SPINE Volume 32, Number 9, pp 1028–1037

Self-Report Outcome Measures for Low Back Pain

Searching for International Cross-Cultural Adaptations

Leonardo Oliveira Pena Costa, PT, MSc, Chris G. Maher, PT, PhD, and Jane Latimer, PT, PhD

Study Design. Systematic review.

Objectives. To describe the available cross-cultural adaptations of low back pain (LBP) self-report outcome measures and the psychometric testing that has occurred for each adaptation.

Summary of Background Data. Self-report measures are commonly used in clinical practice and in research studies. Most existing questionnaires were developed in English, and it is not clear how many have been adapted to other languages.

Methods. Two different searches on MEDLINE, EMBASE, CINAHL, and LILACS were performed. The first search identified questionnaires specifically designed for patients with LBP. The second search combined the name of the questionnaire with 35 different languages in order to locate cross-cultural adaptations of the questionnaire. Data on the psychometric testing of the translated questionnaires were extracted.

Results. Forty questionnaires were identified, only 15 of which had been adapted to a new language. Only 19 of the 35 different languages we searched for were represented in the search results. From 1400 possible adaptations, only 61 have been completed. Psychometric testing of the adapted questionnaires was quite variable and in general suboptimal with testing usually restricted to an assessment of reliability and construct validity.

Conclusions. There is a clear need for further cross-cultural adaptation of LBP self-report measures and for greater attention to the quality of psychometric evaluation of adapted questionnaires. Without appropriately adapted measures, the clinical management of LBP patients who do not speak English is potentially compromised.

Key words: outcome assessment, psychometric properties, low back pain. **Spine 2007;32:1028–1037**

Self-report outcome measures are commonly used in clinical practice, in clinical research and large epidemiologic studies. Many are simple, reliable, and of low cost, making them suitable for quality assurance and research activities. Most existing self-report measures were developed in English, and it is not clear how many have been adapted to other languages.

Cross-cultural adaptation of existing self-report measures is important for a number of reasons. First, not all the world speaks English; and even in English-speaking countries like Australia, United Kingdom, and United States, a significant proportion of the population are not English-speakers. Second, the availability of adapted questionnaires may stop the undesirable, but common, practice of excluding non-English subjects from clinical trials conducted in English-speaking countries. Third, the existence of adapted questionnaires would be of value to researchers conducting systematic reviews by assisting the pooling of data from trials conducted in non–English-speaking countries. Lastly, adaptation of existing questionnaires is potentially more simple and efficient than requiring researchers in non–English-speaking countries to develop their own self-report measures.

Cross-cultural adaptation is more than just simple translation of English text. Researchers need to follow guidelines for cross-cultural adaptation to ensure that the adapted questionnaire is semantically equivalent to the original and that the items and scale are relevant in the new culture. Following this first step, it is also essential to assess whether the adapted questionnaire has retained the content validity of the original questionnaire. This second step requires an assessment of the measurement properties of the new questionnaire.[1]

The measurement properties (or psychometric properties) of a questionnaire include both reliability and validity of the calculated scores. A brief explanation of the concepts in psychometric testing is provided in Table 1.

The objectives of this paper are to describe the cross-cultural adaptations of self-report measures relevant to the management of low back pain (LBP) that are now available and describe the psychometric testing that has occurred for each adaptation.

Materials and Methods

Literature Search 1. The purpose of this first step was to identify self-report outcome measures designed for patients with LBP. A systematic literature search was carried out in MEDLINE, EMBASE, LILACS, and CINAHL databases for the period from January 2001 to July 2006. The following search terms were included: *back pain*, *low back pain*, *scoliosis, spinal stenosis, disc herniation, nerve root compromise*, and *ankylosing spondylitis* these terms were combined with the terms *questionnaire(s), outcome measure(s), index*, and *scale(s)*. The search strategy was: (back pain OR low back pain OR scoliosis OR spinal stenosis, OR disc herniation OR ankylosing spondylitis) AND (questionnaire(s) OR outcome mea-

From the Back Pain Research Group, School of Physiotherapy, University of Sydney, Lidcombe, NSW Australia.
Acknowledgment date: April 13, 2006. First revision date: July 17, 2006. Second revision date: August 11, 2006. Acceptance date: August 15, 2006.
Leonardo Costa is a PhD student supported by CAPES, Ministério da Educação, Brazil; Christopher Maher holds a research fellowship funded by the National Health and Medical Research Council of Australia.
The manuscript submitted does not contain information about medical device(s)/drug(s).
No funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.
Address correspondence and reprint requests to Leonardo Oliveira Pena Costa, PT, MSc, Back Pain Research Group, School of Physiotherapy, University of Sydney, East Street, Lidcombe, NSW Australia 2141; E-mail: lcos3060@usyd.edu.au

Table 1. Concepts in Psychometric Testing[18]

Concept	Explanation
Internal consistency	Internal consistency is the extent to which the items assess the same construct. If a questionnaire has a no. of items addressing the same underlying dimension (*e.g.,* disability status in low back pain patients), then it is reasonable to expect that scores on each item would be correlated with scores on all other items. Internal consistency can be determined by calculating Cronbach's alpha. A low alpha value means that some items may be ambiguous, while a very high Cronbach's alpha means that the items in the questionnaire show too much homogeneity and some items may be redundant
Factor analysis	Factor analysis confirms the structure of the questionnaire by summarizing patterns of correlations among observed variables, to reduce a large no. of items into a smaller no. of factors. Items that do not load on any factor or that load on multiple factors should be omitted or rephrased
Reliability	Reliability is the extent to which an experiment, test, or any measuring procedure yields the same result on repeated trials. Without the agreement of independent observers able to replicate research procedures or the ability to use research tools and procedures that yield consistent measurements, researchers would be unable to satisfactorily draw conclusions, formulate theories, or make claims about the generalizability of their research. A common measure of reliability is the intraclass correlation coefficient (ICC). This measure takes random error as well as systematic error (bias) into account. Other reliability indexes such as Pearson's correlation and coefficient of variation can be used
Responsiveness	Responsiveness is the ability of an instrument to measure real or important change over time, in the concept being measured
Validity	Validity refers to the degree to which a study accurately reflects or assesses the specific concept that the researcher is attempting to measure. *Construct* validity refers to the extent to which scores on a particular instrument relate to other measures in a manner that is consistent with theoretically derived hypotheses concerning the concepts that are being measured. Where an external gold standard is available, to which the self-report measure could be compared, evaluation of *criterion* validity is possible. However, many self-report measures assess attributes such as quality of life or disability where there is no accepted external gold standard. In these cases, only assessment of *construct* validity is possible
Rasch analysis	Rasch analysis addresses unidimensionality by assessing the contribution that items make to the scale hierarchy. The technique also provides an estimate of item difficulty based on the frequency with which patients respond to an item, which can be used to assess the position of items along the scale and to consider any possible redundancy or gaps in the scale hierarchy
Floor or ceiling effects	Floor or ceiling effects are considered to be present if more than 15% of respondents achieved the lowest or highest possible score, respectively. If floor or ceiling effects are present, it is likely that items assessing the extremes of the attribute are missing from the scale. The consequence of a floor effect is that deterioration may be missed and for a ceiling effect improvement may be missed

sure(s) OR index OR scale(s)). In addition, hand searches of journals and textbooks of spinal disorders were carried out. Titles and abstracts of papers from the search were scanned to determine eligibility. The inclusion criteria were: 1) that the questionnaire had the word "*back*" in the text or in the title and 2) the questionnaire was specifically developed for LBP patients. Papers published in non-English journals were only included if they had an English abstract.

Literature Search 2. The purpose of this step was to locate different language versions available for the questionnaires identified in Step 1. The list from the first search was used for the second literature search, the terms used were the *questionnaire's name* (*e.g.*, Roland Morris) and *validation, translation, cross-cultural adaptation* plus 35 different languages (*e.g.*, Portuguese) using AND/OR operands to combine the terms (*e.g.*, Roland Morris AND (validation OR translation OR cross-cultural adaptation) AND Portuguese). This second literature search was performed on the same databases used in the first search. There was no time limit on this search strategy.

Following the searches the list of questionnaires and cross-cultural adaptations was presented to 27 experts in spinal disorders from the United States, France, Canada, Netherlands, Germany, Norway, China, Tunisia, Turkey, Thailand, Italy, Denmark, Hong Kong, Spain, Switzerland, Korea, Japan, Greece, Brazil, Sweden, and Finland to determine whether the list was comprehensive.

Description of the Psychometric Testing of Translated Questionnaires. The following data were extracted from articles describing psychometric testing of the translated questionnaires: citation, sample size, and all types of psychometric properties possible (*i.e.*, test-retest reliability, internal consistency, Rasch analysis, responsiveness, factor analysis, and construct validity). Additionally, psychometric properties were rated by the Quality Criteria for Psychometric Properties of Health Status Questionnaires[2] with the evaluation restricted to the subset of items relevant to cross-cultural adaptation (Appendix available online only through Article Plus). This checklist considers both the quality of psychometric testing and the results of psychometric testing and so is somewhat different from scales used to measure the methodologic quality of trials. We are unaware of a scale that only measures quality of psychometric testing.

■ Results

From the first search, 66 potentially eligible questionnaires were identified from MEDLINE, 12 from CINAHL, 23 from EMBASE, and 1 from LILACS; however, neither CINAHL, nor EMBASE nor LILACS added new papers to the 66 obtained from MEDLINE. Twenty eight questionnaires did not meet inclusion criteria leaving 38 eligible questionnaires. The most common reasons for exclusion were the absence of the word "back" in the title or in the text, followed by no evidence that the questionnaire was specifically designed for LBP patients. The expert committee's review of the list of questionnaires found the list to be quite comprehensive. Committee members added 2 additional eligible questionnaires[3,4] that were not found by the original searches.

From the second search in MEDLINE, 57 cross-cultural adaptation studies were identified. Only 15 of 40 eligible questionnaires had been adapted to a new language, and only 19 of 35 different languages we searched for were represented in the search results. As with the first search, the other databases did not add new international validations to the MEDLINE results. The experts found 4 new cross-cultural adaptations, providing a total of 61 cross-cultural adaptations. Of the 4 new adaptations, 1 was published in a nonindexed journal,[5] 1 was an article "in press,"[6] and 2 were from a paper that did not mention any of the terms that we used in our search.[7,8]

Table 2 describes the 40 original LBP questionnaires and 61 cross-cultural adaptations. From 1400 possible translations (40 questionnaires by 35 languages), only 61 have been completed. The most frequently adapted

Table 2. LBP Questionnaires and Their International Adaptations

	Original Language	Arabic	Cantonese	Mandarin	Danish	Dutch	French	German
Acute Low Back Pain Screening Questionnaire[19]	English							
Back Dysfunction Score[29]	English							
Back Beliefs Questionnaire[22]	English							
Back Illness Pain and Disability 9-Item Scale[23]	English							
Back Pain Functional Scale[24]	English							
Back Pain Interference Scales[25]	English							
Back Performance Scale[3]	English							
Back-specific version of the SF-36 Physical Functioning Scale[26]	English							
Bath Ankylosing Spondylitis Disease Activity Index[27]	English							
Bath Ankylosing Spondylitis Functional Index[31]	English							*14
Bournemouth Questionnaire[35]	English				*36			
CORE outcome measure[37]	English							
Core set[4]	German							
Curtin Back Screening Questionnaire[38]	English							
Dallas Pain Questionnaire[39]	English						*40	
Dougados Functional Index for Ankylosing Spondylitis[41]	English							*14
Fear Avoidance Beliefs Questionnaire[44]	English						*45	*46
Functional Outcomes Questionnaire: Spinal Disorders[48]	English							
Functional Rating Index[49]	English							
General Function Score[50]	English							
Hannover Functional Ability Questionnaire[51]	German							
Low Back Outcome Score[52]	English							
Low Back Outcome Score for Back Pain[53]	English							
Low Back Pain Rating Scale[54]	English							*55
MODEMS Pain and Disability Lumbar Scale[56]	English							
NASS Lumbar Spine Outcome Assessment Instrument[57]	English							*13
Occupational Role Questionnaire[60]	English							
Outcome Measure for Lumbar Spinal Stenosis[61,62]	Swiss							
Oswestry Disability Index[63]	English	*12		*64	*6			*65
Perception of Disability Scale[71]	English							
Quebec Back Pain Disability Scale[72]	English					*73	*74	
Roland Morris Disability Questionnaire[76]	English	*77	*78	*10	*9	*79	*11	*80
Scoliosis Research Society 22[87]	English							
Spinal Function Sort[92]	English							
Spinal Pain Independence Measure[93]	Hebrew							
Aberdeen Low Back Pain Disability Scale[94]	English		*7					*95
The Maine-Seattle Back Questionnaire[96]	English							
Vermont Disability Prediction Questionnaire[97]	English							
Waddell Disability Index[98]	English							
Walter Reed Visual Assessment Scale[99]	English							

questionnaire was the Roland Morris Disability Questionnaire, which was adapted to 17 different languages, followed by the Oswestry Disability Index (11 languages), the Bath Ankylosing Spondylitis Functional Index and the Dougados Functional Index for Ankylosing Spondylitis (5 languages), the Quebec Back Pain Disability Scale (4 languages), the Bath Ankylosing Spondylitis Disease Activity Index, the NASS Lumbar Spine Outcome Assessment Instrument, the Fear Avoidance Beliefs Questionnaire, and the Scoliosis Research Society (3 languages), the Aberdeen Low Back Pain Disability Scale (2 languages), and finally the Acute Low Back pain Screening Questionnaire, the Bournemouth, the Core-set, Low Back Pain Rating Scale, and the Dallas Pain Questionnaire with only one adaptation each

The most numerous languages in the cross-cultural adaptations were Spanish and German (8 times), followed by Turkish (7 times), French and Italian (4 times), Danish, Japanese, Norwegian, Persian and Swedish (3 times), Arabic, Mandarin, Cantonese, Dutch, Greek, and Thai (twice), and the other languages were only in one adaptation.

Table 3 summarizes the reported psychometric testing of the adapted questionnaires. There are large differences in the sample sizes used in psychometric testing ranging from 30 patients (Brazilian Portuguese and Dutch Roland Morris) to 697 patients for the psychometric testing of the Greek Oswestry.

Almost all adaptations have been evaluated for test-retest reliability and internal consistency (48 of 61). Test-retest reliability is good to excellent with intraclass correlation coefficient values ranging from 0.53 (Turkish Bath Ankylosing Spondylitis Disease Activity Index) to 0.98 (Arabic Roland Morris). Other statistics were less commonly used to describe test retest reliability *e.g.*, Pearson's r (German Roland Morris: r = 0.82; P = 0.0001) and coefficient of variation (Norwegian Oswestry and Roland Morris: 12% and 15%, respectively). Internal consistency is also good to excellent with Cronbach's alpha ranging from 0.57 (Japanese Scoliosis Research Society) to 0.96 (Spanish Scoliosis Research Society). For internal consistency, Pearson's r (French Quebec: 0.44–0.76), Spearman's Rho coefficient (German Low Back Pain Rating Scale: 0.98), and Kuder-Richardson 20 coefficient (Chinese Hong Kong Roland Morris: 0.86) were also used.

Only 19 of 61 adapted questionnaires have been tested for responsiveness. Convergent validity was assessed in 57 of 61 questionnaires by comparing questionnaire scores to other measures measuring similar constructs. In many cases, the comparison measure seemed quite similar, *e.g.*, comparison of the Greek Roland Mor-

Table 2. Continued

Greek	Italian	Korean	Japanese	Norwegian	Persian	Portuguese	Romanian	Spanish	Swedish	Thai	Turkish
				*20							
								*28	*29		*30
	*15						*32	*28			*33,34
								*8			
	*15							*28	*42		*33,43
								*47			
	*58							*59			
*66		*17	*67	*68	*69					*5	*70
					*69						*75
*66	*81		*67	*68	*69	*82		*83	*84	*85	*86
			*88					*89,90			*91

ris to the Greek Oswestry where both measures are LBP specific measures of disability. However, some comparisons were more dissimilar, *e.g.*, the comparison of the Korean Oswestry to a Korean pain measure and a Korean quality of life measure. Rasch analysis was only performed on the Turkish Roland Morris version, and factor analysis was only performed on 8 adaptations.

The quality and outcome of the psychometric testing of each adaptation are described in Table 4. In general, there is acceptable evidence for reliability and construct validity but not internal consistency, responsiveness, or floor or ceiling effects. While internal consistency has been widely evaluated, the evaluation has typically not been acceptable. The problem with responsiveness and floor or ceiling effects is that these concepts have not usually been tested.

■ Discussion

Our study aimed to describe the cross-cultural adaptations of self-report measures relevant to the management of LBP. A total of 61 cross-cultural adaptations in 19 different languages were identified. Of the 40 original language questionnaires, only 15 have been adapted, which shows clearly that more effort in this field is required. There were also considerable differences in the psychometric testing that has been undertaken with each adaptation.

When we visited the Oswestry website (http://www.orthosurg.org.uk/odi/) and Roland Morris e-mail (mroland@man.ac.uk), we found more versions than we did in our database searches. We contacted all authors by e-mail and established that all additional versions had only been published as theses or conference abstracts. We did not include these papers because these questionnaires have not been through a peer-review process and their quality thus remains unclear. In our view, the unpublished description and evaluation of such questionnaires are not useful.

In our study, we identified papers in non-English journals only if they contained an English abstract with an English description of the results. Ten questionnaires (8 papers[5,9–15]) fulfilled this criterion. It is possible, however, that there were questionnaires available in other languages that were missed by our search strategy. Another problem is that some questionnaires have different names in different languages, for example: the French RMDQ is known in France as the EIFEL questionnaire; and due to this practice, some additional cross-cultural adaptations may have been missed. To avoid this problem, we advise that,

Table 3. Psychometric Characteristics of LBP Questionnaires

Language	Sample Size	Test-Retest Reliability (ICC)	Internal Consistency (Cronbach's alpha)	Responsiveness	Construct Validity (Pearson correlation)
Acute Low Back Pain Screening Questionnaire					
Norwegian[20]	123	0.90	0.95		RMDQ: 0.46, age: 0.38
Bath Ankylosing Spondylitis Disease Activity Index					
Spanish[28]	144	0.74		Effect size: 1.6	General well-being: 0.7, Pain VAS: 0.53, Morning stiffness: 0.64
Swedish[29]	113			Sensitivity to change: $P < 0.05$	BASFI: 0.64
Turkish[30]	71	0.53–0.85	0.80		BASFI: 0.62 ($P < 0.001$), Physician's assessment: 0.44 ($P < 0.001$)
Bath Ankylosing Spondylitis Functional Index					
German[14]	72	0.92	0.81	SRM: 0.46	Schober test: −0.30 ($P < 0.05$), Fingertip to floor: 0.40 ($P < 0.001$), Pain VAS: 0.38 ($P < 0.01$)
Italian[15]	95	0.91	0.90	ROC curve: 0.90	Pain VAS: 0.73, BASDAI: 0.72
Romanian[32]	41	0.82	0.93		Dougados Index: 0.37 ($P < 0.05$), Pain VAS: 0.44 ($P < 0.01$)
Spanish[28]	144	0.68		Effect size: 1.2	Schober test: −0.4, Occiput to wall: 0.38
Swedish[42]	113			Sensitivity to change: $P < 0.001$	BASDAI: 0.68
Turkish[100]	81	0.76–0.95	0.93		Schober test: 0.44 ($P < 0.001$), Fingertip to floor: 0.56 ($P < 0.001$), Occiput to wall: 0.53 ($P < 0.01$)
Bournemouth Questionnaire					
Danish[36]	118	0.96	0.88–0.89		RMDQ: 0.46–0.73, SF-36: 0.64–0.71
Dallas Pain Questionnaire					
French[40]	59	0.75	0.89–0.91	Sensitivity to change: $P < 0.001$	Pain VAS: 0.78 ($P < 0.001$)
Core set					
Spanish[8]	131		0.91	Effect size: 0.91–6.15	
Dougados Functional Index for Ankylosing Spondylitis					
German[14]	72	0.89	0.85	SRM: 0.33	Schober test: −0.19 (NS), Fingertip to floor: 0.32 ($P < 0.01$), Pain VAS: 0.22 (NS)
Italian[15]	95	0.86	0.87	ROC curve: 0.82	Pain VAS: 0.53, BASDAI: 0.65
Spanish[28]	144	0.87		Effect size: 1.05	BASFI: 0.83, Schober test: −0.36
Turkish[43]	70	0.61–0.88	0.91	Sensitivity to change: NS	Schober test: −0.29 ($P < 0.05$), Fingertip to floor: 0.45 ($P < 0.001$), Occiput to wall: 0.38 ($P < 0.01$)
Fear Avoidance Beliefs Questionnaire					
French[45]	217	0.72–0.88		Effect size: 0.98–1.37	Pain VAS: 0.36, Quebec: 0.36, HADS: 0.29
German[46]	302	0.87	0.89		Pain VAS: 0.26 ($P < 0.002$), FFbH-R: 0.36 ($P < 0.002$)
Spanish[47]	209	0.96	0.93		Pain VAS: 0.398 ($P = 0.00$), RMDQ: 0.522 ($P = 0.00$), Quality of life physical: −0.361 ($P = 0.00$), Quality of life mental D1: −0.361 ($P = 0.00$)
Low Back Pain Rating Scale					
German[55]	126		0.95		RMDQ: 0.91, SF-36: −0.34 to −0.72
NASS Lumbar Spine Outcome Assessment Instrument					
German[13]	56	0.82–0.89			SF-36: 0.83, FFbH-R 0.28, Pain VAS: 0.68
Italian[58]	74	0.89–0.92	0.87–0.90		SF-36: −0.54, Pain VAS: −0.43 to −0.58
Spanish[59]	70	0.63–0.91	0.78–0.90	ROC curve: 0.74 −0.81, Effect size: −0.05–2.02	SF-36: −0.82, Pain VAS: −0.63
Oswestry Disability Index					
Arabic[12]	80	0.98	0.70–0.76		Quebec: 0.86, Pain VAS: 0.57
Mandarin[64]	79	0.86	0.81		
Danish[6]	191	0.91	0.88		RMDQ: 0.78 ($P < 0.01$), Low back pain rating scale: 0.68 ($P < 0.01$), SF-36: 0.75 ($P < 0.01$)
German[65]	100	0.96	0.90		Pain VAS: 0.78 ($P < 0.001$), RMDQ: 0.80 ($P < 0.001$)
Greek[66]	697		0.83		RMDQ: 0.729 ($P < 0.0005$)
Japanese[67]	97	0.93	0.83		RMDQ: 0.785 ($P < 0.01$), JOA: 0.647 ($P < 0.01$)
Korean[17]	206	0.91	0.84		Pain VAS: 0.425 ($P = 0.0001$), WHOQOL-BREF: −0.09 (NS) to −0.48 ($P = 0.001$)
Norwegian[68]	105	0.88	0.94		RMDQ: 0.73–0.60, Pain VAS: 0.39–0.52
Persian[69]	31	0.91	0.75		SF-36: −0.66; Pain VAS: 0.54 ($P < 0.001$)
Thai[5]	63	0.62	0.93		
Turkish[70]	95	0.94	0.89–0.91		Pain VAS: 0.367 ($P < 0.01$), Schober test −0.068 (NS), RMDQ: 0.815 ($P < 0.001$)
Quebec Back Pain Disability Scale					
Dutch[73]	120	0.90	0.95		RMDQ:0.80, Pain VAS:0.70, Course of the complaint:0.35
French[74]	32				Dallas: 0.755, Perceived health status: 0.739, Impairment score: 0.449, Pain VAS: 0.448, HADS: 0.473
Persian[69]	31	0.86	0.92		SF-36: −0.62, Pain VAS: 0.46 ($P < 0.001$)
Turkish[75]	83		0.93		Pain VAS: 0.63 ($P < 0.001$)

(*Continued*)

Table 3. Continued

Language	Sample Size	Test-Retest Reliability (ICC)	Internal Consistency (Cronbach's alpha)	Responsiveness	Construct Validity (Pearson correlation)
Roland Morris Disability Questionnaire					
Arabic[77]	62		0.94	r = 0.83 ($P = 0.000$)	Pain VAS: 0.33 ($P = 0.0001$), Schober test: 0.27 ($P = 0.0001$), General Function Score: 0.56 ($P = 0.0001$)
Cantonese[78]	112	0.94		ROC: 0.71–0.89	Global rating of change in overall condition at discharge: −0.22 ($P = 0.02$), Mean change score of pain: −0.44 ($P < 0.001$)
Mandarin[10]	112	0.81	0.84		SF-36: −0.73, Pain index: 0.32 ($P < 0.05$)
Danish[9]	135		0.94		SF-36: −0.88 ($P < 0.001$), LBP rating scale: 0.89 ($P < 0.001$)
Dutch[79]	30	0.91			
French[11]	80	0.89			Pain VAS: 0.27 ($P = 0.018$)
German[80]	125		0.81		Pain VAS: 0.81 ($P = 0.0001$), Forward bending: 0.48 ($P = 0.0001$), Lateral bending: −0.47 ($P = 0.0001$)
Greek[66]	697		0.88		ODI: 0.729 ($P < 0.0005$)
Italian[81]	70	0.92	0.82		Pain VAS: 0.79 ($P < 0.001$)
Japanese[67]	97	0.95	0.86		ODI: 0.785 ($P < 0.01$), JOA: −0.568 ($P < 0.01$)
Norwegian[68]	105	0.89	0.94		ODI: 0.73, Pain VAS: 0.32
Persian[69]	31	0.86	0.83		SF-36: −0.62, Pain VAS: 0.36 ($P < 0.001$)
Portuguese[82]	30	0.94–0.95			Pain VAS: 0.79 ($P < 0.01$), Pain numerical scale (0–6): 0.80 ($P < 0.01$)
Spanish[83]	195	0.87	0.83–0.91		Pain VAS: 0.347 ($P = 0.0000$), ODI: 0.197 ($P = 0.0061$)
Swedish[84]	72	0.88			Perceived disability: 0.64 ($P < 0.001$), Pain severity: 0.54 ($P < 0.001$), Perceived life control: −0.32 ($P < 0.01$), General activity: −0.27 ($P < 0.05$)
Thai[85]	120		0.83		
Turkish[86]	81	0.79–0.86	0.85–0.89	Effect size: −0.53	Spinal mobility: −0.33 ($P < 0.01$) −0.11 (NS)
Scoliosis Research Society 22					
Japanese[88,101]	141		0.57		Radiologic measures: −0.33 ($P < 0.01$)
Spanish[89,90]	175	0.89	0.96		Radiologic Measures: −0.25 ($P < 0.001$)
Turkish[91]	82	0.76–0.82	0.76–0.83		SF-36: 0.27–0.75 ($P < 0.0001$)
Aberdeen Low Back Pain Disability Scale					
Cantonese[7]	473	0.94	0.85	Effect size: 0.59–0.81	RMDQ at baseline: 0.66, RMDQ at discharge 0.68
German[95]	158	0.95	0.86	SRM: 1.36–1.96	SF-36: −0.48 to −0.76 ($P < 0.0001$), FFbH-R: −0.76 ($P < 0.0001$)

SF-36 indicates Short Form Health Survey; BASFI, Bath Ankylosing Spondylitis Functional Index; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; Pain VAS, Pain Visual Analogue Scale; RMDQ, Roland Morris Disability Questionnaire; SEM, standard error of measurement; SRM, standardized response mean; WHOQOL-BREF, World Health Organization Quality of Life Assessment; ODI, Oswestry Disability Index; HADS, Hospital Anxiety and Depression Scale; FFbH-R, Hannover Functional Ability Questionnaire; NS, not significant.

where possible, the original name of the questionnaire should be retained and supplemented with the name of the language/country to which the questionnaire has been adapted.

This is the first study to systematically review self-report measures for LBP patients that are available in different languages. While there is a large number of original questionnaires available, very few have been translated into other languages, particularly languages commonly spoken like Mandarin (885 million speakers by 2 questionnaires), Spanish (332 million speakers by 8 questionnaires), Bengali (322 million speakers by no questionnaire), Hindi (182 million speakers by no questionnaire), Portuguese (181 million speakers by 1 questionnaire), Russian (145 million speakers by no questionnaire), Japanese (127 million speakers by 3 questionnaires), and German (120 million speakers by 8 questionnaires).[16] This pattern of results suggests that there is no association between the number of people who speak a language and the number of adaptations to that language.

The translation procedures used were quite similar to those recommended to those in the *Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures*[1]; however, the psychometric testing of the adapted questionnaire varied substantially between studies. There are major differences in sample sizes, test-retest periods, statistical analysis, and benchmarks considering reliability and validity.

The Terwee quality criteria[2] do not summate into one overall quality score, as is sometimes done with trial methodologic quality scales. Summing of items presumes that the items assess the same attribute and that each item is equally important, a presumption that may not be true. Additionally, summed scores do not indicate the specific methodologic problems that are most prevalent. Accordingly, we think that it is more informative to separately consider each of the items of the Terwee scale. In this study, most of the adapted questionnaires were properly tested for reliability and construct validity, but not for internal consistency. Responsiveness and the assessment of floor and ceiling effects were usually not tested at all. We recommend, for further cross-cultural adaptations, that special attention be paid to assessment of internal consistency, responsiveness tests, and floor and ceiling effects.

One common problem that can be observed in our results is that some authors validated the self-report measures by correlating them with scales that do not measure the same construct. We hypothesize that the lack of validated questionnaires in some languages reinforces this procedure. For example, the Korean Oswestry[17] was val-

Table 4. Quality Criteria for Psychometric Characteristics of LBP Questionnaires

Language	Internal Consistency	Construct Validity	Reliability	Responsiveness	Ceiling and Floor Effects
Acute Low Back Pain Screening Questionnaire					
Norwegian[20]	+	+	+	0	0
Bath Ankylosing Spondylitis Disease Activity Index					
Spanish[28]	0	+	+	?	0
Swedish[29]	0	+	0	?	0
Turkish[30]	?	+	−	?	0
Bath Ankylosing Spondylitis Functional Index					
German[14]	?	+	+	?	0
Italian[15]	?	+	+	?	0
Romanian[32]	?	−	+	0	0
Spanish[28]	0	+	−	?	0
Swedish[42]	0	+	?	?	0
Turkish[100]	?	+	+	0	0
Bournemouth Questionnaire					
Danish[36]	?	+	+	0	0
Dallas Pain Questionnaire					
French[40]	?	+	+	0	0
Core set					
Spanish[8]	?	0	0	?	−
Dougados Functional Index for Ankylosing Spondylitis					
German[14]	?	−	+	?	0
Italian[15]	?	+	+	?	0
Spanish[28]	0	+	+	?	0
Turkish[43]	?	+	−	?	0
Fear Avoidance Beliefs Questionnaire					
French[45]	0	+	+	?	0
German[46]	+	+	+	0	0
Spanish[47]	?	+	+	0	−
Low Back Pain Rating Scale					
German[55]	+	+	?	0	0
NASS Lumbar Spine Outcome Assessment Instrument					
German[13]	0	+	+	0	0
Italian[58]	?	+	+	0	0
Spanish[59]	+	+	+	+	−
Oswestry Disability Index					
Arabic[12]	+	+	+	0	0
Mandarin[64]	+	0	+	0	0
Danish[6]	?	+	+	0	+
German[65]	?	+	+	0	0
Greek[66]	?	+	0	0	0
Japanese[67]	?	+	+	0	0
Korean[17]	?	+	+	0	0
Norwegian[68]	?	+	+	0	0
Persian[69]	?	+	+	0	0
Thai[5]	?	0	−	0	0
Turkish[70]	?	−	+	0	0
Quebec Back Pain Disability Scale					
Dutch[73]	?	+	+	0	0
French[74]	?	+	0	0	0
Persian[69]	?	+	+	0	0
Turkish[75]	?	+	0	0	0
Roland Morris Disability Questionnaire					
Arabic[77]	?	+	0	?	0
Cantonese[78]	?	+	+	+	0
Mandarin[10]	?	+	+	0	0
Danish[9]	?	+	0	0	0
Dutch[79]	0	0	+	0	0
French[11]	0	+	+	0	0
German[80]	?	+	?	0	0
Greek[66]	?	+	0	0	0
Italian[81]	?	+	+	0	0
Japanese[67]	?	+	+	0	0
Norwegian[68]	?	+	+	0	0
Persian[69]	?	+	+	0	0
Portuguese[82]	0	+	+	0	0
Spanish[83]	?	+	+	0	0
Swedish[84]	0	+	+	0	0
Thai[85]	?	0	0	0	0
Turkish[86]	0	−	+	?	0

(*Continued*)

Table 4. Continued

Language	Internal Consistency	Construct Validity	Reliability	Responsiveness	Ceiling and Floor Effects
Scoliosis Research Society 22					
Japanese[88,101]	?	+	0	0	0
Spanish[89,90]	+	+	+	0	−
Turkish[91]	?	+	+	0	0
Aberdeen Low Back Pain Disability Scale					
Cantonese[7]	?	+	+	?	0
German[95]	?	+	+	?	0

+ indicates positive rating; ?, indeterminate rating; −, negative rating; 0, no information available.

idated by comparison to pain and quality of life measures because at that time there was not another disability measure for LBP available.

This article clearly shows that there is an enormous lack of cross-cultural adaptations of LBP questionnaires. The adapted questionnaires varied considerably in the extent of psychometric testing each had undergone. Our study shows that there is a clear need for further work in this area.

■ Key Points

- There have been no systematic reviews of cross-cultural adaptations of LBP self-report outcome measures.
- Our review located 40 LBP self-report measures, of which only 15 had been adapted to a new language.
- There is an urgent need for further cross-cultural adaptation.

 *document*

Appendix available online only through Article Plus.

Acknowledgments

The authors thank Raymond Ostelo, Kitir Jirarattanaphochochai, Roberto Padua, Steven Linton, Gro Jamtvedt, Kare Birger, Petros Boscaínos, Hanne Albert, Henrik Lauridsen, Annie Mannion, Dennis Turk, Paul Bettie, Bart Koes, Francisco Kovacs, Margreth Grotle, Rieckie de Vet, and Caroline Terwee for assisting us in the Results section.

References

1. Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000;25:3186–91.
2. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria are proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.
3. Magnussen L, Strand LI, Lygren H. Reliability and validity of the Back Performance Scale: observing activity limitation in patients with back pain. *Spine* 2004;29:903–7.
4. Mannion AF, Elfering A, Staerkle R, et al. Outcome assessment in low back pain: how low can we go? *Eur Spine J* 2005;14:1014–26.
5. Surachai SJ, Hunzavong T, Jirarattanaphochai J. Reliability of Thai version of Oswestry Questionnaire for the evaluation of low back pain patients. *Srinagarind Med J* 2002;17:248–51.
6. Lauridsen HH, Hartvigsen J, Manniche C, et al. Danish version of the Oswestry Disability Index for patients with low back pain: 1. Cross-cultural adaptation, reliability, and validity in two different populations. *Eur Spine J* 2006;15:1705–16.
7. Leung ASL, Lam T, Hedley AJ, et al. Use of a subjective health measure on Chinese low back pain patients in Hong Kong. *Spine* 1999;24:961–6.
8. Pellise F, Alvarez L, Escudero O, et al. Metric characteristics of the six-question 'core-set' in the evaluation of back pain. *Eur Spine J* 2003; 12(suppl):12–3.
9. Albert HB, Jensen AM, Dahl D, et al. Criteria validation of the Roland Morris questionnaire: a Danish translation of the international scale for the assessment of functional level in patients with low back pain and sciatica. *Ugeskr Læger* 2003;165:1875–80.
10. Chen SM, Liu MF, Wang BM, et al. Chinese translation and adaptation of the Roland-Morris Low Back Pain Disability Questionnaire. *FJPT* 2003; 28:324–32.
11. Coste J, Le Parc JM, Berge E, et al. French validation of a disability rating scale for the evaluation of low back pain (EIFEL questionnaire). *Rev Rhum* 1993;60:335–41.
12. Guermazi M, Mezghani M, Ghroubi S, et al. The Oswestry index for low back pain translated into Arabic and validated in an Arab population. *Ann Readapt Med Physique* 2005;48:1–10.
13. Pose B, Sangha O, Peters A, et al. Validation of the North American Spine Society Instrument for assessment of health status in patients with chronic backache. *Z Orthop Ihre Grenzgeb* 1999;137:437–41.
14. Ruof J, Sangha O, Stucki G. Evaluation of a German version of the Bath Ankylosing Spondylitis Functional Index (BASFI) and Dougados Functional Index (DFI). *Z Rheumatol* 1999;58:218–25.
15. Salaffi F, Stancati A, Silvestri A, et al. Validazione delle versioni italiane del Bath Ankylosing Spondylitis Functional Index (BASFI) e del Dougados Functional Index (DFI) in pazienti con spondilite anchilosante. *Reumatismo* 2005;57:161–73.
16. Gordon RG. *Ethnologue*: *Languages of the World*, 15th ed. Dallas: SIL International, 2005.
17. Kim DY, Lee SH, Lee HY, et al. Validation of the Korean version of the Oswestry Disability Index. *Spine* 2005;30:E123–7.
18. Streiner DL, Normam GR. *Health Measurement Scales*: *A Practical Guide to Their Development and Use*, 3rd ed. New York: Oxford University, 2003.
19. Linton SJ, Hallden K. Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and sub acute back pain. *Clin J Pain* 1998;14:209–15.
20. Grotle M, Vollestad NK, Brox JI. Screening for yellow flags in first-time acute low back pain: reliability and validity of the Acute Low Back Pain Screening Questionnaire. *Clin J Pain* 2006;22:458–67.
21. Ettinger B, Black DM, Nevitt MC. Contribution of vertebral deformities to chronic back pain and disability. *J Bone Miner Res* 1992;7:449–56.
22. Symonds TL, Burton AK, Tillotson KM, et al. Absence resulting from low back trouble can be reduced by psychosocial intervention at the work place. *Spine* 1995;20:2738–45.
23. Tesio L, Granger CV, Fiedler RC. A unidimensional pain/disability measure for low back pain syndromes. *Pain* 1997;69:269–78.
24. Stratford PW, Binkley JM, Riddle DL. Development and initial validation of the Back Pain Functional Scale. *Spine* 2000;25:2095–102.
25. Rytokoski U, Puukka P, Talo S. Biodisability in relation to other disease consequences in the functional assessment of patients with chronic low back pain. *Int J Rehabil Res* 1997;20:225–44.
26. Davidson M, Keating JL, Eyres S. A low back-specific version of the SF-36 Physical Functioning Scale. *Spine* 2004;29:586–94.
27. Garret S, Jenkinson T, Kennedy LG, et al. A new approach to defining disease status in ankylosing spondylitis: the development of the Bath An-

kylosing Spondylitis Disease Activity Index. *J Rheumatol* 1994;21:2286–91.
28. Cardiel MH, Londoño JD, Gutiérrez E, et al. Translation, cross-cultural adaptation, and validation of the Bath Ankylosing Spondylitis Functional Index (BASFI), the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) and the Dougados Functional Index (DFI) in a Spanish speaking population with spondyloarthropathies. *Clin Exp Rheumatol* 2003;21:451–8.
29. Waldner A, Cronstedt H, Stenstrom CH. The Swedish version of the Bath Ankylosing Spondylitis Disease Activity Index: reliability and validity. *Scand J Rheumatol* 1999;28:10–6.
30. Akkoc Y, Karatepe AG, Akar S, et al. A Turkish version of the Bath Ankylosing Spondylitis Disease Activity Index: reliability and validity. *Rheumatol Int* 2005;25:280–4.
31. Calin A, Garrett S, Whitelock H. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. *J Rheumatol* 1994;21:2281–5.
32. Udrea G, Ciobanu C, Mihai C, et al. Evaluation of a Romanian version of the Bath Ankylosing Spondylitis Functional Index (BASFI) in patients with spondylarthropathies. *Rom J Intern Med* 2004;42:199–209.
33. Karatepe AG, Akkoc Y, Akar S, et al. The Turkish versions of the Bath Ankylosing Spondylitis and Dougados Functional Indices: reliability and validity. *Rheumatol Int* 2005;25:612–8.
34. Yanik B, Gürsel YK, Kutlay S, et al. Adaptation of the Bath Ankylosing Spondylitis Functional Index to the Turkish population, its reliability and validity: functional assessment in AS. *Clin Rheumatol* 2005;24:41–7.
35. Bolton JE, Breen AC. The Bournemouth Questionnaire: a short-form comprehensive outcome measure: I. Psychometric properties in back pain patients. *J Manipulative Physiol Ther* 1999;22:503–10.
36. Hartigsen J, Lauridsen H, Ekstrom S, et al. Translation and validation of the Danish version of the Bournemouth questionnaire. *J Manipulative Physiol Ther* 2005;28:402–7.
37. Deyo RA, Battie M, Beurskens AJ. Outcome measures for low back pain research: a proposal for standardized use. *Spine* 1998;23:2003–13.
38. Harper AC, Harper DA, Lambert LJ, et al. Development and validation of the Curtin Back Screening Questionnaire (CBSQ): a discriminative disability measure. *Pain* 1995;60:73–81.
39. Lawlis GF, Cuencas R, Selby D. The development of the Dallas Pain Questionnaire: an assessment of the impact of spinal pain on behaviour. *Spine* 1989;14:511–6.
40. Marty M, Blotman F, Avouac B, et al. Validation of the French version of the Dallas Pain Questionnaire in chronic low back pain patients. *Rev Rhum Engl Ed* 1998;65:126–34.
41. Dougados M, Gueguen A, Nakache JP, et al. Evaluation of a functional index and an articular index in ankylosing spondylitis. *J Rheumatol* 1988;15:302–7.
42. Cronstedt H, Waldner A, Stenstrom CH. The Swedish version of the Bath Ankylosing Spondylitis Functional Index: reliability and validity. *Scand J Rheumatol* 1999;28(suppl):1–8.
43. Ozer HTE, Sarpel T, Gulek B, et al. Evaluation of the Turkish version of the Dougados functional index in ankylosing spondylitis. *Rheumatol Int* 2005;25:368–72.
44. Waddell G, Newton M, Henderson I, et al. A Fear Avoidance Beliefs Questionnaire (FABQ) and the role of fear avoidance beliefs in chronic low back pain and disability. *Pain* 1993;52:157–68.
45. Chaory K, Fayad F, Rannou F, et al. Validation of the French version of the Fear Avoidance Belief Questionnaire. *Spine* 2004;29:908–13.
46. Pfingsten M, Kroner–Herwig B, Leibing E, et al. Validation of the German version of the Fear–Avoidance Beliefs Questionnaire (FABQ). *Eur J Pain* 2000;4:259–66.
47. Kovacs FM, Muriel A, Medina JM, et al. Psychometric characteristics of the Spanish version of the FAB Questionnaire. *Spine* 2006;31:104–10.
48. Nork SE, Hu SS, Workman KL. Patient outcomes after decompression and instrumented posterior spinal fusion for degenerative spondylolisthesis. *Spine* 1999;24:561–9.
49. Feise JF, Menke M. Functional Rating Index: a new valid and reliable instrument to measure the magnitude of clinical change in spinal conditions. *Spine* 2001;26:78–87.
50. Hagg O, Fritzel P, Romberg K. The General Function Score: a useful tool for measurement of physical disability: validity and reliability. *Eur Spine J* 2001;10:203–10.
51. Kohlmann T, Raspe H. The Hannover Functional Ability Questionnaire for measuring low back pain-related limitations (FFbH-R). *Rehabilitation* 1996;35:I–VIII.
52. Greenough CG, Fraser RD. Assessment of outcome in patients with low–back pain. *Spine* 1992;17:36–41.
53. Holt AE, Shaw NJ, Shetty A, et al. The reliability of the Low Back Outcome for back pain. *Spine* 2002;27:206–10.
54. Manniche C, Asmussen K, Lauritsen B. Low Back Pain Rating Scale: validation of a tool for assessment of low back pain. *Pain* 1994;57:317–26.
55. Nuhr MJ, Crevenna R, Quittan M, et al. Cross-cultural adaptation of the Manniche questionnaire for German-speaking low back pain patients. *J Rehabil Med* 2004;36:267–72.
56. Walsh TL, Hanscom B, Lurie JD, et al. Is a condition-specific instrument for patients with low back pain/leg symptoms really necessary? The responsiveness of the Oswestry Disability Index, MODEMS, and the SF–36. *Spine* 2003;28:607–15.
57. Daltroy LH, Cats-Baril WL, Katz JN, et al. The North American Spine Society Lumbar Spine Outcome Assessment Instrument: reliability and validity tests. *Spine* 1996;21:741–9.
58. Padua R, Padua L, Ceccarelli E, et al. Cross-cultural adaptation of the Lumbar North American Spine Society Questionnaire for Italian-speaking patients with lumbar spinal disease. *Spine* 2001;26:344–7.
59. Sarasqueta C, Gabaldon O, Iza I, et al. Cross-cultural adaptation and validation of the NASS outcomes instrument in Spanish patients with low back pain. *Eur Spine J* 2005;14:586–94.
60. Kopec KA, Esdaile JM. Occupational role performance in persons with back pain. *Disabil Rehabil* 1998;20:373–9.
61. Stucki G, Daltroy LH, Liang MH. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine* 1996;21:796–803.
62. Pratt RK, Fairbank JCT, Virr A. The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. *Spine* 2002;27:84–91.
63. Fairbank JCT, Couper J, Davies JB. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980;66:271–3.
64. Chow JH, Chan CC. Validation of the Chinese version of the Oswestry Disability Index. *Work* 2005;25:307–14.
65. Mannion AF, Junge A, Fairbank JCT, et al. Development of a German version of the Oswestry Disability Index: 1. Cross-cultural adaptation, reliability, and validity. *Eur Spine J* 2006;15:55–65.
66. Boscainos PJ, Sapkas G, Stilianessi E, et al. Greek versions of the Oswestry and Roland-Morris Disability Questionnaires. *Clin Orthop* 2003;411:40–53.
67. Fujiwara A, Kobayashi N, Saiki K, et al. Association of the Japanese Orthopaedic Association score with the Oswestry Disability Index, Roland Morris Questionnaire and short form 36. *Spine* 2003;28:1601–7.
68. Grotle M, Brox JI, Vollestad NK. Cross-cultural adaptation of the Norwegian versions of the Roland–Morris Disability Questionnaire and the Oswestry Disability Index. *J Rehabil Med* 2003;35:214–47.
69. Mousavi SJ, Parnianpour M, Mehdian H, et al. 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* 2006;31:E454–9.
70. Yakut E, Duger T, Oksuz C, et al. Validation of the Turkish version of the Oswestry Disability Index for patients with low back pain. *Spine* 2004;29:581–5.
71. Gallon RL. Perception of disability in chronic back pain patients: a long term follow up. *Pain* 1989;37:67–75.
72. Kopec KA, Esdaile JM, Abrahamowicz M. The Quebec Back Pain Disability Scale: conceptualization and development. *J Clin Epidemiol* 1996;49:151–61.
73. Schoppink LE, van Tulder MW, Koes BW, et al. Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. *Phys Ther* 1996;76:268–75.
74. Yvanes-Thomas M, Calmels P, Bethoux F, et al. Validity of the French-language version of the Quebec back pain disability scale in low back pain patients in France. *Joint Bone Spine* 2002;69:397–405.
75. Bicer A, Yazici A, Camdeviren H, et al. Assessment of pain and disability in patients with chronic low back pain: reliability and construct validity of the Turkish version of the Quebec Back Pain Disability Scale and Pain Disability Index. *J Back Musculoskeletal Rehabil* 2005;18:37–44.
76. Roland M, Morris R. A study of the natural history of back pain: I. Development of a reliable and sensitive measure of disability of low back pain. *Spine* 1983;8:141–4.
77. Bejia I, Younes M, Kamel BS, et al. Validation of the Tunisian version of the Roland-Morris questionnaire. *Eur Spine J* 2005;14:171–4.
78. Tsang RCC. Measurement properties of the Hong Kong Chinese version of the Roland-Morris Disability Questionnaire. *Hong Kong Physiother J* 2004;22:40–9.
79. Brouwer S, Kuijer W, Dijkstra P, et al. Reliability and stability of the Roland Morris Disability Questionnaire. *Disabil Rehabil* 2004;26:162–5.
80. Wiesinger GF, Nuhr M, Quittan M, et al. Cross-cultural adaptation of the Roland Morris questionnaire for German-speaking patients with low back pain. *Spine* 1999;24:1099–103.
81. Padua R, Padua L, Ceccarelli E, et al. Italian version of the Roland Disabil-

ity Questionnaire, specific for low back pain: cross-cultural adaptation and validation. *Eur Spine J* 2002;11:126–9.
82. Nusbaum L, Natour J, Ferraz MB, et al. Translation, adaptation and validation of the Roland Morris questionnaire-Brazil Roland Morris. *Braz J Med Biol Res* 2001;34:203–10.
83. Kovacs FM, Llobera J, Gil Del Real MT, et al. Validation of the Spanish version of the Roland Morris questionnaire. *Spine* 2002;27:538–42.
84. Johansson E, Lindberg P. Sub acute and chronic low back pain. reliability and validity of a Swedish version of the Roland and Morris Disability Questionnaire. *Scand J Rehabil Med* 1998;30:139–43.
85. Jirarattanaphochai J, Jung S, Sumananont C, et al. Reliability of the Roland-Morris disability questionnaire (Thai version) for the evaluation of low back pain patients. *J Med Assoc Thai* 2005;88:407–11.
86. Kucukdeveci AA, Tennant A, Elhan AH, et al. Validation of the Turkish version of the Roland Morris Disability Questionnaire for use in low back pain. *Spine* 2001;26:2738–43.
87. Climent JM, Reig A, Sanchez J. Construction and validation of a specific quality of life instrument for adolescents with spine deformities. *Spine* 1995;20:2006–11.
88. Watanabe K, Hasegawa K, Hirano T, et al. Use of the Scoliosis Research Society Outcomes Instrument to evaluate patient outcome in untreated idiopathic scoliosis patients in Japan: 1. Comparison with non scoliosis group: preliminary /limited review in a Japanese Population. *Spine* 2005; 30:1197–201.
89. Climent JM, Bago R, Ey A, et al. Validity of the Spanish version of the Scoliosis Research Society-22 (SRS-22) Patient Questionnaire. *Spine* 2005; 30:705–9.
90. Bago J, Climent JM, Ey A, et al. The Spanish Version of the SRS-22 Patient Questionnaire for idiopathic scoliosis: transcultural adaptation and reliability analysis. *Spine* 2004;29:1676–80.
91. Alanay A, Cil A, Berk H, et al. Reliability and validity of adapted Turkish version of the Scoliosis Research Society-22 (SRS-22) Questionnaire. *Spine* 2005;30:2464–8.
92. Robinson RC, Kishino N, Matheson L, et al. Improvement in postoperative and nonoperative spinal patients on a self-report measure of disability: the Spinal Function Sort (SFS). *J Occup Rehabil* 2003;13:107–13.
93. Itzkovich M, Catz A, Tamir A. Spinal pain independence measure: a new scale for assessment of primary ADL dysfunction related to low back pain. *Disabil Rehabil* 2001;23:186–91.
94. Ruta DA, Garrat AM, Wardlaw D, et al. Developing a valid and reliable measure of health outcome for patients with low back pain. *Spine* 1994;19: 1887–96.
95. Osthus H, Cziske R, Jacobi E. A German version of the Extended Aberdeen Back Pain Scale. *Spine* 2006;31:571–7.
96. Atlas SJ, Deyo RA, van den Ancker M, et al. The Maine-Seattle Back Questionnaire: a 12-item disability questionnaire for evaluating patients with lumbar sciatica or stenosis. Results of a derivation and validation cohort analysis. *Spine* 2003;28:1869–76.
97. Hazard RG, Haugh LD, Reid S, et al. Early prediction of chronic disability after occupational low back injury. *Spine* 1996;21:945–51.
98. Waddell G, Main CJ. Assessment of severity in low back pain disorders. *Spine* 1984;9:204–8.
99. Sanders JO, Polly DW Jr, Cats-Baril WL, et al. Analysis of patient and parent assessment of deformity in idiopathic scoliosis using the Walter Reed Visual Assessment Scale. *Spine* 2003;18:2158–63.
100. Ozer HTE, Sarpel T, Gulek B, et al. The Turkish version of the Bath Ankylosing Spondylitis Functional Index: reliability and validity. *Clin Rheumatol* 2005;24:123–8.
101. Watanabe K, Hasegawa K, Hirano T, et al. Use of the Scoliosis Research Society Outcomes Instrument to evaluate patient outcome in untreated idiopathic scoliosis patients in Japan: 2. Relation between spinal deformity and patient outcomes. *Spine* 2005;30:1202–5.