J Rehabil Med 2003; 35: 241–247

Taylor&Francis
healthsciences

CROSS-CULTURAL ADAPTATION OF THE NORWEGIAN VERSIONS OF THE ROLAND-MORRIS DISABILITY QUESTIONNAIRE AND THE OSWESTRY DISABILITY INDEX

M. Grotle,[1] J. I. Brox[2] and N. K. Vøllestad[1]

From the [1]Section for Health Science, University of Oslo, [2]Department of Orthopaedics, National Hospital, Oslo, Norway

Objective: **To evaluate reliability and construct validity of the Norwegian versions of the Roland Morris Disability Questionnaire and the modified Oswestry Disability Index.**
Design: **Translation of two functional status questionnaires and a cross-sectional study of measurement properties.**
Methods: **The questionnaires were translated and back-translated following the Guillemin criteria. The Norwegian versions were tested for 55 patients with acute low back pain and 50 patients with chronic low back pain. Test-retest with a 2-day interval was performed in a subsample of 28 patients from the chronic sample. Reliability was assessed by repeatability according to Bland and Altman, intraclass coefficient and coefficient of variation. Internal consistency was assessed by Cronbach's alpha. Concurrent construct validity was assessed with correlations between the questionnaires and the SF-36, Disability Rating Index and pain intensity.**
Results: **Repeatability of the Roland Morris Disability Questionnaire was 4 points, coefficient of variation 15% and intraclass correlation coefficient 0.89, and of the modified Oswestry Disability Index 11, 12% and 0.88, respectively. Internal consistency was 0.94 for both questionnaires. The questionnaires correlated highly with the physical functioning scale of SF-36, moderately with pain, and low with mental scales of the SF-36.**
Conclusion: **The reliability and construct validity of the Norwegian versions of the Roland Morris Disability Questionnaire and the modified Oswestry Disability Index are acceptable for assessing functional status of Norwegian-speaking patients with low back pain.**

Key words: low back pain, functional status, Roland Morris Disability Questionnaire, Oswestry Disability Index, Norwegian version, reliability, construct validity.

J Rehabil Med 2003; 35: 241–247

Correspondence address: Margreth Grotle, Section for Health Science, University of Oslo, PO Box 1153 Blindern, N-0316 Oslo, Norway. E-mail: margreth.grotle@helsefag.uio.no*
**earlier name Soukup*

Submitted January 7, 2003; accepted March 28, 2003

INTRODUCTION

Assessment of functional status by self-report questionnaires and scales has become an important task for back pain clinicians and researchers. A large number of generic and disease-specific assessment instruments have been used to assess the functional status of back pain patients. In order to facilitate a unified and standardized outcome measurement in clinical studies, an international group of back pain researchers proposed that a selected group of questionnaires be used (1). Two condition-specific measures for spinal disorders were recommended; the Oswestry Disability Index (ODI) (2) and the Roland-Morris disability questionnaire (RMQ) (3).

Both the RMQ and ODI are outcome questionnaires developed in the UK during the 1980s. These questionnaires have been used widely in different studies and settings over many years, several of them reporting evidence for good measurement properties (4–7). A challenge with respect to a standardized use of these questionnaires, however, is the many versions developed and used during the last 10–15 years. For the RMQ there exist at least 4 modified versions (7–10) in addition to the original. Also for the ODI at least 4 versions exist (2, 11–13). The modified versions of both the RMQ and ODI are reviewed and discussed by the original authors (6, 14). They suggest using the original version of the RMQ and the modified version 2.0 of the ODI (2).

The RMQ and ODI have been translated into several non-English languages (1, 14–17). Validated translations are strongly encouraged, and should be developed by using recommended guidelines for cross-cultural adaptation of self-reported measures (18). Furthermore, the adapted versions should be evaluated according to basic measurement properties such as reliability and validity. Despite that a number of different Norwegian versions of the RMQ and ODI have been used (19–21), the translation and cross-cultural adaptation process of these questionnaires have never been reported earlier. Furthermore, to our knowledge, there are no studies reporting the measurement properties of the RMQ or the ODI, version 2.0, used in Norwegian patients. One study (20) has reported the measurement properties of version 1.0 of the ODI in patients with chronic low back pain (LBP). The objectives of this paper were to describe the process used to translate and adapt the proposed version of the RMQ and ODI into Norwegian, and to test these Norwegian versions in terms of test-retest reliability and construct validity among Norwegian patients with LBP.

 ISSN 1650–1977
DOI 10.1080/16501970310006452

METHODS

The study was carried out in 2 phases. First a translation and adaptation of the original version of the RMQ (3) and the modified version (2) of the original ODI (11) into Norwegian, including a pilot testing of the Norwegian versions, was carried out. Secondly, a thorough analysis of the measurement properties of the questionnaires was performed. In this paper the results regarding the reliability and concurrent construct validity are presented. In another paper the responsiveness of the questionnaires is analysed in a head-to-head comparison of several functional questionnaires.

Questionnaires

The Roland & Morris Disability Questionnaire (RMQ) is a self-report questionnaire consisting of 24 items related to normal activities of daily living. The RMQ was developed by deriving 24 relevant items from the longer Sickness Impact Profile (SIP) (22). The questionnaire was transformed to a condition-specific measurement of spinal disorders by adding the phrase "because of my back" to each statement. The patients are asked to circle those items, which they perceive as difficult to perform due to back pain. Each answer is scaled simply 0 or 1, thus leaving a range of scores of 0 to 24, a higher score indicating higher disability.

The Oswestry Low Back Pain Disability Index (ODI) was developed in a specialist referral clinic for patients with chronic LBP. The item selection was based on an interview designed to assess limitations of various activities (11). Version 2.0, which is proposed by the original authors and used in the current study, is a modification of the original ODI (2). This version is a self-report questionnaire of a patient's perceived disability based on 10 areas of pain and daily activities (pain intensity, personal hygiene, lifting, walking, sitting, standing, sleeping, sexual activity, social activity and travelling). Each section is scored on a 6-point scale (0–5), with 0 representing no limitation and 5 representing maximal limitation. The subscales combined add up to a total maximal score of 50. The score is then doubled and interpreted as a percentage of the patient-perceived disability (the higher the score, the greater the disability). In cases where patients did not answer all the 10 sections, the sum score of the answered sections were divided by the number of completed sections.

Translation

The English versions of the RMQ and ODI were translated into Norwegian by 2 different and independent bilingual translators, whose first language is Norwegian. One of the translators had a physiotherapy background and had been living and practising in USA, UK and Norway. The other translator was a professional translator with no health professional background. The first translator was aware of the process purpose and the concepts involved in the questionnaires; the other translator was unaware of these circumstances. After delivering each version of the "forward" translations, the 2 translators reviewed and discussed inconsistencies and differences in the translations, and a synthesis of the versions that both agreed upon, was formed.

The 2 synthesized versions (RMQ and ODI) of the 2 forward translators were then translated back into English by 2 other bilingual translators, whose first language is English. One of these translators has a medical background, and the other is a professional translator. As with the forward translators, the first translator was aware of the process purpose and the concepts involved in the questionnaires; the other translator was not.

To produce a version of the RMQ and ODI subjected to pilot testing, the various translations and back-translations were discussed by a review committee (the 2 translators and the researchers from the research group). Discrepancies between the various versions were resolved by consensus to achieve conceptual equivalence between the pre-final Norwegian version and the original English version of the RMQ and ODI. In this process also the other existing Norwegian versions of the RMQ and ODI were discussed. There were only minor differences between our pre-final version of the ODI and other Norwegian versions except to item 4. In other Norwegian versions 100 yards have been translated to 350 metres, which obviously should be 100 metres (approximately). In our pre-final version of the RMQ many items were different when compared with an earlier existing Norwegian version.

After creating the pre-final versions these were pre-tested in a small pilot study of 20 patients. Ten patients with acute LBP in the primary health care and 10 patients with chronic LBP from a back-clinic filled in the questionnaires. Afterwards the patients were asked about the comprehension of the various items and whether they experienced some difficulties with answering the questionnaires. Very few patients had comments that made any changes necessary. The final versions of the Norwegian RMQ and ODI (can be obtained from the corresponding author) were then subjected for further testing with regard to their reliability and validity when used for Norwegian patients with LBP.

Patients

The study was carried out in Fredrikstad and Sarpsborg, 2 cities in the south-east of Norway. Two different cohorts of patients with LBP between 18 and 60 years were used. The patients recruited from primary healthcare, consulted a medical doctor or chiropractor due to acute LBP of less than 2 weeks duration. Patients with chronic LBP had had complaints for more than 3 months duration. The patients were recruited from the Back Clinic in the Regional Hospital of Østfold. Pregnant women and patients with symptom and signs of cauda equina syndrome, progressive paresis, fracture, suspected tumour or local infection, ankylosing spondylitis, rheumatoid arthritis or other inflammation diseases were excluded. All patients gave their informed consent after receiving both written and oral information about the project. The Ethics Committee for Medical Research in Health Region I of Norway approved the study.

Procedure and measurements

The RMQ and ODI were administered to all patients as part of a comprehension questionnaire used in the cohorts. The comprehensive questionnaire consisted of sociodemographic data, medical history and current medical status, and different self-report measures of pain and functional status. The patients' reports of pain intensity in the lower back and the leg(s) were measured by a 100-mm Visual Analogue Scale (VAS) with the end points "no pain" and "severe pain". The patients were asked to mark the line at a point corresponding to the magnitude of their current pain, first related to their lower back and than the leg(s). The patients also completed *the Medical Outcome Study Short Form-36 questionnaire* (SF-36) (23) and *the Disability Rating Index* (DRI) (24).

The SF-36 consists of 36 questions on the general health status of patients, and provides 8 specific categories of physical and emotional scores: physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role emotional and mental health (23). Lower scores indicate limitations in functioning. The SF-36 is widely used as a standard measurement in validation studies. In this study the Norwegian version translated by Loge & Kaasa was used (25).

The Disability Rating Index (DRI) (24) is a self-administered questionnaire with 12 items divided into 3 sections: common basic activities of daily life; more demanding daily physical activities; and work-related or more vigorous activities. The questions are arranged in increasing order of physical demand, particularly with reference to LBP. The patients mark on a 100-mm VAS his/her presumed ability to perform the activity. The sum score of the DRI is presented as the mean of the 12 measurements. The DRI was developed in Sweden for the assessment of physical disability in patients with chronic pain in the neck, shoulder and lower back. A Norwegian translated version carried out at the Ullevål Hospital was used in this study.

A clinical examination was carried out on the patients, including a neurological and a back examination. To examine whether the patients experienced an increase in lumbar and/or leg pain during the procedure of answering several questionnaires, taking them approximately 30–40 minutes, and the clinical examination, we asked the patients to score their lumbar and leg pain (VAS) twice; when filling in the questionnaires, and after the clinical examination of the patients.

Reliability

For test-retest analyses the functional questionnaires were administered to a subsample of the patients with chronic LBP. It was assumed that the back condition of the patients with chronic LBP was more stable than that of the patients with acute LBP. Thirty consecutive patients with chronic LBP were asked to complete the questionnaires after 2 days, and return them by mail.

Since all the patients rated their pain intensity before and after the clinical examination, the test-retest reliability for lumbar and leg pain intensity was calculated for the acute and chronic sample. There was approximately 1 hour between the 2 pain ratings.

Validity

In lack of a gold standard to assess the validity of a construct such as functional status in patients with LBP, a discriminant approach was used. In this approach, *a priori* hypotheses based on well-established associations between functional status and different types of variables that have been found and replicated in several studies, are made. Because these questionnaires intend to assess disability in terms of limitations in daily activities, it was expected that disability assessed by the RMQ and ODI (summary scores) would be significantly correlated with limitations in physical functioning according to the SF-36 and DRI. Secondly, since the questionnaires were constructed to assess pain-related disability, it was hypothesized that the sum scores would be moderately correlated with pain. Furthermore, it was expected that these questionnaires would provide information about a concept that was distinguishable from general psychosocial concerns. It was therefore hypothesized that the sum scores of RMQ and ODI would show minimal correlations with the mental and general health perception scores of the SF-36.

Statistical analysis

The Statistical Package of Social Science (SPSS), version 10.0, was used to analyse the data. The mean and standard deviation (SD) or frequencies were calculated for numerical and categorical variables, respectively. In the test-retest sample mean and SD was calculated for the 2 time-points, as well as for the change scores of the measures. Differences between the scores were compared by paired *t*-tests.

To establish test-retest reliability of the two questionnaires, different statistical measures of reliability are presented. Firstly, as recommended by Bland & Altman (26) the measurement error is presented by repeatability, which expresses the measurement error in the same unit as the questionnaire, in this case RMQ points (range 0–24) and ODI percentages (range 0–100). This measure is based on the standard error of measurement (SEM) between test and retest scores, and was calculated by taking the square root of the mean square error term from the usual reliability study analysis-of-variance-table. The repeatability was calculated by multiplying the SEM by 2.77 to correspond with the 95% confidence intervals, meaning that the difference between 2 measurements for the same subject is expected to be less than $2.77 \times \text{SEM}$ for 95% of pairs of observations. Since this value defines the smallest difference that can be detected between 2 measurements, it is also referred to as the *minimal detectable change* of a measure (27).

Secondly, the test-retest reliability was calculated by the intraclass correlation coefficient (ICC). The model denoted ICC (1, 1) (28), which suggests all within-variation to be measurement error, was used. The estimates of ICC (1, 1) was obtained from a one-way random effects model, in which the one-way analysis of variance yielded a between-subject mean square (BMS) and a within-subject mean square (WMS) in the following form:

$$\text{ICC}(1,1) = \frac{\text{BMS} - \text{WMS}}{\text{BMS} + (\text{k} - 1)\text{WMS}}$$

where k is the number of judges rating each subject. Since k is one in this situation, the (k–1) WMS = 0.

Thirdly, the average coefficient of variance (CV) for paired measurements was calculated. In cases with no statistical difference between the pairs of items, the CV was calculated as the ratio of the standard deviation divided by the mean and multiplied by 100 to yield a unitless percentage. The internal consistency for the questionnaires was assessed with Cronbach's alpha.

In addition, plots of difference between the first and second response on the questionnaires against a mean of the sum scores were constructed according to Bland and Altman's recommendations (26).

Associations between the sum scores of the questionnaires and other parameters were measured by Pearson's correlation test, because the variables were parametric or normally distributed.

Table I. *Demographic and clinical characteristics of the patients with acute low back pain (LBP). Mean values are presented with SD, and frequencies in numbers (n) and percentage (%)*

	Acute LBP ($n = 55$)	Chronic LBP ($n = 50$)
Women/men (*n* (%))	40/15 (73/27)	31/19 (62/38)
Age (years)	38 (10)	40 (9)
Working status (*n* (%))		
Full or part time remunerative work	33 (60)	8 (16)
Home working, students, non-employed, and others	10 (19)	7 (14)
Sick leave (*n* (%))	11 (20)*	35 (70)**
Pain location (*n* (%))		
LBP without radiating pain	31 (57)	6 (12)
LBP with radiating pain to one of the extremities	15 (27)	36 (72)
LBP with radiating pain to both extremities	9 (16)	8 (16)
Used pain medication the last 2 days (*n* (%))	23 (42)	20 (40)
Duration of current episode of LBP (mean days)	9.5 (7)	580 (785)
Pain intensity lower back (VAS 0–100)	48 (23)	43 (23)
Pain intensity lower extremity (VAS 0–100)	13 (21)	36 (24)
Roland Morris Disability Questionnaire (0–24)	9 (5)	10 (4)
Oswestry Disability Index (0–100)	28 (15)	32 (11)
Disability Rating Index	47 (21)	50 (17)
SF-36 (0–100)***		
Physical functioning	62 (24)	57 (20)
Role physical	34 (38)	12 (23)
Social functioning	81 (22)	70 (18)
Bodily pain	40 (22)	32 (18)
Role emotional	64 (41)	42 (42)
Mental health	78 (18)	69 (16)
Vitality	49 (22)	40 (20)
General health perception	73 (25)	59 (23)

*Three of the patients had disability pension due to other diseases/disorders than LBP.
**Four of the patients had disability pension due to LBP.
***Short-Form-36 Questionnaire range from 0–100 with higher scores indicating better health.

Table II. *Test and retest scores (after 2–4 days) of the functional status questionnaires. Test-retest reliability is expressed by intraclass correlation coefficient (ICC (1,1)), repeatability and coefficient of variance (CV)*

Test-retest sample ($n = 28$)	Mean (SD) of test	Mean (SD) of retest	ICC (1,1)	Repeatability	CV (%)
RMQ	9.8 (4.1)	9.9 (4.7)	0.89	4	15
ODI	32 (11)	32 (12)	0.88	11	12

RMQ = Roland Morris Disability Questionnaire; ODI = Oswestry Disability Index, modified version 2.0.

RESULTS

Fifty patients with chronic LBP and 55 patients with acute LBP completed both the RMQ and ODI. Generally, the patients were able to fill in the questionnaires without help. There were few missing values. Five of the patients had 1 missing item in the ODI. The demographic and clinical variables of the samples are presented in Table I.

Twenty-eight of the 30 patients (93.3%) with chronic LBP used for testing the test-retest reliability of the questionnaires returned the retest-questionnaires. There were no significant differences in demographic data between these 30 patients and the rest of the chronic LBP cohort. There were no statistical significant differences between the first and second completions of the functional questionnaires. Table II summarizes the test-retest reliability between the first and second completion of the RMQ and ODI expressed by the ICC, CV and repeatability according to Bland & Altman (26). The ICCs were about similar for the RMQ and ODI. However, the repeatability and CV were slightly better for the ODI than for the RMQ. This tendency is also visualized in Fig. 1. Internal consistency by Cronbach's alpha was 0.94 for both the RMQ and the ODI.

Table III shows the test-retest reliability of pain intensity measured by VAS in the patients with acute and chronic LBP. Both the acute and chronic patients rated their pain statistically significantly lower after the clinical examination than before. High variability in both lumbar and leg pain was found. The repeatability was between 27 and 31. This means that a change within 27–37 mm on a 100-mm VAS could be attributed to measurement error or random variation in a single patient. There was a tendency for lower repeatability in the acute sample compared with the chronic sample.

Table III. *Test and restest scores of pain measured by VAS in the patients with acute and chronic low back pain (LBP). Test-retest reliability is expressed by intraclass correlation coefficient (ICC (1,1)) and repeatability*

	Mean (SD) of test	Mean (SD) of retest	ICC (1,1)	Repeatability
Acute LBP ($n = 52$)				
Lumbar pain	47 (23)	39 (23)	0.67	37
Leg pain	13 (21)	8 (16)	0.51	36
Chronic LBP ($n = 50$)				
Lumbar pain	43 (23)	33 (23)	0.83	27
Leg pain	36 (24)	29 (22)	0.76	31

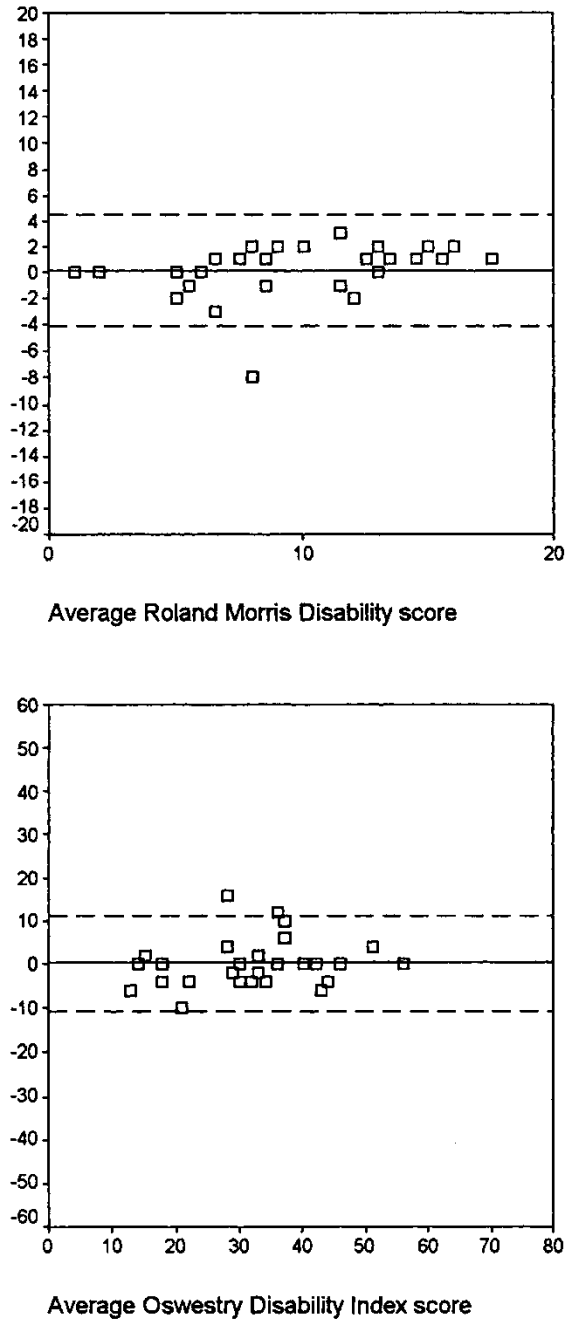

Fig. 1. Intraindividual differences (n = 28) between questionnaire responses on time 1 and 2 plotted against the mean of the sum scores. On each plot, the central horizontal line represents the mean of the intra-individual differences, and the flanking lines represent the 95% limits of agreement (26).

Table IV. *Pearson's correlation coefficient of the sum score of the Roland Morris Disability Questionnaire (RMQ) and Oswestry Disability Index (ODI) with other functional status and pain measures, as well as the subscales of the SF-36*

	Correlation with RMQ		Correlation with ODI	
	Acute LBP $n = 55$	Chronic LBP $n = 50$	Acute LBP $n = 55$	Chronic LBP $n = 50$
Lumbar pain (VAS)	0.32	0.47	0.39	0.52
ODI	0.73	0.60	–	–
Disability rating index	0.68	0.30	0.81	0.54
SF-36				
Physical functioning	−0.74	−0.60	−0.78	−0.77
Role physical	−0.45	−0.39	−0.38	−0.47
Social functioning	−0.50	−0.37	−0.52	−0.60
Bodily pain	−0.19 (n.s.)	−0.33	−0.28	−0.64
Role emotional	−0.21 (n.s.)	−0.19	−0.43	−0.33
Mental health	−0.37	−0.22	−0.57	−0.37
Vitality	−0.19 (n.s.)	−0.08 (n.s.)	−0.35	−0.28
General health perception	−0.13 (n.s.)	−0.13 (n.s.)	−0.31	−0.50

LBP = low back pain; SF-36 = Short-Form-36 Questionnaire. All correlations between the scales were statistically significant; ($p < 0.05$) except for those marked n.s.

Table IV shows the associations between the summary scores of the RMQ and ODI and other variables used to establish the construct validity of the questionnaires. As expected, a strong correlation was found between the summary scores and physical functioning according to the SF-36 and between the RMQ and ODI. In the acute sample high correlations were also found with the DRI, but in the chronic sample the RMQ and ODI correlated only low to moderately with the DRI. The ODI showed higher correlations with pain in patients with chronic LBP than in patients with acute LBP. In particular, there was a high correlation between the ODI and the bodily pain scale of SF-36. In the patients with acute LBP the bodily pain scale correlated poorly both with the RMQ and ODI. Table IV also shows that the RMQ correlated weakly with the Role emotional, Mental health, Vitality and General health perception scales of the SF-36. These scales were, however, moderately correlated with the ODI in the chronic sample, suggesting that the construct assessed in the ODI to some extent is related to psychosocial perceptions.

DISCUSSION

This paper reports the translation process of two frequently used back-specific outcome questionnaires, the original version of the RMQ and version 2.0 of the ODI, and presents the results from the first part of the psychometric testing. In this study the cross-cultural adaptation procedure described by Guillemin et al. (18) is used, which represents a more thorough adaptation process than a mere literary translation. The Norwegian version of the RMQ and modified ODI appeared to be clearly understood and easily administered by the patients participating in this study.

We used 3 different statistical approaches to assess measurement error associated with the RMQ and ODI; the repeatability according to Bland & Altman (26), CV and ICC. The ICC has been frequently used for assessing the reliability of the RMQ (7, 8, 10, 15, 16) and the ODI (2, 11, 20, 29). The ICC values in the current study are in accordance with other studies testing the same questionnaire versions (15, 16). However, there are limitations by using the ICC. Firstly, like the usual Pearson correlation coefficient, it measures the strength of the relation between 2 variables, and not the agreement between them (30). Secondly, the ICC is strongly affected by the variation between the patients in a sample, for example, the ICCs can be increased by extending the variability of the sample (30). Our results show that the use of ICC can give a misleading high estimate of reliability when compared with measurement error according to repeatability and CV.

Repeatability in the current study was 4 points for the RMQ and 11 points for the ODI. Since a change less than these points is indistinguishable from the measurement error, this limit has been labelled *the minimal detectable change*. Stratford et al. (27) have showed that the error measurement is dependent upon where on the scale the change is occurring. A variance of 4–5 RMQ points was sufficiently small to detect change in patients with initial scores in the central portion of the scale (4–20 RMQ points), but too large to detect improvement in patients with scores of less than 4 and deterioration in patients who had scores greater than 20 (27). In the present study most patients scored in the central portion of the RMQ. Only 3 patients of the test-retest sample scored 4 or less and none scored more than 20. Hence, the results of the present study are most valid for scorings in the central portion of the RMQ.

Our estimations of measurement error for the ODI (11 points) are almost identical to what other authors have found for the modified version of ODI (31) and the original version of ODI (20). Further studies should be carried out to examine whether measurement error in the ODI is dependent on the level of disability, as with the RMQ.

The CV, which reflects the relative measurement error in a group, might be important when planning a study in which one of these questionnaires will be used. For example, a CV of 12% for the ODI means that when the ODI is used in repeated measurements, about 12% of the change is attributed to measurement error. The use of CV showed that the relative measurement error was lower in the ODI than in the RMQ.

The test-retest reliability of pain assessed on a VAS was low despite a very short interval between the measurements. Because the results differed significantly between the test and retest, the CV could not be assessed for the pain assessment. The reliability was generally lower for the acute than the chronic patients. The significant reduction of pain was an unexpected finding, because one of the reasons for including the 2 pain ratings was to examine whether the clinical examination provoked an increase in back or leg pain. This finding confirms the considerable random variation of pain reports in patients with LBP.

As expected, there was a strong correlation between the summary score of the RMQ, ODI and the physical functioning scale of the SF-36. Strong correlations between the RMQ and ODI have been demonstrated in several studies (2, 16, 32), and also between the physical functioning scale (SF-36) and the RMQ (17) and the ODI (33, 34). The DRI correlated well with the RMQ and ODI in the patients with acute LBP, but not so well in the patients with chronic LBP. In the original study of the DRI also a low correlation of 0.38 was found between the DRI and the original version of the ODI (24). A possible explanation for this finding might be different content of the questionnaires. The DRI strictly measures limitations in specific daily activities, while both the RMQ and ODI include other dimensions such as pain, sleep and social activities. Our results thus indicate that these dimensions are affecred more in patients with chronic LBP.

An important issue when evaluating construct validity of these questionnaires is whether they actually measure limitations in daily life activities (disability), which they intend to do, or whether other constructs are measured. When completing self-report questionnaires there is a possibility that the patient may be expressing general distress or describing the ability to cope rather than providing specific information about the limitations in daily activities. The findings which showed that the RMQ correlated highly with physical function and low with psychological variables such as role emotional, vitality and mental health, suggest that RMQ mostly reflects the physical aspects of disability. These results are in accordance with one of the earliest validation studies of the RMQ carried out by Deyo (5). In their study the RMQ was compared with its longer parent SIP-questionnaire (22) and the RMQ correlated highly with the physical dimension of the SIP, and less with the psychosocial dimension of the SIP. The ODI also correlated highly with physical function, but moderately with the psychosocial scales of the SF-36. Similar findings have been reported by others (33–35). In the current study both the RMQ and ODI correlated moderately with lumbar pain. These findings are similar to those obtained when testing the German (17), Swedish, (15) and Spanish (16) versions of the RMQ, and in a previous Norwegian study of the original version of the ODI (20). The moderate correlations between lumbar pain and ODI are in keeping with a recent Rasch analysis of the ODI demonstrating that the pain item (item 1) measured another dimension than rest of the items in ODI (36). Further studies are needed concerning the construct validity of the ODI.

CONCLUSION

In conclusion the results of this study suggest that the test-retest reliability and the construct validity of the Norwegian version of the RMQ and ODI are acceptable for assessing self-reported functional status of Norwegian-speaking patients with LBP.

ACKNOWLEDGEMENT

We want to thank the Norwegian Fund for Post Graduate Training in Physiotherapy for financial support which made this research possible.

REFERENCES

1. Deyo RA, Battie M, Beurskens AJ, Bombardier C, Croft P, Koes B, et al. Outcome measures for low back pain research. A proposal for standardized use. Spine 1998; 23: 2003–2013.
2. Baker DJ, Pynsent PB, Fairbank JCT. The Oswestry disability index revisited: its reliability, repeatability and validity, and a comparison with the St Thomas's disability index. In: Roland MO, Jenner JR, eds. Back pain: new approaches to rehabilitation and education. Manchester: Manchester University Press; 1990, p. 175–181.
3. Roland M, Morris R. A study of the natural history of back pain. Part I: Development of a reliable and sensitive measure of disability in low back pain. Spine 1983; 8: 141–144.
4. Delitto A. Are measures of function and disability important in low back care? Physical Therapy 1994; 74: 452–462.
5. Deyo RA. Comparative validity of the sickness impact profile and shorter scales for functional assessment in low-back pain. Spine 1986; 11: 951–954.
6. Fairbank JCT, Pynsent PB. The Oswestry disability index. Spine 2000; 25: 2940–2953.
7. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. Spine 1995; 20: 1899–909.
8. Stratford PW, Binkley JM. Measurement properties of the RM-18. A modified version of the Roland-Morris disability scale. Spine 1997; 22: 2416–2421.
9. Dionne CE, Koepsell TD, Von Korff M, Deyo R, Barlow W, Checkoway H. Predicting long-term functional limitations among back pain patients in primary care settings. J Clin Epidemiol 1997; 50: 31–43.
10. Underwood MR, Barnett AG, Vickers MR. Evaluation of two time-specific back pain outcome measures. Spine 1999; 24: 1104–1112.
11. Fairbank JCT, Davies JB, Couper J, O'Brien JP. The Oswestry low back pain disability questionnaire. Physiotherapy 1980; 66: 271–273.
12. Hudson-Cook N, Tomes-Nicholson K, Breen A. A revised Oswestry disability questionnaire. In: Roland MO, Jenner JR, eds. Back Pain: new approaches to rehabilitation and education. Manchester: Manchester University Press; 1989, p. 187–204.
13. Fairbank JC. Use of Oswestry disability index (ODI). Spine 1995; 20: 1535–1537.
14. Roland MO, Fairbank JC. The Roland-Morris disability questionnaire and the Oswestry disability questionnaire. Spine 2000; 25: 3115–3124.
15. Johansson E, Lindberg P. Subacute and chronic low back pain. Reliability and validity of a Swedish version of the Roland Morris disability questionnaire. Scand J Rehabil Med 1998; 30: 139–143.
16. Kovacs FM, Llobera J, Gil Del Real MT, Abraira V, Gestoso M, Fernandez C, et al. Validation of the Spanish version of the Roland-Morris questionnaire. Spine 2002; 27: 538–542.
17. Wiesinger GF, Nuhr M, Quittan M, Ebenbichler G, Wolfl G, Fialka-Moser V. Cross-cultural adaptation of the Roland-Morris questionnaire for German-speaking patients with low back pain. Spine 1999; 24: 1099–1103.
18. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of

health-related quality-of-life measures: literature review and proposed guidelines. J Clin Epidemiol 1993; 46: 1417–1432.
19. Dullerud R, Amundsen T, Lie H, et al. Clinical results after percutaneous automated lumbar nucleotomy: a follow-up study. Acta Radiologica 1995; 36: 418–424.
20. Holm I, Friis A, Storheim K, Brox JI. Measuring self-reported functional status and pain in patients with chronic low back pain by postal questionnaires. A reliability study. Spine 2003; 28: 828–833.
21. Torstensen TA, Koes B, Glomsrod E, Seifert L, Glomsrod B, Hagen O. The effect of conservative therapies on patients suffering from chronic low back pain. Spine 1996; 21: 656–657.
22. Bergner M, Bobbitt RA, Carter WB, Gibson BS. The Sickness Impact Profile: development and final revision of a health status measure. Med Care 1981; 19: 787–805.
23. Ware JE, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). Med Care 1992; 30: 473–483.
24. Salèn BA, Spangfort EV, Nygren ÅL, Nordemar R. The Disability Rating Index: an instrument for the assessment of disability in clinical settings. J Clin Epidemiol 1994; 47: 1423–1434.
25. Loge JH, Kaasa S. Short Form 36 (SF-36) health survey; normative data from the general Norwegian population. Scand J Soc Med 1998; 26: 250–258.
26. Bland JM, Altman DG. Measuring agreement in method comparison studies. Statistics in Medicine 1999; 8: 135–160.
27. Stratford P, Finch E, Solomon P, et al. Using the Roland-Morris questionnaire to make decisions about individual patients. Physiotherapy Canada 1996; 48: 107–110.
28. Shrout PE, Fleiss JL. Intraclass Correlations: uses in assessing rater reliability. Psychological Bulletin 2002; 86: 420–428.
29. Grönblad M, Hupli M, Wennerstrand P, et al. Intercorrelation and test-retest reliability of the Pain Disability Index and the Oswestry disability questionnaire and their correlation with pain intensity in low back pain patients. Clin J Pain 1993; 9: 189–195.
30. Bland JM, Altman DG. A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. Computers in Biology & Medicine 1990; 20: 337–340.
31. Fritz JM, Irrgang JJ. A comparison of a modified Oswestry low back pain disability questionnaire and the Quebec back pain disability scale. Phys Ther 2001; 81: 776–788.
32. Leclaire R, Blier F, Fortin L, Proulx RA. Cross-sectional study comparing the Oswestry and Roland-Morris functional disability scales in two populations of patients with low back pain of different levels of severity. Spine 1997; 22: 68–71.
33. Grevitt M, Khazim R, Webb J, Mulholland R, Shepperd J. The Short Form-36 health survey questionnaire in spine surgery. Journal of Bone & Joint Surgery – British Volume 1997; 79: 48–52.
34. Taylor SJ, Taylor AE, Foy MA, Fogg AJ. Responsiveness of common outcome measures for patients with low back pain. Spine 1999; 24: 1805–1812.
35. Greenough CG, Fraser RD. Assessment of outcome in patients with low-back pain. Spine 1992; 17: 36–41.
36. Page SJ, Shawaryn MA, Cernich AN, Linacre JM. Scaling of the revised Oswestry low back pain questionnaire. Arch Phys Med 2002; 83: 1579–1584.