

# Clinical Rehabilitation

<http://cre.sagepub.com/>

---

## **Do Numerical Rating Scales and the Roland-Morris Disability Questionnaire capture changes that are meaningful to patients with persistent back pain?**

Julia M Hush, Kathryn M Refshauge, Gerard Sullivan, Lorraine De Souza and James H McAuley

*Clin Rehabil* 2010 24: 648 originally published online 8 June 2010

DOI: 10.1177/0269215510367975

The online version of this article can be found at:

<http://cre.sagepub.com/content/24/7/648>

---

Published by:



<http://www.sagepublications.com>

**Additional services and information for *Clinical Rehabilitation* can be found at:**

**Email Alerts:** <http://cre.sagepub.com/cgi/alerts>

**Subscriptions:** <http://cre.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations:** <http://cre.sagepub.com/content/24/7/648.refs.html>

# Do Numerical Rating Scales and the Roland-Morris Disability Questionnaire capture changes that are meaningful to patients with persistent back pain?

**Julia M Hush, Kathryn M Refshauge, Gerard Sullivan** Faculty of Health Sciences, The University of Sydney, Australia, **Lorraine De Souza** School of Health Sciences and Social Care, Brunel University, UK and **James H McAuley** Faculty of Health Sciences and Faculty of Medicine, The University of Sydney, Australia

Received 29th October 2009; returned for revisions 8th January 2010; revised manuscript accepted 2nd February 2010.

**Objectives:** To investigate patients' views about two common outcome measures used for back pain: Numerical Rating Scales for pain and the Roland-Morris Disability Questionnaire.

**Subjects:** Thirty-six working adults who had previously sought primary care for back pain and who could speak and read English.

**Method:** Eight focus groups were conducted to explore participants' views about the 11-point Numerical Rating Scales and the 24-item Roland-Morris Disability Questionnaire. Each group was led by a facilitator and an interview topic guide was used. Audio recordings of focus groups were transcribed verbatim. Framework analysis was used to chart participants' views and an interpretive analysis performed to explain the findings.

**Results:** Participants reported that neither the Roland-Morris nor the Numerical Rating Scales captured the complex personal experience of pain or relevant changes in their condition. The time-frame of assessment was identified as particularly problematic and the Roland-Morris did not capture relevant functional domains.

**Conclusion:** This study provides empirical data that working adults with persistent back pain consider these clinical outcome measures largely inadequate. These measures currently used for back pain may contribute to misleading conclusions about treatment efficacy and patient recovery.

## Introduction

Optimal measures to evaluate patients' progress in the clinical rehabilitation of back pain remain a topic of considerable debate in the literature.<sup>1–5</sup> Two of the most commonly used instruments are the 11-point Numerical Rating Scale<sup>6</sup> for pain and the 24-item Roland-Morris Disability

Questionnaire.<sup>7</sup> The expert advisory group, the Initiative on Methods, Measurement and Pain Assessment in Clinical Trials (IMMPACT), has endorsed the use of these two patient-reported outcome instruments.<sup>4</sup> IMMPACT has also recommended that the adequacy of such measures, from the patients' perspective, be investigated by consulting relevant patient groups.<sup>5</sup> However, the 2006 IMMPACT report<sup>5</sup> identified that no attempt has been made to consult relevant patient groups about whether current outcome measures are meaningful or whether the instructions or item content of the scales are adequate. To date, patient

---

Address for correspondence: Julia Hush, Discipline of Physiotherapy, Faculty of Health Sciences, University of Sydney, 75 East St, Lidcombe, NSW 2141, Australia.  
e-mail: julia.hush@sydney.edu.au

validation data about Numerical Rating Scales and the Roland-Morris are lacking. This issue is critically important. If patients view these measures as sufficient to capture their progress during clinical rehabilitation, this would provide the first empirical data to validate the use of these instruments as meaningful to patients. If, on the other hand, patients consider these measures to be inadequate, the basis will be established for an agenda to urgently explore more appropriate measures. There is preliminary evidence that patients consider these instruments imperfect.<sup>8,9</sup> However, a thorough exploration of patients' views about these standard patient-reported outcome measures for back pain is lacking.

The aim of this study was to examine the views of people with persisting or previous back pain about whether Numerical Rating Scales and the Roland-Morris Disability Questionnaire capture meaningful changes in their condition.

## Materials

Focus groups were conducted to investigate patients' views about the adequacy of the 11-point Numerical Rating Scale<sup>6</sup> and the 24-item Roland-Morris Disability Questionnaire.<sup>7</sup> This methodology is optimal for understanding views underlying individual experiences in health research.<sup>10,11</sup> This study was approved by the Human Research Ethics Committee at the University of Sydney (HREC: 02-2007/9763).

Participants were recruited from a database of 973 individuals who had completed a large cohort study of primary care management of back pain.<sup>12</sup> A typical case sampling strategy<sup>13</sup> was used to recruit working adults with persistent back pain to the present study, as we wanted to investigate the views of this specific patient group. In order to also investigate the views of those whose condition had considerably changed, we further sampled individuals from this database who had recovered. An inclusion criterion for all participants was the ability to speak and read English. The first eligible 36 participants were included. Each participant was provided with a study information sheet and consent form. In a study previously reported<sup>14</sup> we separately investigated these participants' views of the meaning of recovery from back pain.

Eight separate focus groups were conducted, each of 2 hours duration. Each group was composed of an average of five participants. In accordance with standard focus group methods<sup>15</sup> the discussion in each group was led by a facilitator using pre-determined questions (Table 1). The conversation was flexible and responsive to issues as they arose so that participants' views could be further explored where appropriate. Participants completed four different versions of the 11-point Numerical Rating Scale<sup>6</sup> (Appendix 1) and the 24-item Roland-Morris Disability Questionnaire (Appendix 2).<sup>7</sup> The facilitator specifically explored participants' views about the adequacy of these instruments. Interviews continued until emergent themes from the data were sufficiently conceptualized as to not warrant further interviews. Each focus group was audio recorded.

Audio recordings were transcribed verbatim. Participants were assigned a code number and a pseudonym to ensure confidentiality. Transcripts were checked against audio files for accuracy (first author) and compared with scribe notes taken during each focus group to validate the transcribers' speech allocation to individual participant. The first stage of textual analysis of the interview transcripts was conducted by the first author. This was achieved by an iterative process of identifying and reviewing the evolving themes and coding definitions. Two additional reviewers then reviewed the definitions, coding and themes identified by the first author with careful reference

**Table 1** Interview topic guide

---

**Introductory questions**

Can you briefly introduce yourself to the group?  
Can you please tell me about your back condition?

**Evaluation of outcome measures**

These are questionnaires that researchers use as indicators of recovery from back pain.

Do these questionnaires capture what is important to you with changes in your back condition?

- How well do you think these pain scales [NRS] would capture your improvements or recovery?
- How well do you think this [RMDQ] questionnaire would capture your improvements or recovery?
- What is not captured by these instruments that is important to you?

**Closing question**

Is there anything else you would like to add?

---

to the source transcripts and in consultation with the first author. The framework method of data management<sup>16,17</sup> was used to arrange data from transcripts into charts that summarized themes, together with illustrative quotes. For the qualitative data, descriptors such as few ( $n=3-4$ ) some ( $n=5-6$ ), several ( $n=7-8$ ), many ( $n=9-18$ ), a majority ( $n=19-25$ ) and most ( $n=26-36$ ) are used below to provide an indication of the frequency of ideas expressed by participants.<sup>14</sup>

Descriptive statistics of participants' pain and disability scores were analysed using SPSS version 16.0 (SPSS Inc, Chicago, IL, USA).

## Results

Thirty-six participants from a range of sociodemographic backgrounds were interviewed. Mean (SD) participant age was 41.6 (6.4) years; 21 (58%) participants were male and 15 (42%) female, which reflects the gender distribution in the original cohort study that we recruited from.<sup>12</sup> A breadth of experiences was described regarding their history of back pain, care that had been sought and impact on their lives. Twenty-five participants had persistent or recurrent back pain with similar levels of pain and disability to those typically recruited for treatment efficacy trials.<sup>18</sup> Roland-Morris disability scores ranged from 2 to 19 (mean, SD: 6.6, 4.6), and pain scores on the four Numerical Rating Scales ranged from 0 to 8 (mean, SD: 2.8, 3.2) (Table 2). Eleven participants had recovered from previous back pain, established by the participants' self-rating of recovery.

### Participants' views of Numerical Rating Scales

A general view expressed by the majority of participants was that the Numerical Rating Scale did not adequately capture the complexity of their personal experience of pain. A few, however, stated that the instrument would capture their progress in the acute stage of back pain. Two specific themes about the adequacy of the instrument were: (1) the meaning attributed to the pain score and (2) the time-frame of measurement. Issues relating to these themes are summarized with illustrative participant quotes below.

**Table 2** Participants' pain and disability scores

Variable <sup>a</sup>	Range	Mean (SD)
Participants with persistent pain ( $n=24$ )		
Pain: worst in past 24 hours	0 to 8	3.6 (2.5)
Pain: least in past 24 hours	0 to 5	2.1 (1.5)
Pain: on average	0 to 7	3.2 (1.8)
Pain: right now	0 to 8	2.7 (2.3)
Disability	2 to 19	6.6 (4.6)
Recovered participants ( $n=11$ )		
Pain: worst in past 24 hours	0 to 3	0.6 (1.0)
Pain: least in past 24 hours	0 to 1	0.2 (0.4)
Pain: on average	0 to 2	0.4 (0.7)
Pain: right now	0 to 3	0.5 (0.9)
Disability	0 to 4	1.2 (1.4)

<sup>a</sup>Pain measured on the 11-point Numerical Rating Scales; disability measured on the 24-item Roland-Morris Disability Questionnaire.

#### 1) Meaning attributed to the pain score

It is commonly assumed that the Numerical Rating Scale measures pain intensity, the sensory component of pain. However, participants in this study reported that their score reflected many other aspects of the pain experience. Some indicated that their score could include an evaluative aspect, including the impact of pain on factors such as independence, coping, sleep or work, emphasizing the complexity and idiosyncratic nature of patients' pain experiences.

*For me 'ten' was the pain combined with ... being reliant on other people ... as well as coping with the pain. My zero would be not having to think about everything I do, the consequences of doing something. (Sue, 33)*

*It wasn't that the pain was unbearable, the pain was at, say 70% of what I can handle, but just because it was constant and you couldn't sleep and you couldn't get away from it, it just wore me down. So that was my 'ten'. (Ruth, 37)*

A common view was that Numerical Rating Scale scores are highly dependent on individual experiences of pain that can determine the benchmark used by a patient to rate their pain. Of particular interest were comments that this internal scale by which individuals rate pain can change over time. For example, some participants reported that their pain 'shifted down' as their

back condition persisted, interpreted by a few as an increase in pain tolerance rather than a reduction in pain intensity.

*I think our 'zero' also changes, because you can't remember what your back was like before, so your starting point is that there's something there in your back: that's your new 'zero'. (Peter, 50)*  
*You might call this a 'ten' until you've realized that it was 'eleven'! Until people ... have been exposed to the full range of experiences, they don't know where to benchmark. (Dean, 40)*

## 2) Time-frame of measurement

The majority of participants believed that versions of the Numerical Rating Scale that assess pain 'in the past 24 hours' or 'right now' were unlikely to capture improvements because of symptom fluctuation.

*It [24 hours] is a very small window in your life of pain. (Sue, 33)*  
*Why is it concentrating only on what has happened today, or 24 hours, because I think ... the variation and the fluctuation is so much that what is today may not be anything tomorrow ... but to me it looks like 24 hours is not really capturing exactly what's here. (Ben, 35)*

A few reported that average measures would be more appropriate, particularly if evaluated over a week or a month.

*The pain can be fairly different, so for me, it [pain assessment] should be over a period of a week or a month. (Ruth, 37)*

Three others stated that the spectrum of their pain severity was important to capture. Some participants commented on the absence, but relevance, of assessment of episode duration or frequency, deemed particularly important for those with recurrent pain.

## Patients' views of the Roland-Morris Disability Questionnaire

A general opinion of most participants in this study was that the Roland-Morris was not an adequate measure for chronic back pain.

Five participants stated that the instrument would be most relevant for the acute stage of back pain. Some identified a general discordance between the questionnaire score and their experience of pain-related disability.

*If my sheet [RMDQ] was sent to [the Workers' Compensation authority] ... anyone who reads it would say 'he's good, he's all right, fine, nothing.' But I'm not happy with it because that's not fully describing me. (Tim, 42)*

Two specific themes about the inadequacies of the Roland-Morris were identified: (1) the relevance of items and (2) the time-frame of measurement. Issues relating to these themes are summarized, with illustrative participant quotes below.

### 1) Relevance of items

Participants in this study were all working and many had relatively high levels of daily function, despite persistent pain. These participants pointed out the irrelevance of numerous Roland-Morris items relating to basic levels of activity and self-care.

*You would give this to somebody who you are assessing to see if they can care for themselves ... I think it's probably limited to people who probably aren't like us ... we all work. (Sue, 33)*  
*I stay at home most of the time because of my back': I don't really stay at home ... I still go to work. That doesn't relate to me. (Paul, 45)*

Importantly, participants stated that there were meaningful and important dimensions of functioning that were not captured by the Roland-Morris. These included work, driving, housework, leisure and exercise, social functioning, psychological health and adaptive strategies.

Meaningful aspects of work included job effectiveness and the capacity to perform both light and heavy work tasks.

*Because it's looking at whether we are getting dressed ... but there's nothing there about work ... There's nothing there that can qualify or quantify if we're actually doing our job as effectively because of our back. (Kathie, 51)*



Participants reported that driving was particularly important to assess because of the impact on independence and the ability to work or socialize.

*My main difficulty was because I couldn't drive ... It was only for a week and a half ... I felt totally housebound ... I couldn't visit friends, go shopping. (Leslie, 52)*

A few participants suggested that the assessment of the capacity for housework in the Roland-Morris, which has two items that refer generally to 'jobs around the house', was inadequate. They reported that the ability to do tasks such as clean the bathroom, cook, garden and shop were more relevant.

*Lawn mowing [is missing]. As soon as my back goes: that's it, I don't mow my lawns for a couple of months. (Mick, 40)*  
*Things like [cleaning the] bathroom and shower and stuff because you have to get right in and you're bending over when you're scrubbing. (Angela, 35)*

Many were concerned that leisure or exercise activities were not captured by these instruments. This was judged a serious exclusion by these participants who considered such highly valued activities as vital indicators of progress of their condition.

*If I want to do so many more things, there's a chance it [RMDQ] may or may not be sufficient; ... if pain stopped me from scuba diving, that might be something I'd cry about. (Murray, 37)*  
*I think all of us in some regard said our injury has affected our social or sporting lives ... I think the sport for us has been a great deal of our lives prior to our injury. I think that's affected us more than ... our home lives. (Sue, 33)*

Issues related to social functioning included independence, relationships with family members, work colleagues and friends, and the ability to socialize.

*[What is missing?] Perhaps lifestyle things ... We're talking about socialization ... am I going out and seeing people? Has it stopped me from*

*doing that? So am I able to still see my friends or family or whatever? All these sorts of issues it doesn't cover. (Kathie, 51)*

*One thing it doesn't say, it all affects your lifestyle ... and there's no such thing as that: no question: 'because of my back pain I can't do many things like go play with my kids'. (Alex, 48)*

Many participants stated that assessment of the impact of back pain on their psychological health would be valuable, including thoughts and mood, as well as the impact of persistent pain on their personal identity.

*The thinking time, the emotional side of it ... I suppose it [the RMDQ] doesn't touch on that. (Peter, 50)*

*I think probably I should have seen two [RMDQ] sheets: one for the earlier stage that people have ... for us we're sort of recovered, or we are in it for two, three years, so probably another sheet asking other questions: 'how do you feel, what do you think?' (Joan, 37)*

Participants with persistent pain stated that the use of adaptive strategies to manage living with pain, prevent re-injury and enhance quality of life was also pertinent to assess.

*I think the other thing that doesn't really get captured in some of this is not the pain avoidance side of it, but what attributes you do to make sure you're going to mitigate against that. So people go to the gym, physio ... doing you know ... strategies to cope with it. (Dean, 40)*

## 2) Time-frame of measurement

The Roland-Morris assesses an individual's level of pain-related disability on the day of assessment ('as you read the list, think of yourself today'). Participants explained that because their pain level was not static, their score on one day was unlikely to reflect the overall picture of their pain-related disability. This reveals a fundamental problem with the time-frame of assessment that is embedded in this instrument.

*[For] people who are living with a certain level of pain, that's probably a relevant set of questions;*

*but for a lot of us it would seem, it's not the way: you have these episodic rather than daily problems. (Dean, 40)*

*It is only asking for the information for 'today' and most of that didn't apply to me. (Meagan, 21)*

## Discussion

By investigating the perspective of individuals with persisting and previous back pain, this study has revealed considerable inadequacies with the two most common instruments used to evaluate the progress of individuals with persisting back pain. Participants in the study thought that their experience of pain and the impact of the condition on their lives were not well captured by the Numerical Rating Scale or the Roland-Morris. Many participants also considered the time-frame of these measures to be inappropriate.

The time-frame of assessment was an issue common to both instruments. From the participants' perspective, being asked to rate their symptoms or abilities 'today', 'in the past 24 hours' or 'right now', excludes the reliable capture of improvements if symptoms fluctuate or episodes recur, as is typical in the course of back pain.<sup>19,20</sup> Capturing 'typical' or 'average' pain experiences and function over a longer period of time, may therefore be more appropriate for chronic back pain, as was suggested by participants in this study.

One intriguing finding was that the internal scale or baseline, against which improvements in pain are benchmarked, may be unstable over time. This may reflect habituation or a 'response shift' phenomenon, which is a change in the meaning of one's self-evaluation of a target construct,<sup>21</sup> and has been reported previously in patients' use of health survey research instruments.<sup>8,22</sup> The direction of the response shift was downwards; that is, a similar level of pain was rated lower over time. Curiously, the opposite would be predicted from current pain neurobiology, where central and peripheral neurobiological changes associated with persistent pain, amplify the pain experience.<sup>23</sup> We are not aware that this phenomenon has been reported previously for chronic back pain.

Numerical Rating Scales have been designed to measure the sensory–discriminative dimension of

pain. We found that patients tend to integrate cognitive–evaluative aspects into their assessment of pain intensity. This supports previous findings that patients can incorporate multiple dimensions of pain into their pain ratings, including the impact of pain on function, affect<sup>9</sup> and quality of life.<sup>8</sup> However, when we asked participants what was lacking from their assessment using these two instruments, a common response was 'how you feel' or 'the emotional side of it', suggesting that many patients consider the affective dimension of pain should be explicitly assessed. While the multi-dimensional nature of pain is universally acknowledged<sup>24</sup> and the appropriateness of measuring pain with a single-item scale has previously been challenged,<sup>25</sup> the use of unidimensional 0–10 pain rating scales (verbal, visual or numerical) remains the norm for research and clinical use today.

In a previous study investigating the relationship between pain scores and recovery status in those with back pain, we found that a low pain score does not clearly distinguish those recovered from those unrecovered.<sup>14</sup> Both the previous and current studies indicate that it may be timely to consider alternative approaches to pain assessment for chronic back pain, particularly to capture multiple dimensions of the pain experience. Given the statistical imperative to be parsimonious and avoid the use of multiple outcome measures in research to avoid type I errors,<sup>26</sup> as well as to avoid respondent fatigue, the use of multiple scales is clearly not the answer. The McGill Pain Questionnaire<sup>25,27</sup> is one well-recognized instrument that is designed to capture multidimensional aspects of pain and may address some of these issues. However, patients' views of this and other more comprehensive pain assessment instruments are unknown as they have not yet been investigated.

The Roland-Morris is recommended by IMMPACT to measure the core outcome domain of physical functioning<sup>4</sup> and is a commonly used outcome measure for back pain in clinical practice and research. Our study shows that many of the items in this instrument may be irrelevant to those with persistent back pain and that the assessment time-frame is inappropriate. While there are many alternative instruments designed to assess function in back pain,<sup>28</sup> of those with acceptable measurement properties,

none includes the spectrum of domains that appear to be of greatest concern to patients, as indicated by the findings of the present study. To address these limitations of standardized items, patient-specific measures warrant consideration as a possible alternative. Examples are the Patient-Generated Index of Quality of Life, which has been tested in a back pain population<sup>29,30</sup> and the Patient-Specific Functional Scale.<sup>31</sup> The key advantage of these instruments is that they address the idiosyncratic nature of issues salient to individual patients as patients nominate relevant functional items of concern and so may have greater relevance in the context of the daily lives of patients with back pain.

We have previously reported the significance of capturing relevant functional domains for individuals with persisting back pain.<sup>14</sup> Whether people with back pain see themselves as recovered or not depends on their cognitive appraisal of the impact of symptoms on their ability to perform meaningful daily activities and fulfil social roles, rather than simply the presence or absence of pain.<sup>14</sup>

Based on patients' views in this study, ideal outcome instruments for this population would capture the sensory, affective and evaluative aspects of pain, as well as relevant aspects of function (e.g. work, driving, leisure, social functioning and psychological health), evaluated over an appropriate time-frame. The views of back pain patients reported in this study align well with the findings from the recent IMMPACT focus groups and patient survey that identified a similar scope of relevant domains from the perspective of people who suffer a range of chronic pain conditions.<sup>32</sup> Further research is warranted to investigate patients' views of the adequacy of existing, potentially more relevant instruments to capture the multidimensional experience of pain and the impact on function.

There have been few patient validation studies of outcome measures in the field of back pain. This study contributes innovative and valuable data about patients' views of measures that are routinely used in clinical rehabilitation and for research. A strength of this study is that multiple strategies were used to enhance credibility: (i) the sampling strategy was clearly defined; (ii) we included participants who had persisting pain as well as those recovered to enable valid evaluation

of the adequacy of the outcome measures to capture meaningful changes; (iii) we chose an appropriate method of data collection to understand individual views and obtain information-rich data; and (iv) the audio-taping and transcription methods were reliable and independently verifiable. A limitation of this study is that only one researcher initially coded the key themes and issues, although two additional researchers were involved in further review of both codes and themes. It should be noted that the views of patients with acute back pain, those not working or non-English speaking participants may differ from those reported in this study.

The Roland-Morris Disability Questionnaire and Numerical Rating Scales that are currently recommended as outcome measures for back pain may not capture changes that are meaningful to working adults with persistent back pain, and may contribute to misleading conclusions about treatment efficacy and patient recovery. There is an imperative to identify or develop alternative instruments that better capture the breadth and depth of domains that are more meaningful to patients with persistent back pain to evaluate their progress to recovery.

#### Clinical messages

- Neither the Roland-Morris Disability Questionnaire nor Numerical Rating Scales seem to capture the complex personal experience of back pain or relevant changes in the patients' condition.
- The time-frame of assessment is too narrow and the Roland-Morris does not capture the breadth of functional domains that are relevant to patients.

#### Acknowledgements

We are grateful to the patients who shared their back pain experiences with us. Without their generous and open contributions, this study would not have been possible. The authors also thank Dr Doris McIlwain from Macquarie University, Sydney, and Dr Bronwyn Hemsley from the



University of Sydney for invaluable guidance in the design, conduct and analysis of this study.

This study was wholly conducted at the Faculty of Health Sciences at the University of Sydney.

### Competing interests

There are no conflicts of interest.

### Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

### References

- 1 Deyo RDA, Battie M, Beurskens AJHM *et al.* Outcome measures for low back pain research: a proposal for standardized use. *Spine* 1998; **23**: 2003–13.
- 2 Turk DC, Rudy TE, Sorkin BA. Neglected topics in chronic pain treatment outcome studies: determination of success. *Pain* 1993; **53**: 3–16.
- 3 Turk DC, Dworkin RH, Allen RR *et al.* Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2003; **106**: 337–45.
- 4 Dworkin RH, Turk DC, Farrar JT *et al.* Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2005; **113**: 9–19.
- 5 Turk DC, Dworkin RH, Burke LB *et al.* Developing patient-reported outcome measures for pain clinical trials: IMMPACT recommendations. *Pain* 2006; **125**: 208–15.
- 6 Downie WW, Leatham PA, Rhind VM *et al.* Studies with pain rating scales. *Ann Rheum Dis* 1978; **37**: 378–81.
- 7 Roland M, Morris R. 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* 1983; **8**: 141–4.
- 8 Ong BN, Hooper H, Jinks C *et al.* 'I suppose that depends on how I was feeling at the time': perspectives on questionnaires measuring quality of life and musculoskeletal pain. *J Health Serv Res Policy* 2006; **11**: 81–8.
- 9 Williams ACdeC, Davies HTO, Chadury Y. Simple pain rating scales hide complex idiosyncratic meanings. *Pain* 2000; **85**: 457–63.
- 10 Minichiello V, Sullivan G, Greenwood K *et al.* *Handbook of research methods for nursing and health science*. Sydney: Pearson, 2004.
- 11 Pope C, Mays N. Qualitative research: reaching the parts other methods cannot reach. *BMJ* 1995; **311**: 42–45.
- 12 Henschke N, Maher CG, Refshauge KM *et al.* Prognosis in patients with recent onset low back pain in primary care. *BMJ* 2008; **337**: a171.
- 13 Llewellyn G, Sullivan G, Minichiello V. Sampling in qualitative research. In Minichiello V, Sullivan G, Greenwood K, Axford R. eds. *Handbook of research methods for nursing and health science*. Sydney: Pearson, 2004.
- 14 Hush J, Refshauge K, Sullivan G *et al.* Recovery: what does this mean to patients with low back pain? *Arthritis Rheum Arthritis Care Res* 2009; **61**: 124–31.
- 15 Morgan DL, Krueger RA. *The focus group kit*. Sage: Thousand Oaks, CA, 1998.
- 16 Pope C, Ziebland S, Mays N. Qualitative research in health care: analysing qualitative data. *BMJ* 2000; **320**: 114–16.
- 17 Ritchie J, Lewis J. *Qualitative research practice: a guide for social science students and researchers*. London: Sage Publications, 2003.
- 18 Frost H, Lamb SE, Doll HA *et al.* Randomised controlled trial of physiotherapy compared with advice for low back pain. *BMJ* 2004; **329**: 708.
- 19 Von Korff M, Saunders KW. The course of low back pain in primary care. *Spine* 1996; **21**: 2833–7.
- 20 Carey TS, Garrett JM, Jackman A *et al.* Recurrence and care seeking after acute back pain: results of a long-term follow-up study. *Med Care* 1999; **37**: 157–64.
- 21 Schwartz CE, Sprangers MA. Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Soc Sci Med* 1999; **48**: 1531–48.
- 22 Mallinson S. Listening to respondents: a qualitative assessment of the Short-Form 36 Health Status Questionnaire. *Soc Sci Med* 2002; **54**: 11–21.
- 23 Woolf CJ, Salter MW. Neuronal plasticity: increasing the gain in pain. *Science* 2000; **288**: 1765–8.
- 24 Melzack R, Casey KL. Sensory, motivational and central control determinants of pain: a new conceptual model. In Kenshalo D. ed. *The skin senses*. Springfield: Thomas, 1968.
- 25 Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975; **1**: 277–99.
- 26 Turk DC, Dworkin RH, McDermott MP *et al.* Analyzing multiple endpoints in clinical trials of pain treatments: IMMPACT recommendations. *Pain* 2008; **139**: 485–93.

27 Melzack R, Torgerson WS. On the language of pain. *Anesthesiology* 1971; **34**: 50–9.

28 Grotle M, Brox JI, Vøllestad NK. Functional status and disability questionnaires: what do they assess? *Spine* 2004; **30**: 130–40.

29 Ruta DA, Garratt AM, Leng M *et al.* A new approach to the measurement of quality of life: the Patient-Generated Index. *Med Care* 1994; **32**: 1109–26.

30 Martin F, Camfield L, Rodham K *et al.* Twelve years' experience with the Patient Generated Index (PGI) of quality of life: a graded structured review. *Qual Life Res* 2007; **16**: 705–15.

31 Stratford P, Gill C, Westaway M *et al.* Assessing disability and change on individual patients: a report of a patient-specific measure. *Physiother Can* 1995; **47**: 258–63.

32 Turk DC, Dworkin RH, Revicki D *et al.* Identifying important outcome domains for chronic pain clinical trials: An IMMPACT survey of people with pain. *Pain* 2008; **137**: 276–85.

4) Please rate your pain by circling the one number that tells how much pain you have **RIGHT NOW**.

0 1 2 3 4 5 6 7 8 9 10  
 No Pain as  
 pain bad as  
 you can  
 imagine

**Appendix 1 - The four versions of the 11-point Numerical Rating Scales evaluated by participants**

1) Please rate your pain by circling the one number that best describes your pain at its **WORST** in the past 24 hours.

0 1 2 3 4 5 6 7 8 9 10  
 No Pain as  
 pain bad as  
 you can  
 imagine

2) Please rate your pain by circling the one number that best describes your pain at its **LEAST** in the past 24 hours.

0 1 2 3 4 5 6 7 8 9 10  
 No Pain as  
 pain bad as  
 you can  
 imagine

3) Please rate your pain by circling the one number that best describes your pain on the **AVERAGE**.

0 1 2 3 4 5 6 7 8 9 10  
 No Pain as  
 pain bad as  
 you can  
 imagine

**Appendix 2 - The 24-item Roland-Morris Disability Questionnaire**

When your back hurts, you may find it difficult to do some of the things you normally do.

This list contains some sentences that people have used to describe themselves when they have back pain. When you read them, you may find that some stand out because they describe you **today**. As you read the list, think of yourself **today**. When you read a sentence that describes you today, fill the box to the left of the sentence. If the sentence does not describe you, then leave the box blank and go on to the next one. Remember, only mark the sentence if you are sure that it describes you **today**.

- 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.