

How little pain and disability do patients with low back pain have to experience to feel that they have recovered?

Steven J. Kamper · Christopher G. Maher ·
Robert D. Herbert · Mark J. Hancock ·
Julia M. Hush · Robert J. Smeets

Received: 26 October 2009 / Revised: 5 January 2010 / Accepted: 25 February 2010 / Published online: 13 March 2010
© Springer-Verlag 2010

Abstract Epidemiological and clinical studies of people with low back pain (LBP) commonly measure the incidence of recovery. The pain numerical rating scale (NRS), scores from 0 to 10, and Roland Morris disability questionnaire (RMDQ), scores from 0 to 24, are two instruments often used to define recovery. On both scales higher scores indicate greater severity. There is no consensus, however, on the cutoff scores on these scales that classify people as having recovered. The aim of this study was to determine which cutoff scores most accurately classify those who had recovered from LBP. Subjects from four clinical studies were categorized as ‘recovered’ or ‘unrecovered’ according to their self-rating on a global perceived effect scale. Odd ratios were calculated for scores of 0, 1, 2, 3 and 4 on the NRS and RMDQ to predict perceived recovery. Scores of 0 on the NRS and ≤ 2 on the RMDQ most accurately identify patients who consider themselves completely recovered. The diagnostic odds ratio (OR) for predicting recovery was 43.9 for a score of 0 on the NRS and 17.6 for a score of ≤ 2 on the

RMDQ. There was no apparent effect of LBP duration or length of follow-up period on the optimal cutoff score. OR for the NRS were generally higher than those for RMDQ. Cutoffs of 0 on the NRS and 2 on the RMDQ most accurately classify subjects as recovered from LBP. Subjects consider pain more than disability when determining their recovery status.

Keywords Low back pain · Recovery · Pain numerical rating scale · Roland Morris disability questionnaire

Introduction

Studies on low back pain (LBP) commonly measure the incidence of recovery to make determinations of the effect of treatment [12, 16, 31, 33] or prognosis [11, 20, 25, 28]. However, recovery is a complex construct and there is currently no consensus on how it should be defined or measured. This makes interpretation and comparison of studies problematic.

Two measures that are frequently used and have been endorsed as measures of outcome for LBP studies [9] are the 11-point numerical rating scale for pain [7] (NRS) and the Roland Morris disability questionnaire [30] (RMDQ). These scales can be used to track the course of pain and disability. Importantly, NRS and RMDQ scores are also used to classify people as having recovered or not from an episode of back pain [2, 3, 8, 14–17].

If NRS or RMDQ are to be used to classify people as recovered or not, it is necessary for researchers to set a cutoff value, below which people are considered recovered and above which people are considered not recovered. Currently, this choice of cutoff is made arbitrarily. Thus,

S. J. Kamper (✉) · C. G. Maher · R. D. Herbert
The George Institute for International Health,
University of Sydney, PO Box M201, Missenden Rd,
Camperdown, NSW 2050, Australia
e-mail: skamper@george.org.au

M. J. Hancock · J. M. Hush
Faculty of Health Sciences, University of Sydney,
Sydney, NSW, Australia

R. J. Smeets
Rehabilitation Foundation Limburg, Department
of Rehabilitation Medicine, Caphri, Maastricht University,
Maastricht, The Netherlands

while some studies define recovery as having NRS scores of 0 [23], others define recovery as having NRS scores of ≤ 1 [18], and still others define recovery as having NRS scores of ≤ 2 [5]. The RMDQ score has also been used to define recovery, but similar inconsistencies apply. Cutoff scores used for the RMDQ have included ≤ 2 [2], ≤ 3 [23] and ≤ 4 [15]. At this point, there are no studies that detail the relationship of LBP patients' perceptions of their own recovery with cutoff scores on these commonly used instruments and, hence, no data to guide selection of the most suitable cutoff.

Recovery may also be assessed more directly by asking patients whether they perceive they have recovered. The response options may be binary [19], or on a Likert scale [21], or a variant of the global perceived effect scale [24] (GPE). Studies that concurrently consider patient self-report measures of recovery and measure NRS or RMDQ, therefore, offer an opportunity to consider what score on the NRS or RMDQ best equates to patients' perceptions of recovery. That is the approach we have used in the present study.

Patients with LBP are commonly categorized on the basis of how long they have experienced their symptoms. The European Guidelines for management of LBP [32] suggest classification as follows: acute (<6 weeks of symptom duration), subacute (6–12 weeks) and chronic (>12 weeks). It is not known if the duration of the episode of LBP influences patients' ratings of recovery. Such information could provide guidance as to whether measurement of recovery should be adapted to suit different situations.

Recovery is a complex construct [1, 22, 27] and it is possible that construct-specific instruments such as the NRS and RMDQ may capture different aspects of recovery to direct reports of perceived recovery. Both construct-specific instruments and direct reports offer advantages and disadvantages. While the latter enables the individual to incorporate whatever constructs they consider important into their rating of recovery, such ratings are inherently idiosyncratic, which could make between-subject comparisons difficult. On the other hand, defining recovery in terms of pain and disability may limit individual relevance, but facilitate interpretation.

This study involved a secondary analysis of data from four clinical studies of people with LBP. In each study, subjects reported NRS and RMDQ scores and directly reported whether they perceived they had recovered via a GPE scale. The primary aim was to determine the cutoffs for NRS and RMDQ that most accurately identified when LBP patients consider themselves recovered. A secondary aim was to assess whether the optimal cutoff varied according to the duration of LBP at entry into the study or to the length of follow-up period within the study.

Methods

Data sets

All subjects had non-specific LBP and were enrolled in one of four clinical studies. One was an observational study [4] in which patients were followed for 2 weeks while receiving a course of physiotherapy treatment. The other three studies were RCTs investigating the effects of McKenzie exercises [26], exercise and advice [29] or mobilization and stabilization exercises [10]. Symptom duration and follow-up points varied across the studies (Table 1).

We grouped studies according to the duration of participants' LBP [32] at entry into the study: acute (<6 weeks of symptom duration), subacute (6–12 weeks) and chronic (>12 weeks). Follow-up times were defined as the time from the entry into the study or randomization until follow-up and grouped as follows: short term (<12 weeks), intermediate term (12–26 weeks) and long term (>26 weeks).

Measures

All subjects completed (a) an 11-point NRS to score their pain, either averaged over the previous 24 h or over the previous week depending on the study, (b) the 24-point RMDQ to rate their disability, and (c) a GPE scale to rate recovery at each follow-up point. A full description of the measures is presented in "Appendix". GPE was scored on an 11-point scale from completely recovered (+5), through unchanged (0) to vastly worse (−5). Scores were dichotomized such that subjects who reported that they were completely recovered (+5) were classified as recovered and all others as unrecovered. It is acknowledged that selection of a recovery cutoff score of completely recovered (+5) on the GPE may be considered an arbitrary choice on behalf of the authors, but there is at least face validity for the assumption that subjects selecting the option 'completely recovered' perceived themselves to be recovered.

Analyses

We examined the performance of cutoffs of ≤ 0 , 1, 2 and 3 for NRS scores and ≤ 0 , 1, 2, 3 and 4 for RMDQ scores.

Data from all studies at all follow-ups were pooled. Subjects were categorized according to whether their NRS and RMDQ scores were \leq or $>$ each candidate cutoff score. Diagnostic odds ratios (OR) were calculated for each candidate cutoff score. OR is the odds of being recovered for someone who scores on or below the cutoff divided by the odds of being recovered for someone who scores above the cutoff.

Table 1 Included studies

	Sample size	LBP duration	Follow-up points		
			Short	Intermediate	Long
Costa et al. [4]	99	Acute	1 day, 3 weeks	–	–
Machado et al. [26]	140	Acute	1 and 3 weeks	–	–
Pengel et al. [29]	230	Subacute	6 weeks	12 weeks	52 weeks
Ferreira et al. [10]	220	Chronic	8 weeks	26 weeks	52 weeks

To test whether the optimal cutoff score varied with duration of LBP or length of follow-up, OR were calculated for each category of LBP (acute, subacute, chronic) and for each follow-up period (short, intermediate and long term).

Results

Optimal cutoff scores

Data from all studies at all follow-up points were pooled. In 1,818 assessments, 7.98% patients assigned a GPE score of completely recovered.

The cutoff scores that best discriminated people who felt they were, or were not recovered were an NRS score of 0 and an RMDQ score of 2 (Table 2). For the NRS, the OR for the optimal NRS cutoff was substantially higher than for the other candidates' cutoffs. That is, a cutoff of 0 was clearly superior to other cutoff scores for the NRS. In contrast for RMDQ, the ORs were similar for all candidate cutoffs, indicating that no cutoff was clearly superior to others for the RMDQ. ORs for pain scores were substantially higher than those for RMDQ.

To compare the relative effectiveness of NRS and RMDQ in classifying recovery, we created ROC [6] curves

Table 2 Odds ratios of complete recovery based on cutoff score

	Cutoff	OR
Numerical rating scale	0	43.9
	≤1	32.8
	≤2	20.9
	≤3	20.1
Roland Morris disability questionnaire	0	14.5
	≤1	15.5
	≤2	17.6
	≤3	15.5
	≤4	15.2

Cutoff scores that best discriminate between recovered and unrecovered participants are indicated in bold

OR Odds ratio

(Fig. 1). Data were pooled from all assessments and areas under the curves [6] and (AUC) were 0.897 for the NRS and 0.856 for the RMDQ.

Effect of LBP duration on cutoff

Separate analyses were conducted on subjects with acute LBP (476 assessments), subacute LBP (697 assessments) and chronic LBP (645 assessments). An NRS score of 0 had the highest OR for acute and chronic populations, while a score of 1 had the highest OR in subacute populations. RMDQ scores of 0, 1 and 2 had the highest ORs for subacute, acute and chronic populations, respectively (Table 3). There was no clear pattern of higher optimal cutoff scores with increasing LBP duration for either instrument. As with the data from the whole cohort, optimal cutoffs for NRS have much larger ORs than other candidates' cutoffs, whereas ORs for all candidates are much more similar for RMDQ. Optimal pain cutoffs had higher absolute ORs for predicting recovery than optimal RMDQ cutoffs.

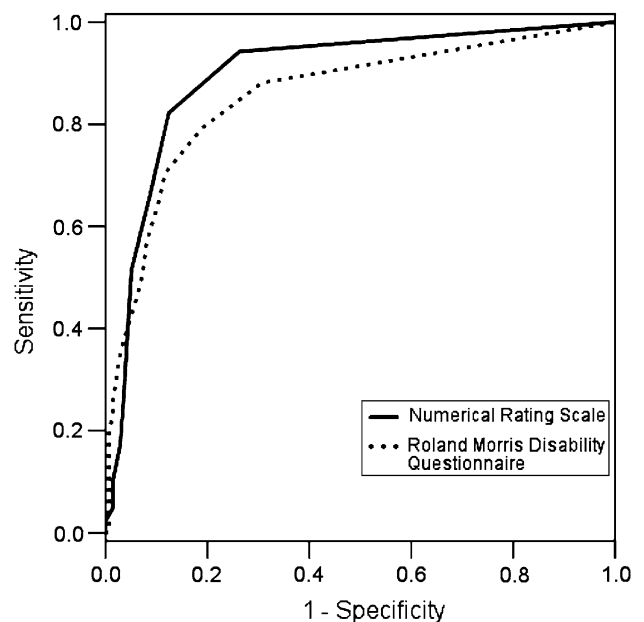


Fig. 1 ROC curves for numerical rating scale and Roland Morris disability questionnaire classifying recovery

Table 3 Odds ratios of complete recovery based on cutoff score and LBP duration

	Cutoff	Acute	Subacute	Chronic
Numerical rating scale	0	42.5	47.9	23.5
	≤1	17.1	62.2	17.8
	≤2	12.8	31.8	11.8
	≤3	15.1	49.1	6.3
Roland Morris disability questionnaire	0	21.0	11.4	7.2
	≤1	23.8	10.5	11.0
	≤2	23.1	11.0	14.7
	≤3	16.4	11.2	11.2
	≤4	19.9	10.4	8.4
Percentage recovered (%)		10.11	11.91	2.16

Cutoff scores that best discriminate between recovered and unrecovered participants are indicated in bold

OR Odds ratio

Effect of length of follow-up on cutoff

Candidate cutoffs were tested separately for all subjects at short-term follow-up (929 assessments), intermediate-term follow-up (447 assessments) and long-term follow-up (442 assessments). NRS scores of 0 or 1 and RMDQ scores of 0 and 4 had the highest ORs. There was no clear pattern of higher optimal cutoff scores with increasing length of follow-up for either instrument (Table 4). As with the data from the whole cohort, optimal cutoffs for NRS have much larger ORs than other candidates, whereas ORs for all candidates are much more similar for RMDQ. Optimal pain cutoffs had higher absolute ORs for predicting recovery than optimal RMDQ cutoffs.

Discussion

This study provides the first data-based recommendations for cutoffs to define recovery from LBP using NRS and

RMDQ scores. Patients who consider themselves completely recovered are most accurately classified by a score of 0 on a pain NRS or a score of 2 or less on the 24-point RMDQ. These cutoffs provide a high level of discrimination between people who consider themselves completely recovered and those who do not. While we do not suggest that pain or disability measures in isolation capture the full complexity of recovery, the AUCs from the ROC curves provide some assurance that NRS and RMDQ are useful in defining recovered subjects.

We expected that subjects with longer LBP duration would consider themselves recovered despite higher levels of pain and disability (that is, we expected they would have higher cutoff scores) but this was not the case. There was no clear evidence that the optimal cutoff score was higher with increasing LBP duration or with increasing length of follow-up. While the optimal cutoff for RMDQ was higher (≤4) for long-term follow-up than for the complete data set, it was also higher (4) for the short-term follow-up, and lower (0) for the intermediate term. This apparently random finding may reflect statistical imprecision associated with the low number of recovered subjects.

Odds ratios for optimal NRS cutoffs were always higher than those for the RMDQ. We did not use inferential statistics to investigate the significance of these differences, because observations were clustered by study and there was more than one observation on some subjects and so we could not assume independence of observations. To some extent, this difference may be due to the different number of response options for the 2 scales, i.e., 11 in the case of the NRS and 25 for the RMDQ. Nonetheless, it appears that NRS pain scores align better with patients' perceptions of recovery than do RMDQ scores. This contention is supported by the AUCs for the ROC curves, which were higher for the NRS than the RMDQ.

Various statistics can be used to quantify test performance. These include accuracy, sensitivity and specificity, positive and negative predictive values, likelihood ratios

Table 4 Odds ratios of complete recovery based on cutoff score and follow-up period

	Cutoff	Short term	Intermediate term	Long term
Numerical rating scale	0	46.2	25.9	73.7
	≤1	25.3	27.5	86.5
	≤2	18.5	16.8	35.5
	≤3	17.2	14.7	42.6
Roland Morris disability questionnaire	0	18.0	13.6	13.0
	≤1	14.6	11.7	28.0
	≤2	18.7	9.6	35.1
	≤3	18.5	7.3	26.7
	≤4	19.2	5.7	44.3
Percentage recovered (%)		8.18	6.70	8.82

Cutoff scores that best discriminate between recovered and unrecovered participants are indicated in bold

OR Odds ratio

- 1. I stay at home most of the time because of my back.
- 2. I change positions frequently to try and get my back comfortable.
- 3. I walk more slowly than usual because of my back.
- 4. Because of my back, I am not doing any of the jobs that I usually do around the house.
- 5. Because of my back, I use a handrail to get upstairs.
- 6. Because of my back, I lie down to rest more often.
- 7. Because of my back, I have to hold on to something to get out of an easy chair.
- 8. Because of my back, I try to get other people to do things for me.
- 9. I get dressed more slowly than usual because of my back.
- 10. I only stand up for short periods of time because of my back.
- 11. Because of my back, I try not to bend or kneel down.
- 12. I find it difficult to get out of a chair because of my back.
- 13. My back is painful almost all the time.
- 14. I find it difficult to turn over in bed because of my back.
- 15. My appetite is not very good because of my back pain.
- 16. I have trouble putting on my socks (or stockings) because of the pain in my back.
- 17. I only walk short distances because of my back pain.
- 18. I sleep less well because of my back.
- 19. Because of my back pain, I get dressed with help from someone else.
- 20. I sit down for most of the day because of my back.
- 21. I avoid heavy jobs around the house because of my back.
- 22. Because of my back pain, I am more irritable and bad tempered with people than usual.
- 23. Because of my back, I go upstairs more slowly than usual.
- 24. I stay in bed most of the time because of my back.

References

1. Beaton DE, Tarasuk V, Katz JN, Wright JG, Bombadier C (2001) Are you better? A qualitative study of the meaning of recovery. *Arthritis Care Res* 45:270–279
2. Burton AK, McClune TD, Clarke RD, Main CJ (2004) Long-term follow-up of patients with low back pain attending for manipulative care: outcomes and predictors. *Man Ther* 9:30–35
3. Collins DL, Evans JM, Grundy RH (2006) The efficiency of multiple impulse therapy for musculoskeletal complaints. *J Manip Physiol Ther* 29:162.e1–162.e9
4. Costa LOP, Maher CG, Latimer J, Ferreira PH, Ferreira ML, Pozzi GC, Freitas LMA (2008) Clinimetric testing of three self-report outcome measures for low back pain patients in Brazil. Which one is the best? *Spine* 33:2459–2463
5. Coste J, Lefranc G, Guillemin F, Pouchot J (2004) Prognosis and quality of life in patients with acute low back pain: insights from a comprehensive inception cohort study. *Arthritis Rheum* 51:168–176
6. Deyo R, Cedor R (1986) Assessing responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. *J Chronic Dis* 39:897–906
7. Downie WW, Leatham PA, Rhind VM, Wright V, Branco JA, Andersen JA (1978) Studies with pain rating scales. *Ann Rheum Dis* 37:378–381
8. Dunn KM, Jordan K, Croft PR (2006) Characterizing the course of low back pain: a latent class analysis. *Am J Epidemiol* 163:754–761
9. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N, Carr DB, Chandler J, Cowan P, Dionne R, Galer BS, Hertz S, Jadad A, Kramer LD, Manning DC, Martin S, McCormick CG,

- McDermott MP, McGrath P, Quessy S, Rappaport BA, Robbins W, Robinson JP, Rothman M, Royal MA, Simon L, Stauffer JW, Stein W, Tolett J, Wernicke J, Witter J (2005) Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 113:9–19
10. Ferreira ML, Ferreira PH, Latimer J, Herbert RD, Hodges PW, Jennings MD, Maher CG, Refshauge KM (2007) Comparison of general exercise, motor control exercise and spinal manipulative therapy for chronic low back pain: a randomized trial. *Pain* 131:31–37
 11. Gheldof ELM, Vinck J, Vlaeyen JWS, Hidding A, Crombez G (2007) Development of and recovery from short- and long-term low back pain in occupational settings: a prospective cohort study. *Eur J Pain* 11:841–854
 12. Giles LGF, Muller R (2003) Chronic spinal pain: a randomized clinical trial comparing medication, acupuncture, and spinal manipulation. *Spine* 28:1490–1502
 13. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PMM (2003) The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol* 56:1129–1135
 14. Grotle M, Brox JI, Vøllestad NK (2005) Concurrent comparison of responsiveness in pain and functional status measurements used for patients with low back pain. *Spine* 29:E492–E501
 15. Grotle M, Brox JI, Glomsrød B, Lønnc JH, Vøllestad NK (2007) Prognostic factors in first-time care seekers due to acute low back pain. *Eur J Pain* 11:290–298
 16. Hancock MJ, Maher CG, Latimer J, McLachlan AJ, Cooper CW, Day RO, Spindler MF, McAuley JH (2007) Assessment of diclofenac or spinal manipulative therapy, or both, in addition to recommended first-line treatment for acute low back pain: a randomised controlled trial. *Lancet* 370:1638–1643
 17. Hancock MJ, Maher CG, Latimer J, Herbert RD, McAuley JH (2008) Independent evaluation of a clinical prediction rule for spinal manipulative therapy: a randomised controlled trial. *Eur Spine J* 17:936–943
 18. Hancock MJ, Maher CG, Latimer J, Herbert RD, McAuley JH (2009) Can rate of recovery be predicted in patients with acute low back pain? Development of a clinical prediction rule. *Eur J Pain* 13:51–55
 19. Heneweer H, Aufdemkampe G, Van Tulder MW, Kiers H, Stappaerts KH, Vanhees L (2007) Psychosocial variables in patients with (sub)acute low back pain: an inception cohort in primary care physical therapy in the Netherlands. *Spine* 32:586–592
 20. Henschke N, Maher CG, Refshauge KM, Herbert RD, Cumming RG, Bleasel J, York J, Das A, McAuley JH (2008) Prognosis in patients with recent onset low back pain in Australian primary care: inception cohort study. *Br Med J* 337:154–157
 21. Heymans MW, De Vet HCW, Bongers PM, Knol DL, Koes BW, Van Mechelen W (2006) The effectiveness of high-intensity versus low-intensity back schools in an occupational setting: a pragmatic randomized controlled trial. *Spine* 31:1075–1082
 22. Hush JM, Refshauge KM, Sullivan G, de Souza L, Maher CG, McAuley JH (2009) Recovery: what does this mean to patients with low back pain? *Arthritis Rheum* 61(1):124–131
 23. Jensen TS, Albert HB, Sorensen JS, Manniche C, Leboeuf-Yde C (2007) Magnetic resonance imaging findings as predictors of clinical outcome in patients with sciatica receiving active conservative treatment. *J Manip Physiol Ther* 30:98–108
 24. Kamper SJ, Maher CG, Mackay G (2009) Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther* 17:163–170
 25. Linton SJ, Boersma K (2003) Early identification of patients at risk of developing a persistent back problem: the predictive validity of the Orebro musculoskeletal pain questionnaire. *Clin J Pain* 19:80–86
 26. Machado LAC, Maher CG, Herbert RD, Clare H, McAuley JH (2010) The effectiveness of the McKenzie method in addition to first-line care for acute low back pain: a randomized controlled trial. *BMC Med* 8(10). [Epub ahead of print]
 27. Norman GR (2003) Hi! How are you? Response shift, implicit theories and differing epistemologies. *Qual Life Res* 12:239–249
 28. Oleske DM, Neelakantan J, Andersson GB, Hinrichs BG, Lavender SA, Morrissey MJ, Zold-Kilbourn P, Taylor E (2004) Factors affecting recovery from work-related, low back disorders in autoworkers. *Arch Phys Med Rehabil* 85:1362–1364
 29. Pengel LHM, Refshauge KM, Maher CG, Nicholas MK, Herbert RD, McNair P (2007) Physiotherapist-directed exercise, advice, or both for subacute low back pain: a randomized trial. *Ann Intern Med* 146:787–796
 30. Roland M, Morris RA (1983) A study of the natural history of low-back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine* 8:141–144
 31. Steenstra I, Anema J, Bongers P, de Vet HC, Knol D, van Mechelen W (2006) The effectiveness of graded activity for low back pain in occupational healthcare. *Occup Environ Med* 63:718–725
 32. van Tulder MW, Becker A, Bekkering T, Breen A, del Real MTG, Hutchison A, Koes BW, Laerum E, Malmivaara A (2006) European guidelines for the management of acute nonspecific low back pain in primary care. *Eur Spine J* 15:S169–S191
 33. Vroomen P, de Krom MC, Wilmsink J, Kester A, Knottnerus J (1999) Lack of effectiveness of bed rest for sciatica. *New Engl J Med* 340:418–423