Scand J Rehab Med 30: 139–143, 1998

SUBACUTE AND CHRONIC LOW BACK PAIN. RELIABILITY AND VALIDITY OF A SWEDISH VERSION OF THE ROLAND AND MORRIS DISABILITY QUESTIONNAIRE

Eva Johansson,, RPT and Per Lindberg, PhD

Department of Public Health and Caring Science, Uppsala University, Uppsala, Sweden

ABSTRACT. The purpose of this study was to investigate test-retest reliability and concurrent validity of a Swedish version of the Roland and Morris Disability Questionnaire (RM-Sw), and to describe demographic factors in patients with low back pain of at least 4 weeks' duration seeking outpatient physical therapy treatment in primary care settings. Seventy-two patients participated in the study. The intraclass correlation coefficient for a one-week test-retest interval was 0.88. There was moderate positive correlation with measures of perceived disability ($r = 0.64$, $p < 0.001$; $r = 0.69$, $p < 0.001$) and pain severity ($r = 0.54$, $p < 0.001$), and low negative correlation with measures of perceived life control ($r = -0.32$, $p < 0.01$) and general activity ($r = -0.27$, $p < 0.05$). Gender, education and occupation were only moderately related to RM-Sw scores, explaining 14% of the variance in the scores. It is concluded that RM-Sw is a reliable and valid measure of functional ability in low back pain.

Key words: back pain; disability; functional measure; measurement; physiotherapy.

INTRODUCTION

The impact of low back pain (LBP) is largely related to patient functioning, that is, how limited a person is in performing everyday activities. Although a number of potential outcomes (e.g. symptoms, range of motion, strength, functional ability, psychologic measures) can be assessed in cases of LBP, measures of functional ability may be the most relevant to both patients and society (7).

Since the 1980s, a number of functional ability questionnaires for patients with LBP have been developed (7, 8, 16). Among the most widely used and evaluated LBP-specific questionnaires are the Oswestry low back disability questionnaire (OSW) (12), the Million visual analogue scale (20), the Roland & Morris disability questionnaire (RM) (23), and the Waddell & Main disability index (30). These instruments seem to have satisfactory reliability and validity (4).

Responsiveness, i.e. the ability of an instrument to detect clinically important changes, has been more extensively studied in OSW and RM than in the other two questionnaires. In one direct comparison, OSW and RM were found to have the same responsiveness in a sample of patients with mechanical LBP who were referred for outpatient physical therapy treatment (26). However, OSW had a higher frequency of blank and multiple response items than RM. In another study, both RM and OSW were found to discriminate between improvement and non-improvement in a sample of patients with LBP duration more than 6 weeks, but OSW appeared to be less responsive to change than RM (3). Other authors have reported that RM seems to be more responsive than OSW in detecting changes when patients have minor disabilities, but less sensitive when patients are more severely disabled (1). Recently, Stratford et al. (25) reported that the minimum level of detectable change for RM varies between 4 and 5 points at the 90% confidence level in a sample of LBP outpatients.

Validity issues in RM have mostly been addressed in patients with chronic LBP referred to hospital-based inpatient or outpatient pain clinics (10, 14, 19). There are indications that patients with chronic LBP treated in pain clinics differ from patients with chronic LBP seen in primary care settings. Demographic factors such as education, occupation, work status, use of medication, and earlier back surgery were especially relevant in discriminating between groups (6, 9).

In summary, there are several LBP-specific functional ability questionnaires, of which RM seems an appropriate choice for evaluation of treatment in LBP patients, especially when relatively low levels of disability might be expected. To our knowledge, there are no published

data on reliability and validity of a Swedish version of RM.

The purpose of this study was to investigate test-retest reliability and concurrent validity of a Swedish version of RM (RM-Sw) in a consecutive sample of patients with LBP of at least 4 weeks' duration referred to or seeking outpatient physical therapy treatment in primary care settings. A further aim, due to reported differences between patients with chronic LBP treated in pain clinics and in primary care settings, was to describe the demographic and background characteristics of the sample (duration, education, occupation, working status, medication, and earlier back surgery). Finally, since it was anticipated that a large proportion of the participants would be female, that the education level would be high, and that non-manual occupations would be reported more frequently than manual occupations, it was important also to explore the relationships among gender, education, occupation and RM scores.

METHOD

Setting, sample, and procedure

The study was conducted in the county council primary health care organization of a Swedish university town with a population of 180,000, and in 3 surrounding rural communities with a total population of 50,000. In the city, public service, administration and "white collar" branches of business dominate. In the rural area, various small industries and farming dominate.

Participants in the study were recruited among persons seeking care at the physical therapy departments within the county council primary health care organization, or at the physical therapy department service for staff at the university hospital. Referral by a physician was not required, and it is estimated that about 30% of persons seeking care were self-referred.

The inclusion criteria were that they be between ages 18–65, with no signs of trauma, no malignant, infectious or systemic diseases, an ability to understand written and spoken Swedish, and a duration of LBP for at least 4 weeks. The latter criterion was chosen because of the recommendation that treatment to restore physical and functional capacity in patients approaching chronicity should start at about this time (21).

Persons seeking care were informed both verbally and in writing about the objectives of the investigation. After giving informed consent to participate, subjects were given the questionnaires and a brief form to obtain demographic and background data. All questionnaires were returned by mail, and data were collected during a period of $1\frac{1}{2}$ years.

All 82 subjects asked agreed to participate in the study. Of these, 72 completed and returned the questionnaires, making the response rate 88%. Due to missing data, n varies in the different analyses. Fifty-four subjects who had to wait for more than one week for their first appointment completed RM-Sw a second time. Those data were used for the one-week test-retest analysis.

Measures

The Roland and Morris Disability Questionnaire (RM) consists of 24 items considered relevant for patients with back problems (23). These items are derived from the Sickness Impact Profile (2), a generic indicator of perceived health status developed for use in a variety of chronic conditions. The RM is structured according to a "yes/no" format, where each "yes" scores one point. The total range is 0–24 points, with higher scores indicating more severe disability. The time frame is defined as "today". The RM was translated into Swedish by the present authors, and the translation was checked by a bilingual person whose native language is English.

The Pain Disability Index attempts to measure the degree to which seven areas of life are presently disrupted by chronic pain (22, 28, 29). The areas covered are family/home responsibilities, recreation, social activity, occupation, sexual behaviour, self-care, and life-supporting activities. The response format is a "calibrated visual analogue scale" where 0 = no disability, and 10 = total disability. A high total score thus indicates more perceived disability. The time frame is not defined. The PDI was also translated into Swedish by the present authors, using the same procedure as described above.

Four sub-scales from the Multidimensional Pain Inventory (15) were also used. The Multidimensional Pain Inventory assesses a range of psychosocial effects in patients with chronic pain. It is a 60-item inventory with 9 sub-scales: pain severity, interference, life control, affective distress, support, punishing responses, solicitous responses, distracting responses, and general activity.

The Pain Severity scale consists of 3 items that assess the perceived severity of pain (pain right now, average pain last week, suffering due to pain). The response format is a seven grade Likert-type scale where 0 = no pain/suffering, and 6 = extreme pain/suffering. The scores of each item are added and then divided by the number of items in the scale to create a mean sub-scale score. A high mean sub-scale score thus indicates worse perceived pain severity.

The Interference scale includes 11 items that assess the extent to which chronic pain interferes with life. The time frame is not defined. The response format (0 = no interference/change, 6 = extreme interference/change) and scoring are the same as for the Pain Severity scale. A high mean sub-scale score thus indicates more perceived interference.

The Life Control scale consists of 4 items that assess the degree to which patients perceive they have control in life (0 = not at all, 6 = extremely well). The response format and scoring are the same as for the Pain Severity scale. A high mean sub-scale score thus indicates higher perceived life control.

The General Activity scale includes 18 items that assess how often patients perform activities at home, around the house, away from home, and in social settings (0 = never, 6 = very often). The response format and scoring are the same as for the Pain Severity scale. A high mean sub-scale score thus indicates higher reported activity level. An already-existing Swedish version of the Multidimensional Pain Inventory was used (5). These sub-scales were chosen because they were hypothesized to be positively (Pain Severity, Interference) or negatively (Life Control, General Activity) correlated with RM.

Data analysis

Statistical analyses were performed in Statistica™ (24). For test-retest reliability, intraclass correlation was calculated. Correlations with other measures were analysed with Pearson's product moment correlation. Relations among gender, educa-

Table I. *Demographic and background characteristics of the sample* (n = 72)

Characteristics		*n*	%
Gender	women	57	79
	men	15	21
Married/cohabiting	yes	51	72
	no	20	27
	no report	1	1
Education	nine-year school	13	18
	senior high school	27	38
	university	31	43
	no report	1	1
Occupation	manual	23	32
	non-manual	36	50
	student	4	6
	self-employed	1	1
	no report	8	11
Working status	working	55	76
	sick-listed	12	17
	unemployed	5	7
Using analgesics[a]	yes	40	56
	no	29	40
	no report	3	4
Earlier back surgery		0	0

[a] NSAID, ASA or paracetamol only were reported.

tion, occupation and RM scores were analysed via standard multiple regression. Prior to regression analysis, the variables "education" and "occupation" were further divided into "low" and "high" education, with senior high school and university in the "high" category, and into "manual" and "non-manual" occupations, with students and self-employed people in the "non-manual" category. The level of significance was set at 0.05.

RESULTS

The participants had a mean age of 43 years (SD 10.7, range 20–64). The median duration of low back pain was 7 months (Q 10.5, range 1–144). The mean score for RM-Sw in the present sample ($n = 72$) was 9.3 points (SD 5.0). The median score was 9 points (Q 4.5). The reported scores ranged between 0 and 23 points, and did not deviate significantly from the normal distribution. The test-retest reliability for a one-week interval ($n = 54$) yielded an intraclass correlation coefficient of 0.88, which indicates that RM-Sw is a stable measure. The RM-Sw showed moderate positive correlation ($n = 64$) with the Pain Disability Index ($r = 0.64$, $p < 0.001$), the Interference scale ($r = 0.69$, $p < 0.001$), and the Pain Severity scale ($r = 0.53$, $p < 0.001$). There was low negative correlation with the Life Control scale ($r = -0.32$, $p < 0.01$) and the General Activity scale ($r = -0.27$, $p < 0.05$).

The demographic and background characteristics of the sample are presented in Table I. The majority of the participants were female, married or cohabiting, participants had a high level of education, and 50% reported having non-manual occupations (Table I).

A standard multiple regression analysis ($n = 64$) showed an overall significant relation between gender, education, occupation and RM-Sw scores ($F[3, 60] = 3.36$, $p < 0.05$), explaining 14% of the variance. Gender was the only variable that significantly contributed to the regression solution ($\beta = -0.25$, $p < 0.05$), implying that RM-Sw scores were predictable by gender—i.e. females had lower RM-Sw scores.

DISCUSSION

The results of this study suggest that RM-Sw is a reliable and valid measure of functional ability for patients with LBP of at least 4 weeks' duration. Also, earlier indications (6, 9) that patients with chronic LBP who are seen in primary care settings have a high education level, are employed and working, report use of non-narcotic analgesics, and usually have had no earlier back surgery, were supported by the results of this study.

Mean scores of RM have been reported earlier. Stratford et al. (26) reported a mean RM score of 11.8 (SD 6.2) in a sample of 88 patients with mechanical LBP

who were referred to the outpatient physical therapy department of a large teaching hospital. Mean duration of symptoms for these patients was 48 days (SD 36). Beurskens et al. (3) reported a baseline mean score of 12.1 (SD 4.7) in a sample of 38 patients with non-specific LBP who improved with treatment, and a baseline mean score of 11.8 (SD 5.1) in 38 patients who did not improve with treatment. Median duration of symptoms for these patients was 6 months. The average age of patients was 41 years in both samples. These RM mean scores are higher than in our sample. The fact that our sample included self-referred patients may be a factor of importance in accounting for this difference. Several of the hospital staff participated in a secondary prevention exercise program for LBP, which may have served as a trigger for seeking care even if their LBP did not warrant seeing a doctor. Generally there may be a higher awareness of health issues in this group so that care/preventive interventions were sought at a lower level of perceived disability; or it may be a selection bias. However, when the hospital staff as a group was excluded from the analysis, the mean RM-Sw score increased to 10.5 (SD 4.1), which is still lower than the scores reported by Stratford et al. (26) and Beurskens et al. (3).

Test-retest reliability for a one-week interval was very good, with an intraclass correlation of 0.88. This is in line with other reports of the reliability for RM. Grönblad et al. (13) reported an intraclass correlation of 0.83 for a one-week interval in a randomly drawn subset of subjects ($n = 20$) from a sample of patients with LBP for at least 3 months ($n = 94$). Kopec et al. (17) reported an intraclass correlation of 0.91 for an interval of "several days" in a sample of 242 patients seeking help for LBP.

The correlation between RM-Sw and the Pain Disability Index in our study ($r = 0.64$, $p < 0.001$) is similar to the correlation reported between these questionnaires ($r = 0.63$, $p < 0.001$) by Millard (19). The study sample consisted of 93 patients with chronic non-malignant LBP referred to two hospital-based pain management services.

Low to moderate correlation between functional status and pain intensity has been reported by several authors (e.g. 11, 18). The negative correlation of RM-Sw with the Life Control scale and the General Activity scale were low, but still in the expected direction. Thus, the results of our study support both convergent (Pain Disability Index, Interference scale, Pain Severity scale) and divergent (Life Control scale, General Activity scale) validity of RM-Sw.

The demographic and background characteristics of this sample reflect the characteristics of the population in a university town where public service and administration dominate. The education level is higher in our sample than in an LBP subgroup ($n = 242$: 54% nine-year school, 37% senior high school, 9% university level) of a random Swedish population sample (unpublished data). Further, 50% of the participants in our study reported having non-manual occupations, as compared to 41% in the general population (Statistics Sweden, 1990). However, the result of the standard multiple regression shows that gender, education and occupation only explained 14% of the variance in RM-Sw scores. Gender was the only variable that significantly contributed to the regression solution.

Some caution is warranted when interpreting these results, since the number of cases in the analysis ($n = 64$) is somewhat lower than the recommended cases to independent variable ratio (27). Thus, although the multiple regression results suggest that the demographic factors gender, education and occupation were only moderately related to RM-Sw scores, the present sample may not be representative for primary care settings in areas where, for example, heavy industry or farming dominate.

In conclusion, the findings concerning RM-Sw in the present study support results from earlier studies where RM has been established as a reliable and valid measure of functional ability in LBP. Longitudinal studies are needed to show if it can be used as a measure in outcome evaluations in primary care settings.

ACKNOWLEDGEMENTS

The authors wish to thank the staff at the physical therapy departments within the Uppsala county council primary health care organization, and at the Uppsala University hospital physical therapy department staff service for their kind cooperation in the data collection phase of this study.

REFERENCES

1. Baker, J., Pynsent, P. & Fairbank, J.: The Oswestry disability index revisited: its reliability, repeatability and validity, and a comparison with the St. Thomas's disability index. In: Backpain: New Approaches to Rehabilitation and Education (ed. M. Roland & J. Jenner), pp. 174–186. Manchester University Press, Manchester, 1989.
2. Bergner, M., Bobbit, R., Carter, W. & Gilson, B.: The Sickness Impact Profile: development and final revision of a health status measure. Med Care *19:* 787–805, 1981.
3. Beurskens, A., de Vet, H. & Köke, A.: Responsiveness of functional status in low back pain: a comparison of different instruments. Pain *65:* 71–76, 1996.

4. Beurskens, A., de Vet, H., Köke, A., van der Heijden, G. & Knipschild, P.: Measuring the functional status of patients with low back pain. Assessment of the quality of four disease-specific questionnaires. Spine *20:* 1017–1028, 1995.
5. Carlsson, S., Bergström, G. & Jensen, I. The Multidimensional Pain Inventory: a Swedish version. Department of Psychology, University of Gothenburg, 1994.
6. Crook, J., Weir, R. & Tunks, E.: An epidemiological follow-up survey of persistent pain sufferers in a group family practice and specialty pain clinic. Pain *36:* 49–61, 1989.
7. Deyo, R.: Measuring the functional status of patients with low back pain. Arch Phys Med Rehabil *69:* 1044–1053, 1988.
8. Deyo, R., Andersson, G., Bombardier, C., et al: Outcome measures for studying patients with low back pain. Spine *19:* 2032S–2036S, 1994.
9. Deyo, R., Bass, J., Walsh, N., Schoenfeld, L. & Ramamurthy, S.: Prognostic variability among chronic pain patients: implications for study design, interpretation, and reporting. Arch Phys Med Rehabil *69:* 174–178, 1988.
10. Deyo, R. & Centor, R.: Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. J Chron Dis *39:* 897–906, 1986.
11. Deyo, R. & Diehl, A.: Measuring physical and psychosocial function in patients with low back pain. Spine *8:* 635–642, 1983.
12. Fairbank, J.: The Oswestry low back pain disability questionnaire. Physiotherapy *66:* 271–273, 1980.
13. Grönblad, M., Hupli, M., Wennerstrand, P., et al: Intercorrelation and test-retest reliability of the Pain Disability Index (PDI) and the Oswestry Disability Questionnaire (ODQ) and their correlation with pain intensity in low back pain patients. Clin J Pain *9:* 189–195, 1993.
14. Jensen, M. & Karoly, P.: Self-report scales and procedures for assessing pain in adults. In: Handbook of Pain Assessment (ed. D. Turk & R. Melzack), pp. 135–151. The Guilford Press, New York/London, 1992.
15. Kerns, R., Turk, D. & Rudy, T.: The West Haven–Yale Multidimensional Pain Inventory (WHYMPI). Pain *23:* 345–356, 1985.
16. Kopec, J. & Esdaile, J.: Functional disability scales for low back pain. Spine *20:* 1943–1949, 1995.
17. Kopec, J., Esdaile, J., Abrahamowicz, M., et al: The Quebec back pain disability scale. Measurement properties. Spine *20:* 341–352, 1995.
18. Kröner-Herwig, B., Jäkle, C. & Frettlöh, J., et al: Predicting subjective disability in chronic pain patients. Int J Behav Med *3:* 30–41, 1996.
19. Millard, R.: Construct validity of practical questionnaires for assessing disability of low back pain. Spine *16:* 835–838, 1991.
20. Million, R., Hall, W., Nilsen, K., Baker, R. & Jayson, M.: Assessment of the progress of the back pain patient. Spine *7:* 204–212, 1982.
21. Nachemson, A.: Ont i ryggen—orsaker, diagnostik, behandling. Statens Beredning för Utvärdering av medicinsk metodik. Stockholm, 1991 (in Swedish).
22. Pollard, C.: Preliminary validity study of the Pain Disability Index. Percept Mot Skills *59:* 974, 1984.
23. 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 *8:* 141–144, 1983.
24. StatSoft STATISTICA™ for the Macintosh. StatSoft Inc., Tulsa, 1994.
25. Stratford, P., Binkley, J., Solomon, P., Finch, E., Gill, C. & Moreland, J.: Defining the minimum level of detectable change for the Roland-Morris questionnaire. Phys Ther *76:* 359–365, 1996.
26. Stratford, P., Binkley, J., Solomon, P., Gill, C. & Finch, E.: Assessing change over time in patients with low back pain. Phys Ther *74:* 528–533, 1994.
27. Tabachnic, B. & Fidell, L. Using multivariate statistics. Harper Collins, New York, 1996.
28. Tait, R., Chibnall, J. & Krause, S.: The Pain Disability Index: psychometric properties. Pain *40:* 171–182, 1990.
29. Tait, R., Pollard, C., Margolis, R., Duckro, P. & Krause, S.: The Pain Disability Index: psychometric and validity data. Arch Phys Med Rehabil *68:* 438–441, 1987.
30. Waddell, G. & Main, C.: Assessment of severity in low back disorders. Spine *9:* 204–208, 1984.

Accepted October 27, 1997

Address for offprints:

Eva Johansson, RPT
Dep. of Public Health and Caring Science
Uppsala Science Park
SE-751 83 Uppsala
Sweden