ICC (95%CI)	SEM	ÖMSQ r_s	P-NRS r_s	TAMPA r_s	SF12-PCS r_s	SF12-MCS r_s
ÖMSQ	0.865 (0.724–0.926)	7.08	–	0.609	0.485	−0.599	−0.469
ÖMSQ F1	0.799 (0.685–0.871)	4.11	0.783	0.407	0.491	−0.545	−0.331
ÖMSQ F2	0.855 (0.737–0.915)	3.04	0.820	0.502	0.303	−0.422	−0.514
ÖMSQ F3	0.875 (0.803–0.919)	2.06	0.743	0.597	0.354	−0.444	−0.275

ICC(2,1): Intraclass correlation coefficient (two-way mixed model, single measurement, absolute agreement).
The ICC score of >0.75 indicates good reliability [40].
SEM (Standard error of measurement)=$SD\sqrt{1-ICC}$, with SD representing the standard deviation of the measure.
r_s: All Spearman rho correlation coefficients were statistically significant ($p < 0.05$). Evaluation criteria for the r_s [41]: 0.00–0.29:Negligible, 0.30–0.49:Low, 0.50–0.69:Moderate, 0.70–0.89:High, 0.90–1.00:Very high correlation.

The ICC values calculated to evaluate the reproducibility were 0.80–0.88 for the total and subscales of the ÖMSQ-12-TR (Table 2). The scale had good reliability (>0.75). The Spearman rho correlation coefficient between ÖMSQ-12-TR and P-NRS, TAMPA, SF12-PCS, and SF12-MCS scales were examined for convergent validity. The positive correlation was determined between the ÖMSQ-12-TR and the P-NRS score ($p < 0.05$). All correlation coefficients were statistically significant and the evaluation according to the criteria for the coefficients are given in Table 2.

The high ÖMSQ-12-TR score of the individuals reveals that the risk of absenteeism from work increases. When the total ÖMSQ-12-TR scores of individuals who received the medical certificates for their pain and did not go to work and who did not receive a medical certificate were compared, a statistically significant difference was observed ($p < 0.001$). While the median ÖMSQ-12-TR score of the individuals who received a medical certificate was 66 (Q1:53–Q3:76), the median score of those who could not receive a medical certificate was 50 (Q1:37–Q3:60) (Supplementary Table 3). The AUC for the risk of absenteeism from work was obtained as 0.738 (95%CI: 0.675–0.800) ($p < 0.001$). The risk of absenteeism from work by receiving a medical certificate was high in individuals with an ÖMSQ-12-TR score of ≥57.5/120 (Figure 2). The sensitivity, specificity and accuracy were respectively 70.2% (95%CI: 65.7–74.8%), 69.7% (95%CI: 65.1%–74.3%), and 69.8% (95%CI: 65.2%–74.4%).

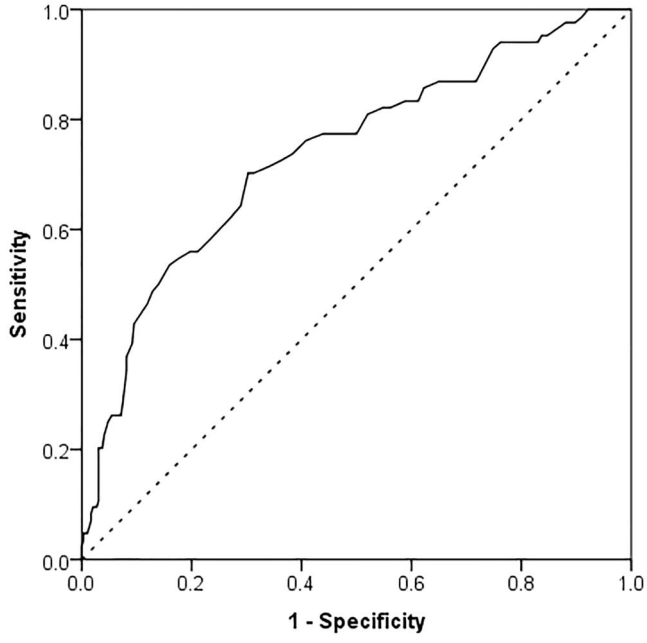

Figure 2. Receiver operating characteristic curve for ÖMSQ-12-TR total score (AUC = 0.738).

Discussion

The current study revealed that the ÖMSQ-12-TR is a valid and reliable multidimensional scale for assessing the risk of absenteeism from work caused by musculoskeletal problems in the Turkish population. The cutoff value of the scale was 57.5/120, and the risk of absenteeism and receiving a medical certificate was high in individuals with a value higher than this.

Musculoskeletal disorders are the leading cause of employee absenteeism, presenteeism, lower quality of life, occupational change, and higher medical expenses owing to disability [42,43]. The absenteeism caused by work-related musculoskeletal disorders is predicted to account for 21%–39% of the days of absence from the workplace for all occupational diseases in the European settings of the United Kingdom, Netherlands, Norway, and Germany [43,44]. Determining the risk factors related to the musculoskeletal problem is of great importance in developing intervention programs against them [43]. In this study, the participants suffered mostly from low back, neck and, general back pains. In a systematic review on global predictions of rehabilitation needs based on the Global Burden of Disease study 2019, LBP was one of the most common musculoskeletal disorders that have the highest disease burden with 568 million individuals and 64 million years of life lived with disabilities [2]. More than half of the participants in this study were actively working. The rate of participants who were absent from work due to a musculoskeletal problem was 22.2%. The results of our study were consistent with the results of studies on musculoskeletal disorders published in literature. The participants also had similar characteristics (21%–39%) with those in the studies in Europe [43,44].

According to the expert opinions on the language and content validity in the study, each item was suitable for the theoretical base of the questionnaire asked and was understandable. In the first expert assessment, Items 4, 6, 7, 11, and 12 required a revision. After the necessary revisions were completed in accordance with the experts' advice on understandability, the questionnaire was at a good-very good level (>80%), which is above the level of acceptance recommended by Davis [45].

The factor analysis is a common approach used to assess the construct validity of a questionnaire [46]. In the study CFA was used to assess the internal construct validity, which is advocated [46,47] and more advantageous than traditional statistical techniques such as correlation and regression analyses [48]. The biggest advantage of CFA is that it determines the latent variables. This allows the ability to predict the measurement error and provide more accurate, reliable, and valid predictions on the relationships between the latent structures, which in turn increases the statistical strength [48]. When the theoretical base established in this study was tested and the data obtained were compared with the reference values in literature, it was concluded that the model was well fit and fit indices were adequate [49]; however, the

factor loading of Item 1 was quite low (0.095). Most of the participants (62%) had their musculoskeletal pain or problem for >1-year, which may be the reason for low factor loading. The related item was not removed from the scale in order not to disrupt the integrity, and because the onset time of the present pain/problem is of critical clinical importance in the development of intervention programs [50].

The internal consistency or reliability of numerous items, measurements, or ratings is measurable by Cronbach's α [51], where the whole questionnaire and its sub-dimensions were calculated with the scores between 0 and 1 where the ideal range is $\alpha = 0.70–0.95$ where high values reveal that items measure the same dimension and values >0.95 indicate item redundancy [52]. For the ÖMSQ-12-TR $\alpha = 0.810$ (good) with subdimensions as: 0.719 (acceptable) for physical, fear-avoidance, and satisfaction; 0.695 (questionable) for psyche and other; and 0.546 (poor) for problem. The α coefficient for the whole questionnaire was 0.75 in the original study, but there was no information for the subdimensions [14]. The Cronbach's alpha of the Hindi version was similar, 0.85 [53]. In the Japanese version, no α values was presented for the questionnaire and its subdimensions [23]. Based on the α values for the ÖMSQ-12-TR, it can be interpreted that the questionnaire is sufficiently reliable.

When a large number of participants get the best/maximum or worst/minimum score, the measure becomes unable to distinguish between subjects at either end of the scale, resulting in ceiling and floor effects [54,55]. According to McHorney and Tarlov, the ceiling and floor effect of the scales assessing the health-related levels of individuals must be at a minimum in order to identify and distinguish between the individuals functional levels [54]. A ceiling and floor effect <15% are the generally accepted limit, though recommended to be <10% or <5% by some authors [56,57]. The ceiling and floor effects of ÖMSQ-12-TR and its subdimensions were assessed and found within the desired limits (<5%). The ceiling and floor effects were not calculated in other studies [23,53].

The test-retest reliability of the questionnaire was determined with the ICC data of the whole questionnaire and its subdimensions. Although the ICC values revealed good reliability (0.80–0.88), they were low compared with the original questionnaire (0.95). The ICC value for the Hindi version was 0.84 and indicating similar reliability [53]. In the Japanese version, ICC data of each item were presented (range 0.71–0.99) and the total score (0.92). The difference in questionnaire repeatability may be due to differences in data collection methods between the studies. This includes: the sample size, as both the original and Japanese studies were smaller samples and potentially had greater homogeneity; completion method, where face-to-face data collection was used in the two previous studies it was performed online in this study due to the COVID-19 pandemic conditions; and that the Snowball sampling method used included recruitment by participant referral, as opposed to medical referral only in both previous studies.

Convergent validity, a subtype of construct validity, confirms whether the scores of a tool assessed were significant or not compared with the scores of the other related tools, which must be at a pre-specified level. Convergent validity was tested with tools that measured the concepts within the ÖMSQ-12-TR subdimensions (P-NRS for pain, TAMPA for Kinesiophobia, and SF-12 for general health). Similar approaches were used in the other version studies of the ÖMSQ-12 and the ÖMSQ-21 [53,58,59]. Total ÖMSQ-12-TR scores had a correlation (low-high) with each of the subdimension scales. The highest correlation was between the P-NRS score and 'problem' subdimension (0.597). Similarly, in the Hindi version, the correlation between the total score of the questionnaire and the P-NRS score was high (0.632) [53]. TAMPA and SF-12-PCS reveal the highest correlation with 'physical, fear avoidance and satisfaction' subdimension (0.491 and 0.545 respectively). SF-12-MCS showed the highest correlation with 'psyche and other' subdimension (0.514). The correlation between the related subdimensions of ÖMSQ-12-TR and scales assumed to measure the same concept was at an acceptable level (>0.50) [60].

Determining a convenient cutoff value is of great importance for effective test use [61]. An ÖMSQ-12-TR cutoff score ≥57.5/120 in the study was associated with the risk of work absenteeism (getting a medical certificate). The cutoff value was 57 in the original questionnaire, which parallels the findings in this study. The AUC for the ROC curve is an accuracy measure of a quantitative diagnostic test [62] and acceptable for the ÖMSQ-12-TR (AUC = 0.738) in screening the absenteeism from work [63]. According to the likelihood ratio obtained with the sensitivity and specificity values, the risk of receiving a medical certificate was 2.32 times higher in individuals with a score >57.5, compared to 4.6 in the original study [14]. Although these screening findings (sensitivity + specificity = 1.4) were lower than the original study, they are acceptable as they approximate the recommend target of an effective test (>1.5) [64].

Limitations and strengths of the study

The study limitations include the participant's 'Problem' onset time, which was chronic for most participants, and the snowball sampling method may have affected this as participants are most likely to recruit those they know who are more likely to be in a similar situation which increases homogeneity and reduces diversity of the sample. Also, this study was performed online which reduced face-to-face contact and consequently diversified the sampling approach from previous studies. However, this can also be study strength as it is the first reported ÖMSQ-12 online use which verified the questionnaires capacity to be used remotely. Further, the questionnaire was translated into the globe's 15th most geographically used language, which expands its potential application, and the positive results for validity and reliability reinforce the findings of previous studies.

Conclusion

The ÖMSQ-12-TR is a successful translation and cultural adaptation of the original English language tool, expanding its potential global use. It was shown as valid and reliable for use in determining the risk of absenteeism in individuals with musculoskeletal problems. The questionnaire is multi-dimensional and convenient for online use, which will attract the attention of researchers of future screening related studies, particularly during these times of required increase in the popularity and use of tele-rehabilitation and the need for increased distance assessment methods. It is recommended for clinicians and for researchers in the field of musculoskeletal risk screening, as well as those performing language translation and cultural adaptation of such tools. Further research will need to consider the onset time of the participants' disorder and to increase the diversity of sampling and recruitment methods.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Author note

The authors declare that they have no conflict of interest. The authors received no specific funding for this work. The authors confirm that the manuscript has been read and approved by all authors and that there are no other individuals who satisfied the criteria for authorship but are not listed.

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Emel Tasvuran Horata http://orcid.org/0000-0002-2471-3713
Pervin Demir http://orcid.org/0000-0002-6652-0290
Gözde Yağcı http://orcid.org/0000-0002-4603-7162
Suat Erel http://orcid.org/0000-0001-7076-7651
Fatma Eken http://orcid.org/0000-0003-2975-7480
Charles Philip Gabel http://orcid.org/0000-0001-8354-4545

References

[1] Palazzo C, Ravaud JF, Papelard A, et al. The burden of musculoskeletal conditions. PLOS One. 2014;9(3):e90633.
[2] Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. Lancet. 2020;396(10258):1204–1222.
[3] Cieza A, Causey K, Kamenov K, et al. Global estimates of the need for rehabilitation based on the global burden of disease study 2019: a systematic analysis for the global burden of disease study 2019. Lancet. 2020;396(10267):2006–2017.
[4] Demou E, Smith S, Bhaskar A, et al. Evaluating sickness absence duration by musculoskeletal and mental health issues: a retrospective cohort study of Scottish Healthcare Workers. BMJ Open. 2018;8(1):e018085.
[5] Amaro J, Magalhães J, Leite M, et al. Musculoskeletal injuries and absenteeism among healthcare professionals-ICD-10 characterization. PLOS One. 2018;13(12):e0207837.
[6] March L, Smith EUR, Hoy DG, et al. Burden of disability due to musculoskeletal (MSK) disorders. Best Pract Res Clin Rheumatol. 2014;28(3):353–366.
[7] Hsu Y-Y. 0243 Screening and disability prevention for musculoskeletal disorders of high-tech industry workers in Taiwan. Occup Environ Med. 2014;71(Suppl 1):A92–A92.
[8] Hill JC, Whitehurst DG, Lewis M, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT back): a randomised controlled trial. Lancet. 2011;378(9802):1560–1571.
[9] Kuorinka I, Jonsson B, Kilbom A, et al. Standardised nordic questionnaires for the analysis of musculoskeletal symptoms. Appl Ergon. 1987;18(3):233–237.
[10] Hill JC, Dunn KM, Lewis M, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. Arthritis Rheum. 2008;59(5):632–641.
[11] Shiri R, Heliövaara M, Ahola K, et al. A screening tool for the risk of disability retirement due to musculoskeletal disorders. Scand J Work Environ Health. 2018;44(1):37–46.
[12] Gabel CP, Melloh M, Yelland M, et al. Predictive ability of a modified örebro musculoskeletal pain questionnaire in an acute/subacute low back pain working population. Eur Spine J. 2011;20(3):449–457.
[13] Gabel CP, Burkett B, Neller A, et al. Can long-term impairment in general practitioner whiplash patients be predicted using screening and patient-reported outcomes? Int J Rehabil Res. 2008;31(1):79–80.
[14] Gabel CP, Burkett B, Melloh M. The shortened örebro musculoskeletal screening questionnaire: evaluation in a work-injured population. Man Ther. 2013;18(5):378–385.
[15] Gabel CP, Melloh M, Burkett B, et al. The örebro musculoskeletal screening questionnaire: validation of a modified primary care musculoskeletal screening tool in an acute work injured population. Manual Therapy. 2012;17(6):554–565.
[16] Linton SJ, Halldén K. New Zealand acute low back pain guide patient information screening questionnaire 1996. [cited 23 August 2004]. Available from: https://www.health-navigator.org.nz/media/1006/nz-acute-low-back-pain-guide-acc.pdf.
[17] Linton SJ, Boersma K. Early identification of patients at risk of developing a persistent back problem: the predictive validity of the örebro musculoskeletal pain questionnaire. Clin J Pain. 2003;19(2):80–86.
[18] Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000;25(24):3186–3191.
[19] Naderifar M, Goli H, Ghaljaie F, editors. Snowball sampling: a purposeful method of sampling in qualitative research. SDME. 2017;14(3):e67670.
[20] Yilmaz ÖT, Yakut Y, Uygur F, et al. Tampa kinezyofobi ölçeği'nin türkçe versiyonu ve test-tekrar test güvenirliği. Fizyoterapi Rehabilitasyon. 2011;22(1):44–49.
[21] Soylu C, Kütük B. SF-12 yaşam kalitesi ölçeği'nin türkçe formunun güvenirlik ve geçerlik çalı şması . Türk Psikiyatri Dergisi. 2022;33(2):108–116.
[22] Yeşilyurt M, Faydalı S. Ağrı değerlendirmesinde tek boyutlu ölçeklerin kullanımı. Anadolu Hemşirelik ve Sağlık Bilimleri Dergisi. 2020;23(3):444–451.
[23] Veirman E, Van Ryckeghem DML, De Paepe A, et al. Multidimensional screening for predicting pain problems in adults: a systematic review of screening tools and validation studies. Pain Rep. 2019;4(5):e775.
[24] Takasaki H, Gabel CP. Cross-cultural adaptation of the 12-item örebro musculoskeletal screening questionnaire to japanese (ÖMSQ-12-J), reliability and clinicians' impressions for practicality. J Phys Ther Sci. 2017;29(8):1409–1415.
[25] Kori S. Kinisophobia: a new view of chronic pain behavior. Pain Manage. 1990;3:35–43.
[26] Crombez G, Vlaeyen JW, Heuts PH, et al. Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability. Pain. 1999;80(1–2):329–339.
[27] Kahl C, Cleland JA. Visual analogue scale, numeric pain rating scale and the McGill pain questionnaire: an overview of psychometric properties. Phys Ther Rev. 2005;10(2):123–128.
[28] Tabachnick BG, Fidell LS. SAS for windows workbook for tabachnick and fidell using multivariate statistics. New York (NY): Allyn and Bacon; 2001.
[29] Bujang M, Baharum N. Sample size guideline for correlation analysis. World J Soc Sci Res. 2016;3(1):37–46.
[30] Bujang M, Baharum N. A simplified guide to determination of sample size requirements for estimating the value of

intraclass correlation coefficient: a review. Arch Orofac Sci. 2017;12(1):1–11.
[31] Team RC. R: a language and environment for statistical computing 2021; 2021. Available from: https://www.R-project.org/.
[32] Revelle W. psych: Procedures for personality and psychological research Northwestern University, Evanston, Illinois, USA2020; Version = 2.0.12. Available from:https://CRAN.R-project.org/package=psych.
[33] Rosseel Y. Lavaan: an R package for structural equation modeling and more. Version 0.5–12 (BETA). J Stat Soft. 2012;48(2):1–36.
[34] Epskamp S. semPlot: Path diagrams and visual analysis of various SEM Packages' Output. R package version 1.1.2. 2019. Available from: https://CRAN.R-project.org/package=semPlot.
[35] Davis L. Instrument review: getting the most from a panel of experts. Appl Nurs Res. 1992;5(4):194–197.
[36] Yusoff MSB. ABC of content validation and content validity index calculation. EIMJ. 2019;11(2):49–54.
[37] Hair J, Black W, Babin B, et al. Exploratory factor analysis. In: Multivariate data analysis, 7th pearson new international ed. Harlow, UK: Pearson Education; 2014.
[38] George D, Mallery P. SPSS for windows step by step: a simple guide and reference. 11.0 update, 2003. Boston: Allyn & Bacon; 2003.
[39] Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42.
[40] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155–163.
[41] Mukaka M. Statistics corner: a guide to appropriate use of correlation in medical research. Malawi Med J. 2012;24(3):69–71.
[42] Habib MM, Yesmin, S, Moniruzzaman . A pilot study of prevalence and distributions of musculoskeletal symptoms (MSS) among paper based office workers in Bangladesh. Work. 2015;50(3):371–378.
[43] Tveter AT, Øiestad BE, Rysstad TL, et al. Risk assessment for prolonged sickness absence due to musculoskeletal disorders: protocol for a prospective cohort study. BMC Musculoskelet Disord. 2020;21(1):326.
[44] Luger T, Maher CG, Rieger MA, et al. Work-break schedules for preventing musculoskeletal symptoms and disorders in healthy workers. Cochrane Database Syst Rev. 2019;7(7):CD012886.
[45] Zamanzadeh V, Ghahramanian A, Rassouli M, et al. Design and implementation content validity study: development of an instrument for measuring patient-centered communication. J Caring Sci. 2015;4(2):165–178.
[46] Yasir A. Cross cultural adaptation & psychometric validation of instruments: step-wise. Int J Psychiatry. 2016;1(1):1–4.
[47] Schmitt TA. Current methodological considerations in exploratory and confirmatory factor analysis. J Psychoeduc Assess. 2011;29(4):304–321.
[48] Gallagher MW, Brown TA. Introduction to confirmatory factor analysis and structural equation modeling. In: Handbook of Quantitative Methods for Educational Research: Brill Sense. 2013. p. 287–314.
[49] DiStefano C, Hess B. Using confirmatory factor analysis for construct validation: an empirical review. J Psychoeduc Assess. 2005;23(3):225–241.
[50] Oliveri M. [Physical therapy of patients in general practice. role of the main groups of physical therapy methods in acute and chronic problems of the locomotor system]. Schweizerische Rundschau Fur Medizin Praxis = Revue Suisse de Medecine Praxis. 1989;78(11):281–289.
[51] Bujang MA, Omar ED, Baharum NA. A review on sample size determination for Cronbach's alpha test: a simple guide for researchers. Malays J Med Sci. 2018;25(6):85–99.
[52] Leontitsis A, Pagge J. A simulation approach on cronbach's alpha statistical significance. Math Comput Simul. 2007;73(5):336–340.
[53] Khan S, Agrawal R, Bhalerao S. Development of the Hindi version of the Orebro musculoskeletal screening questionnaire–12 items-cross cultural adaptation, validity and test–retest reliability in patients with musculoskeletal dysfunction. Indian J Physiother Occup Ther. 2021;15(4):127–135.
[54] McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res. 1995;4(4):293–307.
[55] Stucki G, Liang MH, Stucki S, et al. Application of statistical graphics to facilitate selection of health status measures for clinical practice and evaluative research. Clin Rheumatol. 1999;18(2):101–105.
[56] Hung M, Voss MW, Bounsanga J, et al. Psychometrics of the patient-reported outcomes measurement information system physical function instrument administered by computerized adaptive testing and the disabilities of arm, shoulder and hand in the orthopedic elbow patient population. J Shoulder Elbow Surg. 2018;27(3):515–522.
[57] Hung M, Hon SD, Franklin JD, et al. Psychometric properties of the PROMIS physical function item bank in patients with spinal disorders. Spine. 2014;39:158–163.
[58] Langenfeld A, Bastiaenen C, Brunner F, et al. Validation of the Orebro musculoskeletal pain screening questionnaire in patients with chronic neck pain. BMC Res Notes. 2018;11(1):161.
[59] Cuesta-Vargas AI, González-Sánchez M. Spanish version of the screening örebro musculoskeletal pain questionnaire: a cross-cultural adaptation and validation. Health Qual Life Outcomes. 2014;12:157.
[60] 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(4):659–670.
[61] Habibzadeh F, Habibzadeh P, Yadollahie M. On determining the most appropriate test cut-off value: the case of tests with continuous results. Biochem Med. 2016;26(3):297–307.
[62] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837–845.
[63] Hosmer DW Jr, Lemeshow S, Sturdivant RX. Applied logistic regression. Vol. 398. Newyork (NY): John Wiley & Sons; 2013.
[64] Power M, Fell G, Wright M. Principles for high-quality, high-value testing. Evid Based Med. 2013;18(1):5–10.