Turkish translation of the Roland-Morris Disability Questionnaire

This is a copy of a paper published as: Kucukdeveci AA, Tennant A, Elhan AH, Niyazoglu H. Validation of the Turkish version of the Roland-Morris Disability Questionnaire for use in low back pain. Spine. 2001 Dec 15;26(24):2738-43

Validation of the Turkish Version of the Roland-Morris Disability Questionnaire for Use in Low Back Pain

Ayse A. Küçükdeveci, MD, Associate Professor,
Alan Tennant*, PhD., Professor,
Atilla H. Elhan**, MSc,
Hava Niyazoglu, MD.

Department of Physical Medicine & Rehabilitation, Faculty of Medicine, University of Ankara, Turkey;

*Rheumatology & Rehabilitation Research Unit, School of Medicine, University of Leeds, United Kingdom;

**Department of Biostatistics, Faculty of Medicine, University of Ankara, Turkey.

This study was conducted in the Department of Physical Medicine & Rehabilitation, Faculty of Medicine, University of Ankara, Ankara, Turkey.

Address for correspondance:
Dr. Ayse A. Küçükdeveci
Gökkusagi Sitesi, Mor Blok, 12/25, Karakusunlar 06530 Ankara, Turkey
Tel : +90 312 2853936 Fax : +90 312 3094132
E-mail : ayse@tepa.com.tr
Acknowledgement: We gratefully acknowledge the financial support of the British Council for the collaboration of the authors.
Running Head: Turkish Version of Roland-Morris
Word Count: 3051

ABSTRACT

Study design: A reliability and validity study of a previously translated version of the Roland-Morris Disability Questionnaire (RMDQ).

Objectives: To validate the Turkish version of the RMDQ for use in low back pain.

Summary of Background data: Clinical and epidemiological research related to low back pain in the Turkish population would be facilitated by the availability of well-established outcome measures.

Methods: 81 outpatients with low back pain, 64 of whom were followed up on a second occasion were assessed by the RMDQ. Reliability is assessed using internal consistency and the intra-class correlation coefficient (ICC). Internal construct validity is assessed by Rasch analysis; external construct validity by association with pain and spinal movement. Responsiveness is tested by both the non-parametric and parametric effect size.

Results: Internal consistency of the RMDQ is found to be adequate (> 0.85) at both times, with high ICC's also at both time points. Internal construct validity of the scale is good, indicating a single underlying construct. Expected associations with pain confirm external construct validity. There is little evidence of Differential Item Functioning (DIF). The scale is at the ordinal level. Responsiveness of the RMDQ is good and greater than observed change in spinal movement.

Conclusions: The RMDQ is a robust unidimensional ordinal measure, largely free of DIF, which works well in the Turkish population. Non-parametric effect sizes of ordinal scales are found to over- or underestimate the true effect size depending upon the nature of the scale and the distribution of patients at baseline.

KEY WORDS
LOW BACK PAIN, DISABILITY, REHABILITATION, OUTCOME, RASCH.

KEY POINTS

- The Roland Morris Disability Scale is valid for the Turkish Population.
- The scale works in the same way by gender and age.
- The scale is ordinal.
- Effect size calculations on this scale are invalid due to its ordinal nature.

MINI ABSTRACT

This study investigates the reliability and validity of a previously translated Turkish version of the Roland-Morris disability questionnaire on 81 patients with low back pain. The questionnaire proved to be a robust, unidimensional ordinal scale, reliable and valid for the Turkish population.

INTRODUCTION

Low back pain is a ubiquitous health problem. It represents the most frequent illness of mankind after the common cold.[5,13] The lifetime prevalence of low back pain ranges from 60 to 90 percent and the annual incidence is 5 percent.[6] It is reported to be the leading cause of disability in people younger than 45 years of age and the third cause of disability in those older than 45 years of age.[28] High costs associated with low back pain and its socio-economic impact has made this so-called self-limited and benign condition a considerable health-care policy challenge, especially in industrial countries.[28]

These considerable consequences highlight the importance of measuring the outcome of health care interventions to alleviate low back pain. Measuring outcome in health care in general has been well understood among health care specialists in the last two decades. Potential applications of outcome measurement include the use in clinical practice for planning and monitoring therapy; clinical and epidemiological research; program evaluation, policy analysis to establish priorities and allocate resources; and population monitoring to track trends in levels of health, risk factors, and use of services.[12]

A variety of outcome measures are used for low back pain. Among the most common instruments developed to assess the functional status of patients with low back pain is the Roland-Morris Disability Questionnaire (RMDQ).[20] A self-completed measure of disability for use with people with low back pain, it was developed by selecting statements from the Sickness Impact Profile.[3] The phrase 'because of my back' was added to relevant questions in an attempt to elicit back pain-specific responses. Studies have reported that the test-retest reliability, validity and responsiveness of the RMDQ are adequate.[20,25]

The importance of measuring outcome has also been recognised among the rehabilitation medicine specialists in Turkey over the last decade. Consequently, internationally accepted instruments for functional assessment have been adapted and used, especially in clinical research.[14] One recent adaptation was for the RMDQ, which was subsequently used in a study to investigate the correlation of pain, spinal mobility and disability in chronic low back pain syndrome.[27] For the translation process (which was undertaken a decade ago), using the recent guidelines for cross cultural adaptation,[2] Stage I involved three bi-lingual professionals translating the original version. One professional had a clinical background and was thus an 'informed' translator. The other two translators were an English teacher and a bilingual secretary (educated in England), and were thus 'uniformed' translators. Inconsistencies in the translations were resolved (Stage II) by discussions between the translators. Back-translation (Stage III) and further expert review (Stage IV) was not undertaken at that time. Following pre-testing for face validity (Stage V) in the Turkish population, which includes variable educational levels, the general modification of 'because of my back *pain*' was made to adjust for nuances of the Turkish language. In addition, item 2 'I change position frequently to try and get my back comfortable' was modified to 'I change my standing, sitting and lying position frequently to make my back comfortable', to make it more readily understandable. However, the reliability and validity of this Turkish adaptation was not reported. The aim of this study is to rectify this shortfall and examine the reliability, validity and responsiveness of the RMDQ for Turkish patients with low back pain.

METHODS

Patients and Setting

Eighty-one consecutive outpatients with chronic low back pain of at least 3-months duration who were receiving one or more therapeutic interventions (non-steroid anti-inflammatory drug medication and/or physical therapy) were included in the study. All patients had been previously investigated by physical and neurological examination, spine x-rays and laboratory tests (complete blood count, erythrocyte sedimentation rate, blood biochemistry, urinary analysis) to identify the non-mechanical/medical causes of low back pain. Patients with the suspicion of non-mechanical/medical low back pain and patients having neurological deficit were not included in the study.

Methods

All patients, after giving their consent to participate, were assessed by the same observer (HN). The assessments included lumbar flexibility measured by Schober test, level of pain on a Likert pain scale (0 no pain, 1 mild, 2 moderate, 3 severe, 4 unbearable pain), and functional disability by RMDQ. RMDQ was self-completed by literate patients. Where patients were illiterate, the questionnaire was administered by the observer. Patients were asked to attend for a further assessment two weeks later.

Assessment of Reliability, Validity and Responsiveness

The scientific quality of an instrument is determined through a range of analysis. This includes tests for reliability, validity and, more recently, responsiveness. Reliability is concerned with the consistency of the instrument. Validity is concerned with whether the instrument measures the characteristic it purports to measure. Responsiveness assesses the ability of the instrument to detect change.

Two common forms of reliability for a self-completed questionnaire in a yes/no format, such as the present instrument, are internal consistency and the intra-class correlation coefficient, (ICC) both of which are evaluated in this study.[4,22] Increasingly, two forms of validity should also be considered; ‘internal’ validity where attention is given to the integrity of the defined construct,[18] and external validity which is concerned with expected associations with other key variables.[26] In the former, the data are tested against some model that examines unidimensionality. In the latter, where no ‘gold standard’ exists against which to contrast an instrument, construct validity becomes the usual form of external validity. Here the instrument is compared with other measures where there would be an expected level of agreement (convergent validity) or disagreement (divergent validity).[26]

Internal validity in the current study is assessed by fit of the data to the one-parameter Item Response Theory (Rasch) model.[19] The Rasch measurement model assumes that the data from an instrument are unidimensional and thus the model can be used to test whether the items in the scale do belong to a single underlying construct.[1,19] Testing the fit of the data to the Rasch model is equivalent to a test of the theoretical construct validity and adequacy of the scale.[10] Fit is assessed by two mean square (MNSQ) fit statistics. These statistics are derived for every item and, taken together, provide information on the consistency of the responses to each item. The outlier-sensitive MNSQ fit statistic (OUTFIT) is more sensitive to abnormal responses to items far from the person's ability level, for example, those from a very able person responding to a very easy item. This statistic is weighted to derive an information-weighted statistic (INFIT). As the influence of outliers is reduced, the INFIT is able to provide information about the more central responses, that is, individuals' responses to items at the same difficulty level as their ability level. MNSQ values between 0.6 and 1.4 are taken to reflect adequate fit to the model for this sample size.[23] Values above 1.4 indicate unexpected responses to the item, and may reflect poorly understood items, or those that do not belong to the same construct. Values below 0.6 indicate items where the response is more deterministic, in that there is less variation than expected. The data derived from the RMDQ were consequently fitted to the Rasch model, operationalized by the unconditional maximum likelihood approach.[31]

Another important aspect of the internal integrity of a scale is the absence of item bias or Differential Item Functioning (DIF).[8] At a given level of disability it is important that the response to any item is unaffected by group membership. For example, at the same level of disability it is important that both males and females have the same probability of affirming an item. If this probability differs, then the scale works in different ways by gender, rendering comparison between groups difficult. DIF should be evaluated for age and gender as a matter of routine, and other relevant groups as appropriate. In the current study, DIF is examined by age, gender, duration of low back pain, severity of pain, and time (baseline and follow-up). Items are considered to display DIF if there is a significant difference between groups in the residuals resulting from the Rasch analysis. This means that there should be no group-related patterns in the data once the primary ‘disability’ construct has been removed. Due to the number of tests undertaken, the level of significance is set at 0.01.

Once the internal validity of the scale has been confirmed, including the absence of DIF, external validity can be considered. In this study it has been examined by construct validity through convergent validity of the instrument with a measure of pain. Although impairments do not necessarily give rise to limitation in activities (disability), a moderate association (>.3) would be expected between pain and disability in the current clinical context.[15,27] Some association with the spinal mobility measure may also be expected, although previous work is equivocal as to the presence and extent of such an association.[16]

The scaling properties of the RMDQ are examined by comparing the average item difficulties and their position along the underlying disability construct, as defined by fit of the data to the Rasch model. An understanding of such properties is instructive for considering the limitations of effect size calculations. Finally, in this study, responsiveness is calculated both by the parametric[11] - and the non-parametric effect size.[30]

Data were analysed using the Statistical Package for the Social Sciences (SPSS),[17] and a Rasch-Model Computer program WINSTEPS.[32]

RESULTS

Patient characteristics

A total of 81 patients with a mean age of 37.0 years (SD:10.6), sixty three percent of whom are female, were enrolled in the study. Mean duration of low back pain was 4.6 years (SD 3.7), ranging from 4 months to 15 years. 64 patients attended for follow up assessment after two weeks. Mean spinal movement (Schober) was 14.2cm (SD 1.1) at baseline and 14.3cm (SD 0.9) at follow up. All patients report some pain at baseline, with almost over three in five (63%) reporting 'severe' or 'unbearable' levels of pain, reducing to 19% at follow up.

The baseline RMDQ was not normally distributed (Kolmogorov-Smirnov 0.114; df 81; p= 0.011). Baseline median RMDQ was 15.0 (IQR:8.0) and follow-up median RMDQ was 9 (IQR:9.8). Thus patients were experiencing quite high levels of disability at baseline and also (for the 64 patients who were followed up) a significant improvement over time (Wilcoxon Signed Ranks Test; Z=-5.919: p=<0.01).

Reliability

Internal consistency is adequate at both times, with Cronbach's alpha (α) at 0.85 and 0.89 at time 1 and time 2 respectively. ICC (one way random effect model) is also adequate at 0.79 and 0.86 at times 1 & 2 respectively.

Validity

a) Internal validity

b)

I. Unidimensionality

Fit of the RMDQ to the Rasch model shows adequate fit of items to a single underlying construct (Table 1). Item 10 'I only stand up for short periods of time because of my back' is most problematic. It has a high frequency of endorsement, but an OUTFIT of 1.76 indicates that there are some patients with severe disability who nevertheless say no to this item, when the opposite would be expected. Other than this item, most items display adequate fit to the model and this confirms an underlying single construct. Where misfit occurred, this tended to be low levels of fit – dependency, suggesting some redundancy in the scale. No patients confirmed item 19 'Because of my pain, I get dressed with help from someone else', at follow up, and thus the item was not estimated at that time point.

II. Differential Item Functioning (DIF)

There is no evidence of DIF for age, duration or severity of pain (all split at median). Two items display DIF for gender. For Item 5 'Because of my back, I use a handrail to get upstairs', at any given level of disability, females are more likely to affirm the item than males. The same applied to item 7 'Because of my back, I have to hold on to something to get out of an easy chair'. Finally, for time (baseline and follow-up) two other items showed a change in endorsement frequency. Both item 4 'Because of my back, I am not doing any of the jobs that I usually do around the house' and 11 'Because of my back, I try not to bend or kneel down' are less frequently affirmed at follow-up, given the same level of disability. However, given the number of groups tested, and the small number of items identified as problematic, the RMDQ appears robust with little evidence of serious DIF.

b) External validity

Convergent construct validity between the RMDQ and the Likert pain scale, which is demonstrated by a Kruskal-Wallis non parametric ANOVA, shows a significant association at both times ($p< .05$). The association between spinal mobility and the RMDQ is less clear, with a significant negative association at baseline (-0.337; $p=<.01$), but not at follow up (-0.109; $p=0.391$).

Scaling Properties

Figure 1 shows the way in which each item of the RMDQ marks the disability construct (x axis). Patients who have affirmed all the items will be at the right-hand side of this scale. As their disability reduces, they will move from right to left, reducing their score by one point each time they pass one of the marks (thresholds) on the graph. Consequently, severely disabled patients will have to make a substantial improvement in disability along the underlying metric construct before they reduce their score by five points (i.e. from 5.6 to 0.0 on the metric scale). In contrast, patients who begin with moderate disability at the mid-point of the scale (0.0 on the metric scale) will accumulate five raw score points within a 0.5 range of the metric scale. This is because the thresholds are clustered closely together, and some even duplicate the same level of disability (y axis is the number of thresholds at each point). It is clear from the analysis that the scale is working at the ordinal level. This has considerable importance not only for the use of appropriate statistics, but also for its potential impact on the calculation of effect size.

Responsiveness

Due to the non-normal distribution of the RMDQ at baseline the non-parametric version of the effect size is used. However, as the data fit the Rasch model it is also possible to use a transformation of the raw score to allow comparison between the parametric and non parametric approaches. Figure 2 shows the comparison between the raw change score and the Rasch transformed change measure for the 64 patients followed up. Each point is numbered by the patient identity number. It becomes immediately apparent that there are two groups of associations. One group of patients (the lower parallel set on the graph) show a much greater change in the transformed score than they do on the raw score, and these patients are characterised by a location at either end of the disability construct, with either high or low levels of disability. The second, much larger group of patients have a higher raw score change than a transformed change score. They are characterised by a location in the middle of the scale. Referring back to Figure 1 it becomes immediately obvious why this is the case. Patients moving across the margins of the scale accumulate fewer raw score points for a given change in metric disability and therefore their change is underestimated. In contrast, patients moving across the central part of the scale accumulate many points for the same change in metric disability, and their change is overestimated.

The effect of distribution of patients across the construct at baseline is shown in Table 2. The effect size for SCHOBER is negligible but that of the interval level Rasch transformed RMDQ (-1.28) is twice that of its equivalent non-parametric effect size (-0.533). When patients are grouped into those in i) the upper- ii) inter- and iii) lower-quartile groups on the RMDQ at baseline, non-parametric effect sizes differ considerably, reflecting the decreasing potential for reduction on the raw score count. With the transformed data, which takes account of the distortion in the underlying metric, these differences disappear, and effect sizes are similar for each group. The higher effect size reflects the reduction in standard deviation as a result of the grouping.

DISCUSSION

The sequential process of adaptation of outcome measures for use in different cultures is well documented.[9,2] Following this process, it is necessary to establish the psychometric credentials of the newly adapted instrument The adaptation of the RMDQ for the Turkish language has produced an instrument which is reliable, and demonstrates both internal and external validity. Levels of reliability were similar to those found elsewhere. For example, Cronbach's alpha was 0.81 for a German version,[29] and ranged between 0.84-0.91 at different time points for a Spanish version.[7] The latter reported an ICC of 0.874, also similar to the results presented above.

Internal validity, as defined by fit to the Rasch model, was good. There was some evidence of redundancy in the item set, suggesting a shorter scale may be feasible. Indeed, there is an 18-item version of the scale available in English,[24] and further work could be undertaken to match the redundant items in the existing scale with this shorter version.

External validity was also demonstrated. Significant increases in RMDQ scores were found for increasing levels of pain as defined by the Likert pain scale, again similar to that found in the original validity study for the RMDQ.[20]

The adapted version showed adequate responsiveness. However, it has become clear during this analysis that there are major problems with the established procedures for the calculation of effect size. Change scores, or median change scores do not take account of the true nature of ordinal data and seriously underestimate the contribution of change at the margins. This suggests that most published data on effect sizes may be corrupt and invalid. However, it is possible that the distortion of the underlying metric may vary from scale to scale and thus it becomes an empirical question to determine how much the effect size calculation is compromised on each scale.

In conclusion, the RMDQ is a robust, unidimensional ordinal measure, largely free of DIF, which works well in the Turkish population. The full adapted version is published in the Appendix. This version, including details of the translation process have been sent to Dr Roland for inclusion in the Spine Website.[21] Future work on comparative hierarchical ordering of items across cultures would inform on the extent of cross-cultural validity of this versatile scale.

REFERENCES

1. Andrich D. Rasch models for measurement. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 68. Beverly Hills: Sage Publications, 1988.
2. Beaton D, Bombardier C, Guillemin, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000; 25: 3186-3191.
3. Bergner M, Bobbitt RA, Kressel A, Pollard WE, Gilson BS, Morris JR. The Sickness Impact Profile: conceptual formulation and methodology for the development of a health status measure. Int J Health Services 1976; 6: 393-415.
4. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951; 16: 297-334.
5. Frymoyer JW, Pope MH, Costanza MC, Gogging JE, Wilder DG. Epidemiological studies of low back pain. Spine 1980; 5: 19-23.
6. Frymoyer JW. Back Pain and sciatica. N Engl J Med 1988; 318: 291-300.
7. Gil del Real MT, Lioberai J, Kovacs FM, et.al., Validation of thee Spanish version of the Roland Morris questionnaire for quality of life in patients with low back pain. Quality of Life Research 1999; 8: 612.
8. Groenvold M, Bjorner JB, Klee MC, Kreiner S. Test for item bias in a quality of life questionnaire. J Clin Epidemiol 1995; 48: 805-816.
9. Guillemin EG, 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.
10. Gustafsson J-E: Testing the fit of data to the Rasch model. Br J Math Stat Psychol 1980; 33: 205-33.
11. Kaziz L, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989; 27: S178-S189.
12. Keith RA. Functional status and health status. Arch Phys Med Rehabil 1994; 75: 478-83.
13. Kelsey JL, White AA. Epidemiology and impact of low back pain. Spine 1980; 5: 133-42.
14. Küçükdeveci AA, Yavuzer G, Tennant A, Süldür N, Sonel B, Arasıl T. Adaptation of the modified Barthel index for use in physical medicine and rehabilitation in Turkey. Scand J Rehabil Med 2000; 32: 87-92.
15. Lankhorst GJ, Van der Stadt RJ, Van der Korst JK. The natural history of idiopathic low back pain. Scan J Rehab Med 1985; 17: 1-4.
16. Million R, Hall W, Haavik-Nilsen K. Assessment of the progress of the back pain patient. Spine 1982; 7: 204-212.
17. Norusis MJ. SPSS for Windows. Advanced Statistics. Release 6.0. Chicago: SPSS inc, 1993.
18. Nunally JC Jr. Psychometric theory. (2^{nd} edn). New York, McGraw-Hill, 1978.
19. Rasch G. Probabilistic Models for some intelligence and attainment tests. Chicago: University of Chicago Press, 1980.
20. Roland M, Morris M. A study of the natural history of back pain. Part 1: Development of a reliable and sensitive measure of disability in low-back pain. Spine 1983; 8: 141-4.
21. Roland M, Fairbank J. The Roland-Morris Questionnaire and the Oswestry Disability Questionnaire. Spine 2000; 25: 3115-3124.

22. Shrout PE Fleiss JL. Intraclass correlation coefficient: uses in assessing rater reliability. Psychol Bull 1979; 86: 420-428.
23. Smith RM, Schumacker RE, Bush MJ. Using item mean squares to evaluate fit to the Rasch model. J Outcome Measurement 1998; 2: 66-78.
24. Stratford PW, Binkley JM. Measurement properties of the RM-18. A modified version of the Roland-Morris disability scale. Spine 1997; 22: 2416-2421.
25. Stratford PW, Binkley JM, Riddle DL, Guyatt GH. Sensitivity to change of the Roland-Morris back pain questionnaire: Part 1. Phys Ther 1998; 78: 1186-96.
26. Streiner DL, Norman GR. Health Measurement Scales. Oxford: Oxford University Press, 1995.
27. Tuncer S, Arasil T, Alpar R. The correlation of pain, spinal mobility and disability in chronic low back pain syndrome. J Rheum Med Rehab 1992; 3: 2-9.
28. Weinstein SM, Herring SA. Rehabilitation of the patient with low back pain. In: DeLisa JA, Gans BM (eds). Rehabilitation Medicine: Principles and Practice. Lippincott Company, Philadelphia, 1993; p: 996-1017.
29. Wiesinger GF, Nuhr M, Quittan M, Ebenbichler G, Wöhlfl 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.
30. Wilcox,RR; Muska,J Measuring effect size: A non-parametric analogue of w2. Brit. J. Math Stat Psychol 1999; 52: 93-110.
31. Wright BD, Stone MH. Best Test Design. Chicago: Mesa Press, 1979.
32. Wright BD, Linacre JM: A user's guide to WINSTEPS. Chicago: Mesa Press, 1998.

Appendix: Turkish Version of the RMDQ

Note: This version may be used without permission of the original authors or from the journal *Spine*.

1. Bel ağrım yüzünden zamanımın büyük çoğunluğunu evde geçiriyorum.
2. Belimi rahatlatmak için sık sık ayakta duruş, oturuş veya yatış şeklimi değiştirmek zorunda kalıyorum.
3. Bel ağrım yüzünden eskisinden daha yavaş yürüyorum.
4. Bel ağrım yüzünden evde yaptığım birçok işi artık yapmıyorum.
5. Bel ağrım yüzünden merdivenleri çıkarken trabzanlara tutunuyorum.
6. Bel ağrım yüzünden dinlenmek için sık sık uzanıyorum.
7. Bel ağrım yüzünden sandalyeden kalkarken bir yere tutunmak ihtiyacı duyuyorum.
8. Bel ağrım yüzünden bazı işlerimi başkalarına yaptırıyorum.
9. Bel ağrım yüzünden eskisinden daha yavaş giyiniyorum.
10. Bel ağrım yüzünden sadece kısa süre ayakta kalabiliyorum.
11. Bel ağrım yüzünden eğilmekten ve çömelmekten kaçınıyorum.
12. Bel ağrım yüzünden sandalyeden kalkarken zorluk çekiyorum.
13. Belim hemen hemen her zaman ağrıyor.
14. Bel ağrım yüzünden yatakta dönmekte güçlük çekiyorum.
15. Bel ağrım yüzünden iştahım azaldı.
16. Bel ağrım yüzünden çoraplarımı giymekte zorluk çekiyorum.
17. Bel ağrım yüzünden sadece kısa mesafeleri yürüyebiliyorum.
18. Bel ağrım yüzünden rahat uyuyamıyorum.
19. Bel ağrım yüzünden bir başkasının yardımıyla giyiniyorum.
20. Bel ağrım yüzünden günün büyük bir kısmını oturarak geçiriyorum.
21. Bel ağrım yüzünden evdeki ağır işleri yapmaktan kaçınıyorum.

22. Bel ağrım yüzünden eskisine göre huzursuz ve sinirliyim.

23. Bel ağrım yüzünden merdivenleri her zamankinden daha yavaş çıkıyorum.

24. Bel ağrım yüzünden zamanın çoğunu yatakta geçiriyorum.

Table 1. Fit of items of the RMDQ to the Rasch model.

	Time1			Time2		
Item Number	Item calibration (error)	INFIT MNSQ	OUTFIT MNSQ	Item calibration (error)	INFIT MNSQ	OUTFIT MNSQ
10	-3.15 (.49)	1.28	**1.76**	-1.50 (.33)	1.13	0.98
2	-2.93 (.46)	0.98	0.72	-1.72 (.33)	1.33	1.30
21	-2.26 (.37)	0.88	1.04	-2.85 (.39)	0.87	0.93
11	-1.77 (.33)	0.88	0.63	-3.35 (.43)	1.06	1.30
6	-1.56 (.32)	1.11	0.91	-1.39 (.32)	1.08	1.04
4	-1.56 (.32)	0.86	0.63	0.03 (.32)	0.79	0.68
23	-0.84 (.28)	0.89	0.69	-1.50 (.33)	0.85	0.69
17	-0.69 (.28)	1.02	1.02	-0.18 (.32)	0.92	0.82
3	-0.61 (.28)	1.03	0.95	0.34 (.33)	0.74	**0.54**
14	-0.16 (.27)	0.87	0.75	0.57 (.34)	0.79	**0.55**
22	-0.09 (.27)	1.17	1.56	-0.38 (.32)	1.13	1.45
8	-0.09 (.27)	1.09	1.24	0.46 (.34)	0.98	1.15
5	-0.02 (.27)	0.84	0.74	0.24 (.33)	1.16	1.23
7	0.12 (.27)	0.61	**0.50**	0.03 (.32)	1.01	1.28
1	0.19 (.26)	1.01	0.89	0.57 (.34)	0.96	0.74
18	0.26 (.26)	1.25	**1.49**	-0.18 (.32)	1.00	0.97
20	0.26 (.26)	1.28	1.37	0.34 (.33)	1.00	0.95
16	0.33 (.26)	1.01	1.00	0.57 (.34)	1.18	1.41
12	0.33 (.26)	0.9	0.79	1.07 (.37)	0.77	**0.54**
9	1.34 (.28)	0.8	0.62	1.67 (.42)	0.94	0.73
13	1.74 (.29)	1.27	1.65	1.51 (.40)	1.31	1.02
24	2.01 (.30)	1.11	0.98	1.20 (.38)	1.04	0.80
15	3.19 (.40)	0.87	**0.46**	4.47 (1.05)	1.15	1.05
19	5.96 (1.06)	1.16	**0.29**			

Bold numbers indicate miss-fitting items

Table 2. Comparison of parametric and non-parametric Effect Sizes.

	Parametric*	Non-Parametric
SCHOBER	0.1168	0.0000
RMDQ-all	-1.2830	-0.5333
RMDQ upper-quartile	-3.1273	-4.1481
RMDQ inter-quartile	-3.7486	-2.3333
RMDQ lower-quartile	-2.1075	-0.4444

*RMDQ is Rasch transformed

Figure 1. The measurement imprint of the RMDQ.

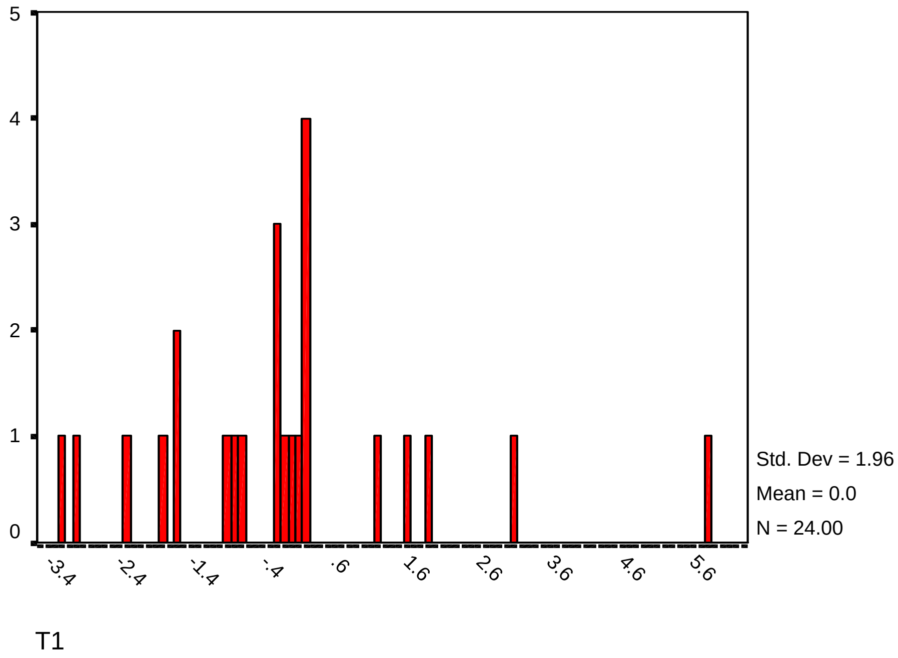

Figure 2. Comparison of raw change score (TOTAL_CH) and Rasch-transformed change score (CHANGE1) of RMDQ.