

COMPARATIVE EVALUATION OF THE PSYCHOMETRIC PROPERTIES OF ORTHOPEDIC SCALES FOR LOW BACK PAIN

AVALIAÇÃO COMPARATIVA DAS PROPRIEDADES PSICOMÉTRICAS DE ESCALAS ORTOPÉDICAS PARA DOR LOMBAR

EVALUACIÓN COMPARATIVA DE LAS PROPIEDADES PSICOMÉTRICAS DE ESCALAS ORTOPÉDICAS PARA DOLOR LUMBAR

GABRIEL JUSTI SCHMIDT,¹ ANA PAULA BONILAURO FERREIRA,² RICARDO PIETROBON,³ JOÃO RICARDO NICKENIG VISSOCI,⁴ RICARDO ANDRÉ ACÁCIO DOS SANTOS,¹ RODRIGO FETTER LAUFFER,¹ CARLOS HENRIQUE MAÇANEIRO¹

1. Orthopedics and Traumatology at the Instituto de Ortopedia e Traumatologia de Joinville/ Hospital Municipal São José - Joinville, SC, Brazil.

2. Department of Scientific Research, Instituto de Ortopedia e Traumatologia de Joinville (IOT), Joinville, SC, Brazil.

3. University of Maryland, College Park, MD, USA.

4. Duke University, Department of Surgery, Emergency Medicine Division, Durham, North Carolina, USA.

ABSTRACT

Objective: To evaluate the reliability, response capacity and validity of four scales for low back pain and correlate these scales with each other and the Self-Administered Comorbidity Questionnaire (SACQ). **Methods:** We evaluated the psychometric properties of four previously selected scales for low back pain: the Roland-Morris Disability Questionnaire (RMDQ), the Quebec Back Pain Questionnaire (QBPDS), the Waddell Disability Index (WDI) and the Back Pain Functional Scale (BPFS) and Self-Administered Comorbidity Questionnaire (SACQ) comorbidity scale. Exploratory and confirmatory factor analyses were conducted. Reliability and internal consistency were measured by Cronbach's alpha. Validity was measured through correlation of the scales with the Self-Administered Comorbidity Questionnaire comorbidity scale and an analysis of the structural equations between them. **Results:** The scales showed adequate indicators based on the factor structure and showed Kaiser-Meyer-Olkin values above 0.90. After the exploratory factor analysis, all scales showed fit indicators suited to a factor model, following the same pattern as the original validations. Similarly, they showed good internal consistency (Cronbach's alpha greater than .78). The only scale that showed factor loadings suggesting the exclusion of any item was the Roland-Morris. In terms of validity, the scales showed positive correlation coefficients similar to the Self-Administered Comorbidity Questionnaire and between them. **Conclusion:** Regarding the scales evaluated, they showed similar indications of reliability and internal consistency, such that we did not find sufficient evidence to indicate one scale over another. **Level of Evidence I; Diagnostic studies – Investigation of a diagnostic test.**

Keywords: Low back pain; Reliability; Factor analysis.

RESUMO

Objetivo: Avaliar a confiabilidade, capacidade de resposta e validade de quatro escalas para dor lombar e correlacionar essas escalas entre si e com a escala Self-Administered Comorbidity Questionnaire (SACQ). **Métodos:** Foram avaliadas as propriedades psicométricas de quatro escalas para dor lombar previamente selecionadas: Roland-Morris Disability Questionnaire (RMDQ), Quebec Back Pain Disability Scale (QBPDS), Waddell Disability Index (WDI) e Back Pain Functional Scale (BPFS) e a escala de comorbidades Self-Administered Comorbidity Questionnaire. Foram realizadas análises fatoriais exploratória e confirmatória, a confiabilidade e consistência interna foram medidas através de alfa de Cronbach e a validade através da correlação de escalas com a escala de comorbidades Self-Administered Comorbidity Questionnaire e através da análise das equações estruturais entre elas. **Resultados:** As escalas apresentaram indicadores adequados com base na estrutura fatorial e mostraram valores Kaiser-Meyer-Olkin acima de 0,90. Após a análise fatorial exploratória, todas as escalas apresentaram indicadores de aptidão adequados para um modelo de fator seguindo o mesmo padrão que as validações originais. Do mesmo modo, apresentaram boa consistência interna (alfa de Cronbach superior a 0,78). A única escala que apresentou cargas fatoriais que sugeriam a exclusão de algum item foi a Roland-Morris. Em relação à validade, as escalas apresentaram coeficientes de correlação positiva semelhantes à escala Self-Administered Comorbidity Questionnaire e entre si. **Conclusão:** Quanto às escalas avaliadas, essas apresentaram indicadores de confiabilidade e consistência interna semelhantes, de modo que não encontramos evidências suficientes para indicar uma escala sobre a outra. **Nível de Evidência I; Estudos diagnósticos–Investigação de um exame para diagnóstico.**

Descritores: Lombalgia; Confiabilidade (Epidemiologia); Análise fatorial.

RESUMEN

Objetivo: Evaluar la confiabilidad, capacidad de respuesta y validez de cuatro escalas para dolor lumbar y correlacionar estas escalas entre sí y con la escala Self-Administered Comorbidity Questionnaire (SACQ). **Métodos:** Fueron evaluadas las propiedades de cuatro escalas para dolor lumbar previamente seleccionadas: Roland-Morris Disability Questionnaire (RMDQ), Quebec Back Disability Scale (QBPDS), Waddell Disability Index (WDI) y Back Pain Functional Scale (BNPFS) y la escala de comorbilidades Self-Administered Comorbidity Questionnaire. Fueron realizados análisis factoriales exploratorio y confirmatorio, fueron medidas la confiabilidad y consistencia interna a través de

Study conducted at the Instituto de Ortopedia e Traumatologia (IOT), Joinville, Brazil.

Correspondence: Ana Paula Bonilauri Ferreira. Odontology Department, Universidade da Região de Joinville (UNIVILLE). Rua Paulo Malschitzki, 10, Joinville, SC, Brazil. 89219-501.

apbonilauri@gmail.com, ana.p@univille.net



alfa de Cronbach y la validez a través de la correlación con la escala de comorbilidades Self-Administered Comorbidity Questionnaire, y a través del análisis de las ecuaciones estructurales entre ellas. Resultados: Las escalas presentaron indicadores adecuados con base en la estructura de factorial y mostraron valores Kaiser-Meyer-Olkin por encima de 0,90. Después del análisis factorial exploratorio, todas las escalas presentaron indicadores de aptitud adecuados para un modelo de factor siguiendo el mismo patrón que las validaciones originales. Del mismo modo, presentaron buena consistencia interna (alfa de Cronbach mayor que 0,78). La única escala que presentó cargas factoriales que sugerían la exclusión de algún ítem fue la Roland-Morris. Con relación a la validez, las escalas presentaron coeficientes de correlación positiva similares a la escala Self-Administered Comorbidity Questionnaire y entre sí. Conclusión: Cuanto a las escalas evaluadas, éstas presentaron indicadores de confiabilidad y consistencia interna semejantes, por lo que no encontramos evidencias suficientes para indicar una escala sobre otra. **Nivel de evidencia I; Estudios diagnósticos - Investigación de un examen para diagnóstico.**

Descriptor: Dolor de la región lumbar; Fiabilidad; Análisis factorial.

INTRODUCTION

Low back pain is currently an international health issue. It is estimated that 70% of the population in developed countries will have this symptom at some point in their lives.¹⁻³ Among the most frequent causes of medical care, low back pain is second only to upper respiratory tract disorders.^{4,5} It is the most common cause of work disability in the United States in people under 45 years of age.^{6,7} According to US statistics, the individual cost per person with low back pain is \$8000 and the total annual cost of this condition ranges from 38 to 50 billion dollars.⁸

Self-reporting low back pain questionnaires, such as the Roland-Morris Disability Questionnaire (RMDQ),⁹ the Quebec Back Pain Disability Scale (QBPDS),¹⁰ the Waddell Disability Index (WDI)¹¹ and the Back Pain Functional Scale (BPFSS)¹² are used routinely in medical clinics and clinical studies, as they allow patient conditions to be evaluated before and after a given treatment, monitoring the course of the disease, in addition to allowing a comparison of results in multicenter studies.¹³⁻¹⁵ The psychometric properties of these low back pain scales have been tested with outpatient subjects and published. However, to the best of our knowledge, this is the first study to compare the psychometric properties of these scales in the general US population presenting symptoms of low back pain using the Amazon Mechanical Turk (MTurk) online platform. MTurk is a website that contains a simplified study design for data collection and extensive participant recruiting, that is, the main requirement for conducting research.¹⁶

The objective of this study is to compare the psychometric properties of four self-administered questionnaires for low back pain via the Mechanical Turk platform and to normalize the results comparing the scales with each other and with the Self-Administered Comorbidity Questionnaire (SACQ).

METHODS

This study was approved by the Institutional Review Board of the Faculdade Uningá (Maringá - PR) as protocol number FR 489024. The informed consent form was presented as the first page of the online questionnaire with a description of the study and its objective. Subsequently there were two alternatives: "Yes, I accept" or "No, I don't accept". The study participants agreed to sign the informed consent form. If the respondent did not agree with the informed consent form, he/she was automatically taken to the end of the survey.

Data collection was conducted through the MTurk website (<http://www.mturk.com>), which allows research data to be obtained in a short amount of time at a low cost. It functions through HITs (human intelligence tasks),¹⁷ which are tasks created by requesters to be completed by paid task workers.¹⁶ After completing the questionnaire the worker receives financial compensation, usually less than one US dollar.¹⁸

Studies conducted using MTurk suggest that the participants are internally motivated to participate in the tasks and not by the monetary compensation. The quality of data obtained through MTurk meets or exceeds the psychometric standards associated with published research.¹⁶

A total of 395 participants filled out a demographic questionnaire, four questionnaires specific to low back pain, and one addressing comorbidities, which will be described below. We used the Qualtrics (<http://www.qualtrics.com>) internet survey software to load the questionnaires from MTurk. The participants accessed a link that took them directly to Qualtrics. A filter was programmed so that only participants with low back pain could participate. We also advised that the results would be used in medical research and suggested that the answers be honest. Qualtrics collects the responses and formats the data in an Excel spreadsheet.

LOW BACK PAIN QUESTIONNAIRES

Roland-Morris Disability Questionnaire

This scale consists of 24 Yes / No questions related to physical functions in order to specifically evaluate the disability caused by the low back pain and how much it has affected the individual within the past 24 hours. It is a self-administered questionnaire and can be completed in less than 5 minutes. In the scale, 1 point is assigned to each positive response and the final score can range from 0 (without disability) to 24 (severe disability).⁹

Quebec Back Pain Disability Scale

This 20-item specific condition scale evaluates the level of disability in daily activities in patients with low back pain. It is a self-administered scale and can be completed in 5 to 10 minutes. Each of the 20 activities of daily living is scored on a scale ranging from 0 points ("without any difficulty") to 5 points ("unable to perform"). The points are summed to obtain the disability score, which ranges from 0 to 100.¹⁰

Waddell Disability Index

This scale is comprised of 9 Yes / No questions that assess activities of daily living commonly restricted by low back pain. The final result is calculated as the sum of the positive items and varies from 0 to 9. The questionnaire is easy to administer and can be completed in about 5 minutes.¹¹

Back Pain Functional Scale

This 12-item questionnaire evaluates activities of daily living related to low back pain. Each item is scored on a 6-point scale, where 0 indicates disability and 5 without difficulty. The final score ranges from 0, representing the worst functional level, to 60, representing the highest functional level.¹²

Self-Administered Comorbidity Questionnaire

This short, easy-to-understand, self-administered questionnaire measures comorbidities. It is highly reproducible and is moderately correlated with the Charlson Index,¹⁹ a standard medical comorbidity index.²⁰

Statistical analysis

Descriptive analysis and the characteristics of all the subjects were applied through means and percentages with a confidence interval of 95%.

Psychometric properties

Reliability (Internal Consistency)

Internal consistency is an interrelationship between the items of a scale.²¹ Different items in a questionnaire may ask the same questions slightly differently to obtain the opinion or functional level of the interviewee reliably. We used Cronbach's alpha to determine internal consistency, with coefficients higher than 0.70 indicating good internal consistency.

Validity

Validity refers to the degree to which a tool measures what it is intended to measure.²² Construct validity refers to how well a tool measures when compared to tools of similar or different purpose or scope.²³ We used construct validity to find a correlation with each result. Correlation between the low back and SACQ scales was made by the correlation adjusted for normal distribution of data.

Factor Analysis (Exploratory Factor Analysis and Confirmatory Factor Analysis)

To investigate the internal structure of the scales we used exploratory factor analysis (EFA) in order to reduce the variables to variance factors. The number of factors tested was determined by the eigenvalues (numbers greater than 1.00), plot analysis, commonality, and interpretability factor (model with theoretical foundation). The EFA was conducted with the principal axis factoring method and promax rotation (oblique) and the cutoff of 0.40 was defined for factor loadings.

The models developed by EFA were tested by confirmatory factor analysis (CFA) in the form of a measurement model composed of latent variable models. This procedure evaluated the fit and adequacy model from the aptitude indicators, factor loadings, and individual reliability of each item. Maximum likelihood was the estimation method used due to the multivariate normality. The aptitude model indicators were chi-squared, the root mean square error of approximation (RMSEA) (values less than 0.05 are considered as an adequate fit), the comparative fit index (CFI) (values greater than 0.95 are accepted as a good fit), the goodness-of-fit index and adjusted goodness-of-fit index (GFI/AGFI) (values greater than 0.90 are interpreted as an acceptable fit), the Tucker-Lewis index (TLI) (acceptable fit with values greater than 0.97), and Akaike information criterion, Bayesian information criterion, and the expected cross validation index (AIC/BIC/ Modified expected cross-validation index (MECVI)) (lower values indicate a better model as compared to the others).²⁴

RESULTS

Most of the participants were male ($n = 256$, 64%), Asian ($n = 229$, 57%), single ($n = 213$, 53%) and with a college education ($n = 269$, 68%) (Table 1).

All scales presented adequate indicators based on their factor structure and showed Kaiser-Meyer-Olkin (KMO) values above 0.90. The exploratory factor analysis (EFA) confirmed the original factor structure for all the scales, assessing only one latent construct – pain. Only the BPFs had eigenvalues higher than 1, which suggests the possibility of more than one factor to be retained in the EFA for all the other scales. However, the scree plot and KMO confirmed the hypothesis of only one factor for all scales. The factor structure of the BPFs presented the greatest amount of explained product variance, however, it was the only one with commonality values that suggested a problem. The only scale that presented factor loadings that suggested exclusion (commonality above 0.50 or no factor loading less than 0.40) was the RMDQ in item 2, which showed a small loading factor (<0.40). After analyzing the behavior of the items in the EFA, all the scales presented adequate aptitude indicators for a factor model, following the same model as the original validations (Table 2).

All the scales had good internal consistency (Cronbach's alpha greater than 0.78), indicating high consistency in the response pattern. Considering the validity, all the scales had similar positive moderate correlation coefficients (0.45-0.47), with positive strong correlations between the scores of the scales (Table 3).

Table 1. Demographics.

Demographics	N = 395 (%)
Age (mean + SD)	31.89±11.22
Sex	
Male	256 (64.8)
Female	139 (35.2)
Race	
Caucasian	120 (30.3)
Afro-American	8 (2)
Asian	229 (57.9)
Native American	6 (1.5)
Oceanic	2 (0.5)
Other	29 (7.3)
Married status	
Single	213 (53.9)
Married	172 (43.5)
Separated	8 (2)
Widowed	1 (0.3)
Education level	
Preschool Incomplete	1 (0.3)
Preschool Complete	2 (0.5)
Elementary School Incomplete	1 (0.3)
Elementary School Complete	3 (0.7)
High School Incomplete	5 (1.2)
High School Complete	39 (9.8)
College Incomplete	74 (18.7)
College Complete	269 (68.1)

DISCUSSION

To our knowledge, this is the first study to compare the psychometric properties of four questionnaires used for self-assessment of low back pain in a generic population through Mechanical Turk and normalizing the results, comparing the scales against each other and with the SACQ. In general, the four scales evaluated demonstrated adequate internal consistency, reliability and validity to evaluate patients with low back pain.

All the scales studied presented indications for the same original factor structure, that is, a single construct. However, the factor analyses conducted in our study showed that all the scales presented indications that there could be more factors. The same was reported in studies of the WDI, in which the analysis of the components did not succeed in demonstrating a single satisfactory construct. According to the author, four measures of motion could be combined, but the 4-factor factorial structure was weak.¹¹⁻²⁵ The QBPDS scale validated for Hungarian²⁶ and Greek²⁷ had an EFA performed and found four and six factors, respectively. We did not find articles in the literature that had conducted an EFA of the RMDQ and BPFs scales. No article cites performing CFA, which we performed for all the scales in our study.

The reliability of the scales was measured by the Cronbach's alpha coefficient, considered adequate for values above 0.7. The results obtained in our study ranged from 0.78 (WDI) to 0.93 (BPFs). A previous study that evaluated the psychometric properties of the RMDQ scale found a similar value (0.96).²⁸ Similarly, studies that translated and validated the RMDQ scale into other languages obtained values close to ours (0.91). In the Persian translation the value found was 0.83;²⁹ for the German, 0.81;³⁰ Spanish, 0.84;³¹ Turkish, 0.85;³² Greek, 0.88;³³ and Japanese, 0.86;³⁴ thus confirming the homogeneity of the items of this scale. Articles that translated the QBPDS scale to Dutch,³⁵ Turkish³⁶ and Arabic³⁷ found values between 0.92 and 0.95, that is, values close to ours (0.92). As for the WDI scale, it obtained the lowest internal consistency value (0.78) as compared to the others. This finding corroborates a previous study that reported Cronbach's alpha values of between 0.86 and 0.96.²⁵

Table 2. Exploratory Factor Analysis and Confirmatory Factor Analysis Values.

	Roland Morris Disability Questionnaire	Quebec Back Pain Disability Scale	Waddell Disability Index	Back Pain Functional Scale
EFA				
KMO	0.939	0.943	0.834	0.935
Eigenvalues*	8.23, 1.51, 1.19	8.05, 1.22, 1.18	3.30, 1.02	6.86
Scree Plot**	1	1	1	1
Commonalities	0.43 - 0.13	0.49 - 0.23	0.36 - 0.18	0.65 - 0.32
Explained Variance	0.32	0.37	0.29	0.53
Factor Loading	Q2 excluded	Without exclusions	Without exclusions	Without exclusions
CFA				
X2 (Df)	562.97 (230), p=6.4329e-30	508.36 (170) p=1.976e-35	106.84 (27) p=1.8904e-11	383.17 (54) p=3.9378e-51
GFI/AGFI	0.877 / 0.853	0.875 / 0.846	0.939 / 0.898	0.845 / 0.776
RMSEA	0.060	0.071	0.086	0.124
TLI	0.876	0.874	0.842	0.861
CFI	0.888	0.887	0.881	0.886
SRMR	0.050	0.052	0.056	0.054

* Values greater than 1. ** Number related to the Scree plot. EFA= Exploratory Factor Analysis; KMO=Kaiser-Meyer-Olkin; CFA= Confirmatory Factor Analysis; Df=Degree of Freedom; GFI=Goodness of Fit Index; AGFI=Adjusted Goodness of Fit Index; RMSEA=Root Mean Square Error of Approximation; TLI= Tucker-Lewis Index; CFI=Comparative Fit Index; SRMR=Standardized Root Mean Square Residual.

Table 3. Validity and Reliability of the four low back pain scales.

	Cronbach's alpha α (CI)	Correlation with SACQ R (P-value)
Roland Morris Disability Questionnaire	0.91 (0.89;0.92)	0.45 (0.001)
Quebec Back Pain Disability Scale	0.92 (0.91;0.93)	0.51 (0.001)
Waddell Disability Index	0.78 (0.74;0.81)	0.47 (0.001)
Back Pain Functional Scale	0.93 (0.91;0.94)	0.47 (0.001)

CI= Confidence Interval; R=Reliability.

The fact that this scale does not take work, physical and personal care activities into account and that the questions are not associated with a certain period of time might explain the lower internal consistency values. And finally, the BPFs scale had a high internal consistency value (0.93), which is in line with a prior study that also reported a value of 0.93.¹²

When we consider the psychometric property of validity by comparing the scales investigated with the SACQ comorbidity scale, our results indicate that all the scales show different correlation patterns between low back pain and several comorbidities. Making the correlations between the scales themselves, they ranged from moderate to high. In a previous study to validate the QBPDS scale for Arabic, the QBPDS was correlated with the Oswestry Disability Index (ODI) scale and the Numeric Pain Rating Scale (NPRS), showing high correlation.³⁷ This same scale translated into Persian was correlated with the RMDQ and presented an excellent correlation, QBPDS = 0.92 and RMDQ = 0.83.²⁹

In their statistical analysis of their Persian translations of both the RMDQ and QBPDS scales, Mousavi et al.²⁹ found, as we did, a relationship between low back pain and degrees of disability.

The same authors also found a significant relationship between these scales and the measurement of patient pain.

However, we cannot guarantee that there is a direct relationship to pain that the patient feels and their actual physical disability, a fact that was also observed by Maaroufi et al.²⁸ and Stratford, Binkley, Riddle,¹² who succeeded in establishing good correlation between the RMDQ scale and pain measurement ($r = 0.32$, $p=0.005$). However, they failed to associate the scale with other variables such as pain duration, for example.

As for the BPFs scale, it showed a correlation in the same direction as the disability variable.¹²

CONCLUSION

All the scales presented good indicators of reliability and internal consistency, the WDI scale set apart with the worst reliability value, even though it was still acceptable. Therefore, since the reliability was similar, we cannot indicate one scale over another based on the magnitude of absolute internal consistency values.

One of the limitations of this study was the difficulty in finding articles in the literature that explore the factor structure of scales. However, this did not interfere with our results.

A positive aspect was the use of MTurk, which allowed us to group a heterogeneous sample of participants quickly. This reinforces MTurk as an effective means of data collection for further studies.

Further studies using our sample will be conducted towards equalization of the scale values.

All authors declare no potential conflict of interest related to this article.

CONTRIBUTION OF THE AUTHORS: Each author made significant individual contributions to this manuscript. The concept and design were executed by GJS (0000-0002-7055-6018)*. The data collection and compilation were conducted by APBF (0000-0002-9885-5183)*, RP (0000-0003-2996-2193)* and JRNv (0000-0001-7276-0402)*. Data analysis and interpretation were performed by CHM (0000-0003-4838-3031)*, RFL (0000-0002-1105-486X)*, RAAS (0000-0003-2901-5583)*, GJS (0000-0002-7055-6018)* and RP (0000-0003-2996-2193)*. The manuscript was written by GJS, APBF and JRNv. All authors read and approved the final manuscript. *ORCID (Open Researcher and Contributor ID).

REFERENCES

- Jarvik JG, Comstock BA, Bresnahan BW, Nedeljkovic SS, Nerenz DR, Bauer Z, et al. Study protocol: The back pain outcomes using longitudinal data registry. *BMC Musculoskeletal Disord.* 2012;13:64.
- Seon-Yu K, Lee IS, Kim BR, Lim JH, Lee J, Koh SE, et al. Magnetic Resonance Findings of Acute Severe Lower Back Pain. *Ann Rehabil Med.* 2012;36(1):47-54.
- Chanda ML, Alvin MD, Schnitzer TJ, Apkarian AV. Pain characteristic differences between subacute and chronic back pain. *J Pain.* 2011;12(7):792-800.
- Casazza BA. Diagnosis and treatment of acute low back pain. *American Family Physician.* 2012;85(4):343-50.
- Cruz RM, Sardá Junior JJ. Diagnóstico de Aspectos Emocionais Associados á Lombalgia e á Lombociática. *Aval Psicol.* 2003;2(1):29-33.
- Bradley Junior WG. Low Back Pain. *Am J Neuroradiol.* 2007;28:990-2.
- Deyo RA, Weinstein JN. Low Back Pain. *N Engl J Med.* 2001;344(5):363-70.
- Atlas SJ, Deyo RA. Evaluating and Managing Acute Low Back Pain in the Primary Care Setting. *J Gen Intern Med.* 2001;16(2):120-31.
- Roland MO, Morris RW. A study of the natural history of back pain. Part 1: Development of a reliable and sensitive measure of disability in low back pain. *Spine (Phila Pa 1976).* 1983;8(2):141-4.

10. Kopec JA, Esdaile JM, Abrahamowicz M, Abenham L, Wood-Dauphinee S, Lamping DL, et al. The Quebec Back Pain Disability Scale: Conceptualization and Development. *J Clin Epidemiol.* 1996;42(2):151-61.
11. Waddell G, Main CJ. Assessment of Severity in Low-Back Disorders. *Spine (Phila Pa 1976).* 1984;9(2):204-8.
12. Stratford PW, Binkley JM, Riddle DL. Development and Initial Validation of the Back Pain Functional Scale. *Spine (Phila Pa 1976).* 2000;25(16):2095-102.
13. Stratford PE, Binkley JM, Riddle DL. Development and Initial Validation of the Back Pain Functional Scale. *Spine (Phila Pa 1976).* 1999;25(16):2095-102.
14. de Groot V, Beckerman H, Lankhorst GJ, Bouter LM. How to measure comorbidity: a critical review of available methods. *J Clin Epidemiol.* 2003;56(3):221-9.
15. Longo UG, Loppini M, Denaro L, Maffulli N, Denaro V. Rating scales for low back pain. *Br Med Bull.* 2010;94:81-144.
16. Buhrmester M, Kwang T, Gosling S. Amazon's Mechanical Turk A New Source of Inexpensive, Yet High-Quality. *Perspectives on Psychological Science.* 2011;6(1):3-5.
17. Mason W, Suri S. Conducting behavioral research on Amazon's Mechanical Turk. *Behav Res Methods.* 2012;44(1):1-23.
18. Rand DG. The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments. *J Theor Biol.* 2012;299:172-9.
19. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis.* 1987;40(5):373-83.
20. Sangha O, Stucki G, Liang MH, Fossel AH, Katz JN. The Self-Administered Comorbidity Questionnaire: A New Method to Assess Comorbidity for Clinical and Health Services Research. *Arthritis Rheum.* 2003;49(2):156-63.
21. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 2010;63(7):737-45.
22. Polit DF, Hungler BP. *Nursing research: Principles and methods.* 3rd edition. Philadelphia: Lippincott; 1993.
23. Michener LA, Leggin BG. A review of self-report scales for the assessment of functional limitation and disability of the shoulder. *J Hand Ther.* 2001;14(2):68-76.
24. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *MPR- Online.* 2003;8(2):23-74.
25. Waddell G, Somerville D, Henderson I, Newton M. Objective Clinical Evaluation of Physical Impairment in Chronic Low Back Pain. *Spine (Phila Pa 1976).* 1992;17(6):617-28.
26. Valasek T, Varga PP, Szoverfi Z, Kumin M, Fairbank J, Lazary A. Reliability and validity study on the Hungarian versions of the Oswestry disability index and the Quebec back pain disability scale. *Eur Spine J.* 2013;22(5):1010-8.
27. Christakou A, Andriopoulou M, Asimakopoulos PJ. Validity and reliability of the Greek version of the Quebec Back Pain Disability Scale. *J Back Musculoskelet Rehabil.* 2011;24(3):145-54.
28. Maaroufi H, Benbouazza K, Faik A, Bahiri R, Lazrak N, Abouqal R, et al. Translation, Adaptation, and Validation of the Moroccan Version of the Roland Morris Disability Questionnaire. *Spine (Phila Pa 1976).* 2007;32(13):1461-5.
29. Mousavi JS, Parnianpour M, Mehdian H, Montazeri A, Mobini B. The Oswestry Disability Index, the Roland-Morris Disability Questionnaire, and the Quebec Back Pain Disability Scale: Translation and Validation Studies of the Iranian Versions. *Spine (Phila Pa 1976).* 2006;31(14):E454-9.
30. Weisinger GF, Nuhr M, Quittan M, Ebenbichler G, Wöfl G, Fialka-Moser V. Cross-cultural adaptation of the Roland-Morris questionnaire for German-speaking patients with low back pain. *Spine (Phila Pa 1976).* 1999;24(11):1099-103.
31. Kovacs FM, Liobera J, Gil del Real MT, Abreira V, Gestoso M, Fernández C, et al. Validation of the Spanish version of the Roland-Morris questionnaire. *Spine (Phila Pa 1976).* 2002;27(5):538-42.
32. Kueukdevec AA, Tennant A, Elhan AH, Niyazoglu H. Validation of the Turkish version of the Roland-Morris Disability Questionnaire for use in low back pain. *Spine (Phila Pa 1976).* 2001;26(24):2738-43.
33. Boscainos PJ, Sapkas G, Stilianessi E, Prouskas K, Papadakis SA. Greek versions of the Oswestry and Roland-Morris Disability Questionnaires. *Clin Orthop Relat Res.* 2003;411:40-53.
34. Fujiwara A, Kobayashi N, Saiki K, Kitagawa T, Tamai K, Saotome K. Association of the Japanese Orthopaedic Association score with the Oswestry Disability Index, Roland-Morris Disability Questionnaire, and short-form 36. *Spine (Phila Pa 1976).* 2003;28(14):1601-7.
35. Schoppink LEM, van Tulder MW, Koes BW, Beurskens SA, de Bie RA. Reliability and Validity of the Dutch Adaptation of the Quebec Back Pain Disability Scale. *Phys Ther.* 1996;76(3):268-175.
36. Melikoglu AM, Kocabas H, Sezer I, Bilgiliyosoy M, Tuncer T. Validation of the Turkish Version of the Quebec Back Pain Disability Scale for Patients With Low Back Pain. *Spine (Phila Pa 1976).* 2009;34(6):E219-24.
37. Alnahhal A, May S. Validation of the Arabic Version of the Back Pain Disability Scale. *Spine (Phila Pa 1976).* 2012;37(26):E1645-50.