ELSEVIER

The Spine Journal ■ (2013) ■

THE SPINE JOURNAL

Clinical Study

Confirmatory factory analysis of the Neck Disability Index in a general problematic neck population indicates a one-factor model

Charles Philip Gabel, PhD, MDT, PT[a,*], Antonio I. Cuesta-Vargas, PhD, PT[b], Jason W. Osborne, PhD[c], Brendan Burkett, PhD[a], Markus Melloh, MD, MPH, PhD, MBA[d]

[a]*Faculty of Science, Health and Education, Centre for Healthy Activities, Sport and Exercise, University of the Sunshine Coast, Sippy Downs Dve, Sippy Downs, Sunshine Coast Qld 4556 Australia*

[b]*Department of Psychiatry and Physiotherapy, Faculty of Medicine, Malaga University, SpainAvenida Estación de el Palo, Málaga, 29017 Spain*

[c]*Educational and Counseling Psychology, University of Louisville, 2323 S Brook St, Louisville, KY, 40208 USA*

[d]*Centre for Medical Research (CMR), Faculty of Medicine, Dentistry and Health Sciences, The University of Western Australia, 35 Stirling Highway, Crawley WA 6009, Perth, Australia*

Received 2 April 2012; revised 11 July 2013; accepted 21 August 2013

Abstract

BACKGROUND CONTEXT: The Neck Disability Index frequently is used to measure outcomes of the neck. The statistical rigor of the Neck Disability Index has been assessed with conflicting outcomes. To date, Confirmatory Factor Analysis of the Neck Disability Index has not been reported for a suitably large population study. Because the Neck Disability Index is not a condition-specific measure of neck function, initial Confirmatory Factor Analysis should consider problematic neck patients as a homogenous group.

PURPOSE: We sought to analyze the factor structure of the Neck Disability Index through Confirmatory Factor Analysis in a symptomatic, homogeneous, neck population, with respect to pooled populations and gender subgroups.

STUDY DESIGN: This was a secondary analysis of pooled data.

PATIENT SAMPLE: A total of 1,278 symptomatic neck patients (67.5% female, median age 41 years), 803 nonspecific and 475 with whiplash-associated disorder.

OUTCOME MEASURES: The Neck Disability Index was used to measure outcomes.

METHODS: We analyzed pooled baseline data from six independent studies of patients with neck problems who completed Neck Disability Index questionnaires at baseline. The Confirmatory Factor Analysis was considered in three scenarios: the full sample and separate sexes. Models were compared empirically for best fit.

RESULTS: Two-factor models have good psychometric properties across both the pooled and sex subgroups. However, according to these analyses, the one-factor solution is preferable from both a statistical perspective and parsimony. The two-factor model was close to significant for the male subgroup ($p<.07$) where questions separated into constructs of mental function (pain, reading headaches and concentration) and physical function (personal care, lifting, work, driving, sleep, and recreation).

CONCLUSIONS: The Neck Disability Index demonstrated a one-factor structure when analyzed by Confirmatory Factor Analysis in a pooled, homogenous sample of neck problem patients. However, a two-factor model did approach significance for male subjects where questions separated into constructs of mental and physical function. Further investigations in different conditions, subgroup and sex-specific populations are warranted.

Keywords: Outcome; Clinimetrics; Neck; Measurement; Factor analysis

FDA device/drug status: Not applicable.

Author disclosures: ***CPG:*** Nothing to disclose. ***AC-V:*** Nothing to disclose. ***JWO:*** Nothing to disclose. ***BB:*** Nothing to disclose. ***MM:*** Nothing to disclose.

Funding and conflict of interest: There was no external funding for this study and the authors have no conflict of interests.

* Corresponding author. University of the Sunshine Coast, Centre for Healthy Activities, Sport and Exercise, Faculty of Science, Health and Education, PO Box 760, Coolum Beach, Queensland 4573, Australia. Tel.: (61) 75-446-1022; fax: (61) 75-471-7022.

E-mail address: cp.gabel@bigpond.com (C.P. Gabel)

1529-9430/$ - see front matter
http://dx.doi.org/10.1016/j.spinee.2013.08.026

Introduction

Neck problems represent a notable financial and societal burden to today's industrial countries. The individual experience varies considerably in both prevalence and severity. It may include nonspecific onset, posttrauma such as whiplash-associated disorder (WAD), postural-based causes, and a sex bias may be present [1,2]. The estimated lifetime prevalence is up to 70%, where as many as 19% report chronic symptoms at any one time [3]. Monitoring such conditions helps to track an individual's current status, to reflect the effectiveness of evidence-based care from individual- and group-based interventions, and to justify service provision. Patient-reported outcome (PRO) tools such as the Neck Disability Index (NDI) [4,5] are commonly used to measure the neck's functional status and outcome [6–9]. The NDI is also advocated in clinical guidelines by insurer groups [10,11] and professional organizations [12] for both patients with nonspecific symptomatic neck and those with WAD.

Despite this level of acceptance and support, the NDI was not developed by a clinimetric process [13] but rather by the less-robust qualitative item selection process [5,14]. The recent literature is also divided in its conclusions' regarding the NDI's factor structure. Factor structure is defined as the underlying themes present within any given PRO. It is important to know how many factors or constructs a PRO measure has, because these factors indicate whether a parsimonious structure is present [15,16]. A one-factor solution is required if the questionnaire is to be used to construct and validate total scores. These total or ''cumulative'' scores are in turn useful for assessing how severe a problem affects an individual's functional status [4].

Factor analysis is ideally performed in two stages. Initially exploratory factor analysis (EFA) is used to expose underlying traits and suggest a model of the number of factors present. The preferred methodology for EFA is maximum likelihood extraction (MLE) for normally distributed samples and principle component analysis for non-normally distributed samples. This analysis usually requires moderately large samples in the order of 150 to 200 participants or 10 to 15 responses for each item that is present [15,16]. Confirmatory factor analysis (CFA) is then ideally used to clarify and validate the suggested model and requires an even larger sample procedure in the order of >500 participants and preferably >1,000 [15,16]. For PROs in which a single summated score is used, it is critical to demonstrate a one-factor structure. This ensures that all questions within the tool are reporting on the same underlying theme or construct and one score can be used [17].

Recent articles have disagreed whether the NDI should be considered a one-factor or two-factor scale [13,18]. Unfortunately, most studies have used relatively small samples, which make them particularly unstable for EFA analysis [19] and not assessable with CFA. For example, Young et al. [20] analyzed a sample of 61 adult neck pain patients and Pickering et al. [21] analyzed a sample of 88 patients with mechanical neck pain. According to most guidelines for EFA sample size, these are inadequate and could be the cause of conflicting results, which raises questions about the robustness of the NDI factor structure and whether it should be modified by item reduction [22] or item addition to improve performance [13] and relevance [1].

To date, only one published study has analyzed the NDI with the more rigorous CFA procedure [23], but the sample was smaller than what is usually recommended [15,16]. Yet, given a large enough sample, CFA can test more definitively whether a one-factor or two-factor model is the best fit for NDI data. Consequently, the aims of this study were: to analyze the NDI factor structure using CFA, in an adequately large sample, and test the two competing models to see which is most supported by the data; and to determine whether this model is consistent across male- and female-only subgroups or whether a sex difference is present.

Methods

Participants

To collect a sufficiently large and diverse population, eight separate research groups were approached through their representative tertiary institution or primary publishing author in North America, Europe, and Oceania. The study was a secondary analysis of these pooled data where each contributor's study had been completed under the respective tertiary institutional ethics committee supervision. All data were deidentified by the original researchers, forwarded to the first author for pooling, then combined in a consistent format to provide a preliminary homogenous sample of patients with neck problems (n=1,384). The final data file was reduced to account for missing responses or missing fields (n=1,278; females=862, males=416) where 803 had nonspecific neck problems and 475 had WAD (Table 1). Of the 106 respondents excluded for missing data, there was no systematic association with any specific NDI questionnaire item. The exclusions were due a lack of item responses, 51% with all 10 questions represented across all contributors, and demographic data, 49% that included age, sex, and injury classification.

Assessment tools

The NDI is a 10-item questionnaire that requests the user to select one of six statements per question that best describes their individual status at that time [5,14]. According to some authors, the NDI considers functional impairments that can be separated into psychological constructs (eg, reading, concentration) and physical constructs (eg, lifting, driving) [6,20]. Each question has six potential responses ranging in severity from zero (no disability) to five (most severe disability) with a total score maximum of 50 points. This is multiplied by two to provide a percentage

Table 1
Demographic data of patients from the six contributing studies

Study source	Patient type	n	Female, %	Age, y, mean±SD	Study	Compensation
University of the Sunshine Coast, Queensland, Australia	Nonspecific and whiplash	106	57.5	40.8±12.8	Acute patients <6 wk	Permitted
Faculty of Medicine Malaga University, Spain	Nonspecific	111	60.4	54.8±8.4	Acute <6 wk, 42%	Not permitted
University of Queensland, Queensland, Australia	Whiplash	304	63.5	36.0±13.6	Acute whiplash patients <6 wk	Permitted
Mobile Spine and Rehabilitation, Mobile, Alabama	Whiplash	60	68.0	49.4±13.1	Acute <6 wk, 36%	Not permitted
Melbourne Whiplash Center, Victoria, Australia	Whiplash	373	66.8	40.6±11.5	Acute <6 wk, 11%	Permitted
Complejo Hospitalraio, Ciudad de Jaen, Spain	Nonspecific neck pain	324	77.5	41.3±11.0	Acute <6 wk, 17%	Not permitted
Totals		1,278	67.5	41.2±12.9		

All studies were prospective and conducted within a physiotherapy outpatients setting with participants referred by a medical practitioner as consecutive patients from a sample of convenience.

scale in which 100% indicates maximum functional loss and 0% indicates no functional loss with cut-off scores of <8 NDI-points (16%) for no disability or recovered and >28 NDI-points (56%) for moderate-to-severe disability [5,24]. The NDI has been used in more than 300 publications and is the most commonly cited outcome tool for measurement of neck-related conditions [14].

Concept-retention analysis

This study also performed a ''concept-retention analysis'' [25,26] on the NDI. This process used consensus agreement by a panel of assessors to subjectively allocate each question to a relevant construct based on the panel members' interpretation of the item-content. In this study the authors were the panel members. Two physical therapists and a rheumatologist physician served as representatives of medical users and potential administrators whereas two nonmedical academics were representative of general patients. Unanimous consensus was required, and the questions were allocated to three potential underlying constructs, physical function and psychological function, as suggested by previous authors [1,20], and pain [13]. In addition, the panel assessed whether all questions could be allocated to a single construct, that of neck associated problems, or if the three suggested categories could be reduced to two.

Factor analysis

The pooled data were analyzed through CFA by two independent assessors by using separate statistical software packages, LISREL 8.80 (Scientific Software International, Skokie, IL, USA) and IBM's AMOS 19.0 (Armonk, NY, USA). The CFA was used to calculate the factor loadings of the variables included in the model. The MLE method was used to estimate the model with a sampling of several fit indices to evaluate fit. These indices were the root mean square error of approximation (RMSEA), the comparative fit index, and the normed fit index. For the RMSEA, values ≤0.08 reflect reasonable fit, whereas ≤0.05 indicate an excellent fit [27]. For the normed fit index and comparative fit index, values vary along a continuum of 0 to 1, in which ≥0.90 is considered satisfactory and ≥0.95 is considered excellent [28]. Because the components/factors of the NDI are continuous variables and the factor loadings obtained by CFA cannot be used directly to assess the level of severity with other reference values, an NDI Reference Values (NDI_REVA) was developed.

During CFA, modification indices (Wald indices) were examined to determine whether allowing error terms to covary would significantly improve the model fit. In the one-factor model, errors with >4.00 modification index were allowed to correlate, whereas in the two-factor model errors >4.00 were allowed only if they were assigned to the same factor. Cross-factor correlations of errors are not generally consistent with a test of a two-factor model. Thus, these analyses compare the best possible model fit for both factor structures. Once the models were optimally fit to one- and two-factors they were compared statistically to determine superiority. The difference in χ^2 between the one- and two-factor solutions can be assessed as a χ^2 with df of the difference in parameters estimated between the two models. A significant difference between the two models indicates a significant difference in model fit.

Results

Concept-retention analysis

The five-person assessment panel had unanimity in agreement with allocation of questions to two broad constructs of physical and mental function [20,29]. The physical function construct contained the six items: personal care, lifting, work, driving, sleep, and recreation. The mental function construct contained four items that could be further considered within two sub-constructs: ''mental–experience'' (pain and headache) [13] and ''mental–action'' (concentration and reading). The panel also agreed all

Table 2
Median and percentiles of Neck Disability Index references values (NDI_REVA) classified into five categories

	NDI_REVA			
	Total, N=1,278	Female=861	Male=416	Threshold categories
Median (90% CI)	−0.06 (−1.35 to 1.40)	0.14 (−1.33 to 1.43)	−0.09 (−1.30 to 1.48)	
Percentile				
>20th	−0.96	−0.99	−0.91	Very low to low
>40th	−0.34	−0.33	−0.39	Low to average
>60th	0.21	0.22	0.18	Average to high
>80th	0.93	0.90	0.89	High to very high

questions could logically be representative of a single construct, ''neck problems.''

NDI_REVAs

The NDI_REVAs were calculated from standardized scores and values. These can be classified into five categories: very low, low, average, high, and very high. These are represented in Table 2.

CFAs: a one-factor or two-factor solution?

The two independently CFA analyses reached unanimous agreement (Table 3). In the first set of analyses, all participants were included. The two-factor solution is acceptable in overall model fit but significantly worse than the one-factor model (exact $p<.000000000006$). This item separation into the two factors or constructs of Physical (NDI_Phys) and Mental (NDI_Ment) are demonstrated in the Figure.

Males only versus females only

In the second set of analyses of males only (Table 3), both models fit well but a near-significant trend favored the one-factor model ($p<.07$). In the third set of analyses of females only, again both models fit well, but there was a significantly better fit for the one-factor model (exact $p<.00000001$). Thus, although two-factor models have good psychometric properties across all subgroups according to these analyses, the one-factor solution is preferable from a statistical perspective and also from the point of parsimony and tradition.

Discussion

Main findings

This study is the first to assess factor structure via confirmatory factor analysis using a sufficiently large patient sample. Its findings support the tradition of using the NDI as a one-factor measurement scale. The determination of a one-factor structure enables practitioners and researchers to use the NDI as a single construct that produces a valid single summated score. The single-factor solution supports the previous findings of several authors [6,21], particularly as an appropriately sized sample is used. This is important for CFA analysis to test competing hypotheses about factor structure. However, the findings are not absolute as a tendency towards a two-factor structure for the male subgroup will require further investigation as will consideration of subgroup diagnoses, particularly WAD.

Table 3
Summary of model comparisons between one-factor and two-factor solutions for the Neck Disability Index

Sample	N	RMSEA*	CFI†	NFI‡	χ^2	df	Significant difference between models?§
Total sample							
One-factor	1,279	.024	.996	.991	36.74	21	p<.000000000006
Two-factor	1,279	.047	.983	.977	91.75	24	
Males only							
One-factor	416	.007	1.000	.979	27.53	27	p<.07
Two-factor	416	.020	.996	.973	36.09	31	
Females only							
One-factor	862	.004	1.000	.993	18.26	18	p<.00000001
Two-factor	862	.045	.987	.980	54.39	20	

CFI, comparative fit index; NFI, Normed Fit index; RMSEA, root mean square error of approximation.

For both models, modification indices were examined and the model was optimized to fit that sample to the extent allowed. In the one-factor model all errors were allowed to covary if the modification indices was greater than 4.00. In the two-factor model, errors were allowed to covary with errors from other items within the same factor, but not across factors.

* Closer to 0.000 is better.

† Closer to 1.00 is better.

‡ Closer to 1.00 is better.

§ Calculated as a χ^2 ($\chi^2_2-\chi^2_1$) with df (df2−df1).

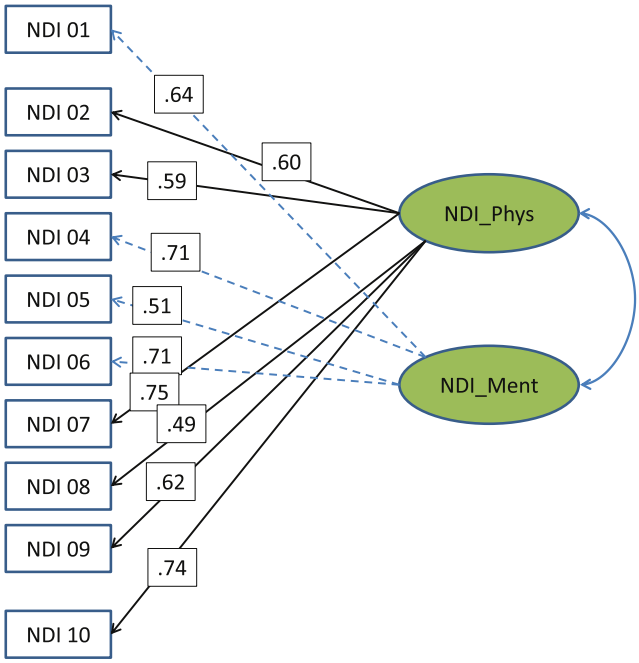

Figure. CFA two-factor solution.

Is a two-factor solution appropriate?

Our results show the two-factor model is inferior but also fits the data well. This structure may have utility and validity in certain research or other applications. The four mental and six physical function items all loaded consistently under the separate constructs in the three different analyses. These separations were consistent with the findings of the concept-retention methodology. Consequently, the two-factor model had good psychometric properties across all subgroups, but the one-factor solution was preferable from the perspective of both statistics and parsimony.

The finding of a potential two-factor solution supports other researchers that used MLE in their EFA [6,20] or Rasch analysis [22]. The breakdown in content toward mental and physical constructs supports the dual nature described by Sterling et al. [30], Young et al. [20] and Nieto [1]. It is also consistent with the summation of items and construct separation in the systematic review of MacDermid et al. [6,13] that suggests seven items are consistently allocated: pain intensity, headaches, and reading to one construct (mental function in this study); and personal care, lifting, driving, and recreation to another (physical function in this study). The three unaccounted items would then fit consistently with both the two-factor solution and the concept-retention analysis with concentration to mental function and work and sleep to the physical function construct. This consistency with the concept-retention findings provides support to this qualitative methodological process [25]. This may also lend support to the recommendations and findings of previous authors. Young et al. [20] suggested a two-factor structure be considered, particularly for "psychologically affected adults," and that an "outcome measure with items related only to activity limitations and participation restrictions might give a truer picture of disability associated with neck pain for patients with psychological distress." Furthermore, Sterling et al. [30] recommended that "co-relationships between biological and psychological factors [of neck pain] be disentangled if improved outcomes are to be gained and clinicians are to … adequately address all aspects of the patient's condition."

Our findings are in contrast to those of van der Velde et al. [22], who noted the items of lifting and headaches were problematic and suggested the need to reduce the NDI to an eight-item tool. This may not be required but remains something that should be examined further. Possible differences between these two studies may be summarized into two categories, demographic and statistical. From the demographic perspective, van der Velde et al. [22] used a pooled population of 521 participants analyzed from two previously published studies on chiropractic populations of nonspecific neck pain; our study used 1,278 medical and physiotherapy participants that included nonspecific and WAD participants. The groups differed substantially in geographic location and, consequentially, potential social differences because the earlier study was from Southern California in the late 1990s with compensation participants excluded whereas our study used an international population from different areas of Australia, the Eastern United States, and Spain collected after 2006 and included compensation patients. However, the average age and sex distribution were similar. From a statistical perspective, the sample size could be an influence as van der Velde et al., with two separate pooled populations, may have observed their discrepant results because of heterogeneous or incompatible populations. This confounding factor would similarly apply to our study with six separate data pools from different social contexts.

Furthermore, Rasch analysis and CFA ask slightly different questions of the data. Our paper essentially asked whether the structure of the questionnaire was unitary or multifaceted using ML estimation and certain model specifications and assumptions. Rasch analysis essentially asks whether all items are equally informative and seeks to create a single "true" score by evaluating differential informativeness of each item consequently using different estimations, assumptions and modeling.

Our results also indicate that new reference values can be identified pragmatically by the use of a continuous index based on the components of the one-factor model obtained through CFA. This NDI_REVA can offer a standard frame of reference and the possibility of assessing the evolution of neck disability patients in a more objective way than simply interviewing them or arbitrarily allocating a descriptive indicator of status and functional loss based only on subjective criteria. The NDI_REVA can cover this as its simplicity of use improves practicality within the primary care setting. This will impact the everyday use of the NDI in the clinical setting by providing simple user friendly

and easily accessible interpretations of scores and what they mean to the individual patient. Furthermore, to assist clinicians and researchers to obtain immediate results based on crude patient data, a software application that provides the value of the index is under development (see http://www.salud.uma.es/calculaNDI_REVA/).

Limitations and strengths

The limitations of this study include the inability to further assess different subgroups within the total data pool, including their socio-demographic and clinical characteristics because to the restricted number of variables that were provided and that the study was a secondary analysis of existing pooled data. The specific implications of a separate psychological function construct that may become dominant and influence the overall validity of the NDI in providing a single summated score could not be specifically accounted for in the analysis because of the lack of a psychological specific criterion [20]. Furthermore, all contributing data were from physical therapy outpatient clinical studies. This sample may consequently not be truly representative of the total general population with neck pain or problems. In addition, the results cannot be extrapolated to the inpatient or community care settings or specific symptomatic neck subgroups such as WAD and possibly headache.

The strengths of this study include the data pooled from multiple centers and that the sample is adequately large and meets the minimum requirements for CFA. Furthermore, all data were analyzed separately with the use of different software systems that provided the same result, and there was consistency between the subjective based content-retention methodology of item construct identification and the statistical findings of a two-factor structure-model.

Implications for future research

There is a need for larger studies in specific clinical and research settings that investigate the NDI alone, in condition- and sex-specific subgroups, and concurrently with other PRO measures that consider the symptoms associated with neck pain and problems as a one-factor model. Such research should consider neck specific PROs and regional whole-spine PROs that can be assessed as a single summated score. Such studies may also concurrently use both Rasch analysis and CFA to examine the presence and nature of underlying psychological factors that may be assessed specifically by questionnaires designed and targeted for this purpose [20,30]. They may also use tools with a capacity for multidimensional construct determination—but are scored based upon these underlying constructs [21] and not as a single summated score [17]. There may also be a need to consider social desirability [31,32] and whether this influences the expression of the two constructs identified within this study both subjectively and statistically through an influence on the psychological construct. Although it has been shown that one-factor regional PROs are not affected by social desirability [26], there are no studies that have reported this influence on conditions or regions in patient groups susceptible to expressions of a two-factor nature when influenced by an underlying psychological trait.

Conclusion

This study is the first to use an adequately large sample to investigate the factor structure of the NDI through CFA. The findings are conclusive in that although two-factor models have good psychometric properties across the full sample and sex subgroups, the one-factor solution is preferable from a statistical perspective, and also from the point of parsimony. The item separation for the two-factor solution was consistent in all analyzed subgroups to two constructs of mental and physical function determined from concept-retention methodology. Potential NDI users must be aware of this for use in both condition or gender specific subgroups in either the research and/or clinical settings. Measures of a single summated score may be influenced by underlying changes in patients' mental functioning state. Consequently, more research is needed to investigate the NDI factor structure with CFA in different conditions, subgroups and gender specific populations.

Acknowledgments

We thank all participating researchers and patients who contributed data to the study. This included the University of the Sunshine Coast, Queensland, Australia; Faculty of Medicine Malaga University–Spain; Michelle Sterling from University of Queensland, Queensland, Australia; Sharon Young of Mobile Spine and Rehabilitation Center, Mobile, Alabama; The Melbourne Whiplash Center staff and patients, Victoria, Australia; and Dr. Andrade-Ortega Complejo Hospitalraio Ciudad de Jaen, Spain. The study was a secondary analysis of existing deidentified data that had Human Research Ethics Committee Approval from each of the contributing researchers' institution.

References

[1] Nieto R, Miró J, Huguet A. Disability in subacute whiplash patients: usefulness of the neck disability index. Spine 2008;33:E630–5.
[2] Sterling M. Neck pain: much more than a psychosocial condition. J Orthop Sports Phys Ther 2009;39:309–11.
[3] Côté P, Cassidy JD, Carroll LJ, Kristman V. The annual incidence and course of neck pain in the general population: a population-based cohort study. Pain 2004;112:267–73.
[4] Cleland JA, Childs JD, Whitman JM. Psychometric properties of the Neck Disability Index and Numeric Pain Rating Scale in patients with mechanical neck pain. Arch Phys Med Rehabil 2008;89:69–74.
[5] Vernon H, Mior S. The neck disability index: a study of reliability and validity. J Manipulative Physiol Ther 1991;14:409–15.

[6] MacDermid JC, Walton DM, Avery S, et al. Measurement Properties of the Neck Disability Index: A Systematic Review. J Orthop Sports Phys Ther 2009;39:334–50.
[7] Sterling M, Jull G, Vicenzino B, Kenardy J. Characterization of acute whiplash-associated disorders. Spine 2004;29:182–8.
[8] Vernon H. The neck disability index: patient assessment and outcome monitoring in whiplash. J Musculoskelet Pain 1996;4:95–104.
[9] Vernon H, Guerriero R, Kavanaugh S, et al. Psychological Factors in the Use of the Neck Disability Index in Chronic Whiplash Patients. Spine 2010;35:E16–21.
[10] WorkSafe Victoria and the Traffic Accident Commission of Victoria (TAC). Clinical framework for the delivery of health services. 2009. Available at: http://www.worksafe.vic.gov.au/wps/wcm/connect/e93587004071f3d096b8dee1fb554c40/tac_worksafe_clinical_framework.pdf?MOD=AJPERES. Accessed October 2, 2013.
[11] South Australian Centre for Trauma and Injury Recovery. Clinical guidelines for the management of acute and chronic whiplash-associated disorders. 2008. Available at: http://www.nhmrc.gov.au/publications/synopses/_files/cp112.pdf. Accessed October 2, 2013.
[12] Australian Physiotherapy Association (APA). Policy on outcome measures and treatment justification. 2003. Available at: http://www.physiotherapy.asn.au/quality-practice/outcome-measures/http://www.tac.vic.gov.au/upload/TAC+Physio+Chart.pdf. Accessed October 2, 2013.
[13] MacDermid JC. Response to letter by Gabel et al. on factor analysis findings for the NDI. J Orthop Sports Phys Ther 2009;39:828–31.
[14] Vernon H. The Neck Disability Index: state-of-the-art, 1991-2008. J Manipulative Physiol Ther 2008;31:491–502.
[15] Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. Psychol Methods 1999;4:272–99.
[16] Costello AB, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Pract Assess Res Eval 2005;10:1–9.
[17] Doward LC, McKenna SP. Defining patient-reported outcomes. Value Health 2004;7:S4–8.
[18] Gabel CP, Burkett B, Yelland M. Balancing fidelity and practicality in short version musculoskeletal outcome measures. Phys Ther Rev 2009;14:221–5.
[19] Osborne JW, Costello AB, Kellow JT. Best practices in exploratory factor analysis. In: Osborne J, ed. Best practices in quantitative methods. Thousand Oaks, CA: Sage Publishing, 2008: 205–18.
[20] Young SB, Aprill C, Braswell J, et al. Psychological Factors and Domains of Neck Pain Disability. Pain Med 2009;10:310–8.
[21] Pickering PM, Osmotherly PG, Attia JR, McElduff P. An Examination of Outcome Measures for Pain and Dysfunction in the Cervical Spine: A Factor Analysis. Spine 2011;36:581–8.
[22] van der Velde G, Beaton D, Hogg-Johnston S, et al. Rasch analysis provides new insights into the measurement properties of the neck disability index. Arthritis Rheum 2009;61:544–51.
[23] Hains F, Waalen J, Mior S. Psychometric properties of the neck disability index. J Manipulative Physiol Ther 1998;21:75–80.
[24] Sterling M, Kenardy J, Jull G, Vicenzino B. The development of psychological changes following whiplash injury. Pain 2003;106:481–9.
[25] Beaton DE, Wright JG, Katz JN, Group UE. Development of the QuickDASH: comparison of three item-reduction approaches. J Bone Joint Surg Am 2005;87:1038–46.
[26] Gabel CP, Michener L, Burkett B, Neller A. The Upper Limb Functional Index (ULFI): Development and Determination of Reliability, Validity and Responsiveness. J Hand Ther 2006;19:328–49.
[27] Schumacher RE, Lomax RGA. Beginner's guide to structural equation modeling. Mahwah, NJ: Lawrence Erlbaum, 1996.
[28] McDonald RP, Marsh HW. Choosing a multivariate model: noncentrality and goodness of fit. Psychol Bull 1990;107:247–55.
[29] Sterling M, Jull G, Kenardy J. Physical and psychological factors maintain long-term predictive capacity post-whiplash injury. Pain 2006;122:102–8.
[30] Sterling M, Jull G, Vicenzino B, et al. Physical and psychological factors predict outcome following whiplash injury. Pain 2005;114:141–8.
[31] Edwards A. The social desirability variable in personality assessment and research. New York, NY: Dryden, 1957.
[32] Crowne DP, Marlowe D. A new scale of social desirability independent of psychopathology. J Consult Psychol 1960;349–54.