

ORIGINAL ARTICLE

The SF-36 walk-wheel: a simple modification of the SF-36 physical domain improves its responsiveness for measuring health status change in spinal cord injury

BB Lee^{1,2,3}, JM Simpson², MT King⁴, MJ Haran⁵ and O Marial¹

¹Spinal Medicine Department, Prince of Wales Hospital, Sydney, Australia; ²School of Public Health, University of Sydney, Sydney, Australia; ³School of Public Health, University of New South Wales, Sydney, Australia; ⁴Centre for Health Economics Research & Evaluation, University of Technology, Sydney, Australia and ⁵Department of Aged Care and Rehabilitation Medicine Royal North Shore Hospital, Sydney, Australia

Objective: To evaluate the validity and responsiveness of a modified SF-36 within a spinal cord-injured (SCI) population.

Study Design: SF-36 scores collected at baseline and on completion of a randomized controlled trial in 305 patients with SCI and neuropathic bladder.

Setting: New South Wales, Australia.

Methods: Subjects were administered the standard SF-36 plus three additional questions, in which 'walk' was replaced with 'wheel' for three of the physical function (PF) questions. Discriminant validity was determined by comparing participants with paraplegia and tetraplegia using the effect size (ES). Responsiveness was assessed in the subset of patients who developed a urinary tract infection (UTI) during the trial using the standardized response mean (SRM).

Results: Compared with the standard SF-36, the SF-36 walk-wheel modification (SF-36ww) increased the mean PF score from 18 to 39 ($P < 0.001$) and the physical composite score from 33 to 37 ($P < 0.001$). Discriminant validity was similar for both versions (PF paraplegia/tetraplegia: ES 1.09(SF-36) vs 1.08(SF-36ww), $n = 305$). Among 138 SCI patients who developed a UTI, the SF-36ww almost doubled PF responsiveness for all neurological levels (SRM increased from 0.36 to 0.68), more so in tetraplegic (SRM, 0.11 vs 0.58; $n = 77$) than paraplegic groups (SRM, 0.77 vs 0.86; $n = 61$).

Conclusion: The SF-36ww is a simple, pragmatic modification of the SF-36 PF items, which addresses some problems of content validity and floor effect for SCI subjects and greatly improves responsiveness, particularly for those with tetraplegia. Because it comprises a simple addition to the standard SF-36, external comparisons are preserved.

Spinal Cord (2009) 47, 50–55; doi:10.1038/sc.2008.65; published online 17 June 2008

Keywords: spinal cord injury; SF-36; validity; SF-36ww; responsiveness; UTI

Introduction

The SF-36 is a generic health status measurement tool, which has been widely used for research in spinal cord injury (SCI) as well as for many other disease groups. It comprises eight domains: physical functioning (PF), physical role limitation, emotional role limitation, bodily pain, general health, vitality, social functioning and mental health; physical and mental composite scale scores (PCS, MCS) can also be derived from these domains.¹ When used in SCI, the SF-36 has been demonstrated to have sufficient discrimination to compare

the health status of those with SCI with that of other health populations² and to be able to detect disease-state change for urinary tract infection (UTI) within the SCI population.³

The SF-36 is not without problems when used in SCI and other severely disabled populations. Particular problems have been demonstrated with the PF domain including significant floor effects due to the inability of many patients to perform some of the physical tasks described. This has the effect of limiting responsiveness and creates problems when correlating the PF domain with other SF-36 domains.³ Another major difficulty has been content validity and acceptability within the SCI population. PF questions that relate specifically to walking and stair climbing (SF-36 items 3d, 3e and 3g–i) may be considered insulting or irrelevant for some SCI individuals.^{3–7} Tate *et al.*⁷ and Meyers and

Correspondence: Dr BB Lee, Spinal Medicine, Prince of Wales Hospital, High Street Entrance level 2, Randwick, Sydney, New South Wales 2031, Australia.
E-mail: blee@medicalinformatics.net
Received 21 December 2007; revised 18 March 2008; accepted 20 March 2008; published online 17 June 2008

Andresen⁸ have suggested replacing the words 'walk' and 'climb' with 'go' and 'go up', respectively, to address this problem. Tate *et al.*⁷ states that construct validity remains adequate with these changes, but to date there has been no published validation study of this type of modification in SCI populations compared with the standard SF-36. In this paper, we aim to validate a modified SF-36 against the standard SF-36 within an SCI sample. We determine whether this modification improves internal SCI discriminant validity, responsiveness to disease-state change and physical floor effects while retaining comparability with other population groups. Supplementary PF, PCS and MCS scores (PFww, PCSww and MCSww, respectively) were thus generated.

Methods

The SF-36 scores were collected as part of the spinal-injured neuropathic bladder antiseptis randomized controlled trial (RCT) in patients with SCI and neuropathic bladder. Between November 2000 and August 2002, 543 eligible patients (mostly community dwelling) were invited to participate in the study, of whom 305 (56%) agreed.⁹ Subjects completed the standard SF-36 plus three additional questions, replacing

the word 'walk' with 'wheel' for PF questions 9–11 (items 3g–i) of the SF-36. The original questions were also asked, allowing coding to either SF-36 walk-wheel (SF-36ww) or the original SF-36 (Box 1).

Measurements were made at baseline, on enrolment in the RCT and then either on development of the first UTI or (if no UTI occurred) at 6-month follow-up. The characteristics and inclusion criteria of this SCI sample have been described previously.^{2,9} The SF-36ww was collected with the assistance of a research officer. This included physically assisting completion of the questionnaire where necessary.

Content validity was assessed by a retrospective review of datasheets for unprompted comments or indications of problems recorded by the patient or assistant on the data collection form during the course of administering the SF-36. The number of participants who made a comment on either baseline or follow-up questionnaire (for the walking questions only) was recorded.

Discriminant validity was determined by comparing participants with paraplegia and tetraplegia. Effect sizes were assessed using the formula: effect size (ES) = $(m_1 - m_2) / s_1$, where m_1 = reference group mean, m_2 = comparison mean, s_1 = reference group standard deviation.¹⁰ Internal consistency was assessed using Cronbach's α .^{11,12}

Box 1 SF-36ww^a

The following questions are about activities you might do during a typical day. In the past 1-week does your health limit you in these activities? If so, how much?

(Please circle one number on each line)

| ACTIVITIES | Yes Limited A lot | Yes Limited A little | No, Not Limited At All |
|---|-------------------|----------------------|------------------------|
| 3a: Vigorous activities, such as running, lifting heavy Objects, participating in strenuous sports | 1 | 2 | 3 |
| 3b: Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf | 1 | 2 | 3 |
| 3c: Lifting or carrying groceries | 1 | 2 | 3 |
| 3d: Climbing several flights of stairs | 1 | 2 | 3 |
| 3e: Climbing one flight of stairs | 1 | 2 | 3 |
| 3f: Bending, kneeling, or stooping | 1 | 2 | 3 |
| 3g: Walking more than one kilometre | 1 | 2 | 3 |
| 3h: Walking half a kilometre | 1 | 2 | 3 |
| 3i: Walking 100 metres | 1 | 2 | 3 |
| 3g ww: Wheeling more than one kilometre | 1 | 2 | 3 |
| 3h ww: Wheeling half a kilometre | 1 | 2 | 3 |
| 3i ww: Wheeling 100 metres | 1 | 2 | 3 |
| 3j: Bathing or dressing yourself | 1 | 2 | 3 |

^aModified from SF-36¹: Items 3 (a to j) are the original SF-36 questions, while 3g ww to 3i ww (shaded area) comprise the supplementary SF-36ww modification.

Table 1 Cross-sectional comparisons at baseline of the SF-36 physical activity question (question 3): walking (q3g-i) compared to the SF-36ww wheeling items (q3gww-q3iww)

| Walking | Wheeling | | | Total (%) |
|--|---------------|------------------|--------------------|-----------|
| | Limited a lot | Limited a little | Not limited at all | |
| <i>(a) Walking (q3g) compared to wheeling (q3gww) more than 1 km</i> | | | | |
| Limited a lot | 80 | 60 | 151 | 291 (96) |
| Limited a little | — | — | 10 | 10 (3) |
| Not limited at all | — | — | 4 | 4 (1) |
| Total (%) | 80 (26) | 60 (20) | 165 (54) | 305 (100) |
| <i>(b) Walking (q3h) compared to wheeling (q3hww) more than half a kilometer</i> | | | | |
| Limited a lot | 50 | 51 | 186 | 287 (94) |
| Limited a little | — | — | 9 | 9 (3) |
| Not limited at all | — | — | 9 | 9 (3) |
| Total (%) | 50 (16) | 51 (17) | 204 (67) | 305 (100) |
| <i>(c) Walking (q3i) compared to wheeling (q3iww) more than 100 m</i> | | | | |
| Limited a lot | 31 | 30 | 221 | 282 (93) |
| Limited a little | — | — | 4 | 4 (1) |
| Not limited at all | — | — | 19 | 19 (6) |
| Total (%) | 31 (10) | 30 (10) | 244 (80) | 305 (100) |

Shaded areas show a higher score on the SF-36ww than the equivalent SF-standard 36 item.

Responsiveness was assessed in the subset of patients who developed a UTI during the trial. UTI onset is a suitable condition for assessing responsiveness in the patient population because it is a common condition in SCI that leads to clinically relevant consequences to test health-state changes. Responsiveness was analysed by calculating change scores and standardized response means¹³ (SRM = mean change/standard deviation of change) for those patients ($n = 138$) who developed a UTI during the course of the clinical trial. The number of patients necessary to detect the (paired sample) change in health status associated with developing a UTI was calculated using the derived formula:^{14,15}

$$N = (Z_{1-\alpha/2} + Z_{1-\beta})^2 / (\text{SRM})^2$$

where $\alpha = 0.05$ and $\beta = 0.20$: $Z_{1-\alpha/2} = 1.96$, $Z_{1-\beta} = 0.84$ and N is the number of patients who subsequently develop a UTI. The total number of patients required for a study is then obtained by dividing N by the proportion expected to get a UTI.

Results

All participants had SCI and neuropathic bladder. The mean age was 44 with a mean elapsed time since SCI of 14 years. Participants were mostly male (83%), 55% were tetraplegic and 49% had complete spinal injury (by ASIA Impairment Scale definition¹⁹). There were no post-randomization losses.

Content validity

Retrospective review of the SF-36 data entry sheets found that 20 of 305 participants (7%) had marked the SF-36 physical activity question (3g-i) as not applicable or problematic. The SF-36ww modification (3g ww-i ww) was

problematic in just 9 of 305 participants (3%). Reasons for problems with the SF-36ww were mainly enforced 'bed rest' (six subjects). One participant stated that the use of an electric wheelchair led to problems interpreting the SF-36ww questions, whereas the remaining two responses had no reasons recorded.

Table 1 shows a comparison of the baseline responses to the standard SF-36 physical activity questions (3g-i) and the equivalent SF-36ww items (3g ww-i ww). The floor effects of the standard SF-36 walking questions are clearly demonstrated, with 93–96% of this sample being maximally limited at healthy (non-UTI) baseline. In comparison, SF-36ww scores were more evenly distributed among the response categories, with 10–26% of participants being maximally limited, and 54–80% not being limited at all.

SF-36ww summary statistics at baseline

The simple change of 'walk' to 'wheel' in the physical activity questions (3g-i) increased the overall mean baseline PF score from 18 (s.d. = 18.8) to 39 (s.d. = 22.4) and the PCS from 33 (s.d. = 7.7) to 37 (s.d. = 8.3), whereas the MCS was only slightly altered from 56 (s.d. = 12.1) to 54 (s.d. = 11.9). Using the paired *t*-test, these differences were all statistically significant ($P < 0.001$).

Ceiling and floor effects

The overall sample ($N = 305$) had a large floor effect of 29% (that is, subjects recorded '1-limited a lot' for every item in Box 1). Post-modification (PFww) improved to 8%. The tetraplegic subgroup ($N = 167$) accounted for most of the floor effect, which was substantially reduced from 49 to 14% with the modification. The paraplegic subgroup ($N = 138$) contributed little to the floor effect in the standard PF

Table 2 Discriminant validity of SF-36ww domain scores for patients with SCI (paraplegia compared to tetraplegia) at baseline ($n = 305$): mean modified (ww)^b and standard physical function domain and composite scores^a

| Domains | Paraplegia $n = 138$ (s.d.) | Tetraplegia $n = 167$ (s.d.) | Effect size ^c | Significance P-value |
|--------------------|--------------------------------|---------------------------------|--------------------------|-------------------------|
| PF ^a | 29 (18.36) | 9 (13.34) | 1.09 | <0.001 |
| PFww ^b | 51 (20.32) | 29 (18.69) | 1.08 | <0.001 |
| PCS ^a | 35 (8.25) | 32 (6.98) | 0.36 | 0.001 |
| PCSww ^b | 39 (8.78) | 36 (7.64) | 0.34 | 0.002 |
| MCS ^a | 55 (12.22) | 57 (12.01) | -0.16 | 0.15 |
| MCSww ^b | 53 (12.05) | 55 (11.65) | -0.17 | 0.14 |

Abbreviations: MCS, mental composite score; PCS, physical composite score; PF, physical function; ww, walk-wheel.

^aStandard SF-36 domains: PF, PCS and MCS.

^bDomains modified with the ww modifications: PFww, PCSww, MCSww.

^cSignificance test: two-sample *t*-test for all domains.

Table 3 Responsiveness: mean SF-36ww domain and composite scores for patients with SCI at baseline and after they developed a UTI ($n = 138$)

| Domain/composite score | SCI baseline (s.d.) | SCI with UTI (s.d.) | Mean change (s.d.) | SRM ^c |
|---------------------------|---------------------|---------------------|--------------------|------------------|
| <i>Overall n = 138</i> | | | | |
| PF ^a | 14.49 (15.11) | 9.49 (13.27) | 5.00 (13.87) | 0.36 |
| PFww ^b | 36.52 (21.36) | 22.64 (20.45) | 13.88 (20.43) | 0.68 |
| PCS ^a | 32.00 (7.08) | 26.90 (7.10) | 5.10 (8.78) | 0.58 |
| PCSww ^b | 36.48 (7.99) | 29.57 (7.85) | 6.91 (9.55) | 0.72 |
| MCS ^a | 55.82 (12.65) | 46.15 (13.35) | 9.67 (13.63) | 0.71 |
| MCSww ^b | 53.51 (12.43) | 44.77 (13.13) | 8.74 (13.41) | 0.65 |
| <i>Paraplegia n = 61</i> | | | | |
| PF ^a | 25.57 (13.91) | 16.31 (11.29) | 9.26 (12.07) | 0.77 |
| PFww ^b | 49.67 (18.75) | 34.67 (18.02) | 15.00 (17.54) | 0.86 |
| PCS ^a | 33.08 (7.43) | 27.56 (7.64) | 5.52 (8.98) | 0.62 |
| PCSww ^b | 37.99 (8.36) | 31.30 (8.30) | 6.69 (9.69) | 0.69 |
| MCS ^a | 55.24 (13.01) | 44.99 (12.48) | 10.25 (11.83) | 0.87 |
| MCSww ^b | 52.72 (12.87) | 43.07 (12.03) | 9.65 (11.51) | 0.84 |
| <i>Tetraplegia n = 77</i> | | | | |
| PF ^a | 5.71 (9.06) | 4.09 (12.24) | 1.62 (14.34) | 0.11 |
| PFww ^b | 26.10 (17.20) | 13.12 (17.01) | 12.99 (22.54) | 0.58 |
| PCS ^a | 31.14 (6.71) | 26.37 (6.64) | 4.76 (8.67) | 0.55 |
| PCSww ^b | 35.29 (7.53) | 28.21 (7.24) | 7.08 (9.50) | 0.75 |
| MCS ^a | 56.27 (12.42) | 47.06 (14.02) | 9.21 (14.97) | 0.62 |
| MCSww ^b | 54.13 (12.12) | 46.12 (13.86) | 8.02 (14.78) | 0.54 |

Abbreviations: MCS, mental composite score; PCS, physical composite score; PF, physical function; SCI, spinal cord injury; SRM, standardized response mean; UTI, urinary tract infection; ww, walk-wheel.

^aStandard SF-36 domains: PF, PCS and MCS.

^bDomains modified with the ww modifications: PFww, PCSww, MCSww.

^cSRM = mean change/standard deviation of change.

domain scores, but this effect was further reduced from 4.3 to 1% using the PFww.

Discriminant validity and internal consistency

Table 2 shows that the ability of the SF-36ww to discriminate tetraplegia from paraplegia is similar to that of the standard SF-36. Mean differences and effect sizes in the PF domain and the PCS and MCS composite scores between the groups were similar for the SF-36 and the SF-36ww. Cronbach's α was slightly better for the PFww scores (0.85) than for the standard PF domain scores (0.83), demonstrating good internal consistency for the walk-wheel modified PFww domain.

Responsiveness of the SF-36ww to disease-state change

The scores of the 138 patients who went on to develop a UTI were analysed for responsiveness. Table 3 demonstrates that, when compared with the standard SF-36, the SF-36ww modification almost doubled the SRM (our indicator for responsiveness) in the PF domain for all neurological levels (from SRM = 0.36–0.68) and increased the PCS responsiveness by 24% (from SRM = 0.58–0.72). A slight decrease in responsiveness in the modified mental composite score (MCSww) was noted. When the sample was stratified into paraplegic and tetraplegic neurological levels, the least responsive domain was the standard PF domain in the tetraplegic group. With the walk-wheel modification, the responsiveness of this group improved by over five times

(from $SRM = 0.11-0.58$, $n = 77$). In contrast, the responsiveness for the paraplegic group increased by only 12% (from $SRM = 0.77-0.86$, $n = 61$).

We used the SRMs to calculate the sample sizes required to detect a change in health status (as reflected by changes in the PF domain and the PCS and MCS scores) associated with UTI. These sample sizes assume that all subjects will develop a UTI. The results in Table 4 demonstrate the difficulty in detecting disease-state change using the PF domain in tetraplegic persons and the marked improvement that the SF-36ww modification has on sample size ($N = 611$ vs 24). Over all neurological groups, the SF-36ww had a smaller but still marked reduction on sample size ($N = 60$ vs 17).

To determine sample size estimates for a study in which only some of the subjects will develop UTI, it is necessary to divide the numbers in Table 4 by the proportion expected to get UTI. For example, using the above PF and PFww results (all neurological levels) and our 45% UTI rate gives a total standard SF-36 sample size of 133 ($60/0.45$) compared to 38 ($17/0.45$) for the SF-36ww modification.

Discussion

The SF-36 is widely used as a health status measure across many disease groups including SCI.¹⁶ This is despite criticism of the content validity and floor effects of the physical domain of the SF-36.^{7,17} Our SF-36ww modification differs from that of Meyers and Andresen,⁸ Andresen and Meyers¹⁷ and Tate *et al.*⁷ in that it alters only the walking items and ignores modification of the items about climbing stairs. Our justification for this decision is that the climbing tasks are more likely to be affected by the environment, whereas the locomotion items are more likely to depend on transportable devices; that is, if a wheelchair is the main mode of locomotion, the subject is likely to travel with it. This reduces problematic situations where scores on health status scales may alter simply by being away from a suitable environment, such as when people travel on holidays. The SF-36ww modification is simple, contains only one task type and is quick to perform. We acknowledged that, while pragmatic, our solution is not as complete philosophically as that suggested by Andresen and Meyers and Tate *et al.* Further studies to review any additional effect of modifying the stair-climbing variables (3d and e) should be performed to see if this also improves responsiveness.

We found that asking participants both the standard SF-36 and the modified SF-36ww items in sequence (in essence asking the problematic physical questions twice, with modified wording to maintain broader SF-36 compatibility) was less annoying to participants than asking questions about walking in isolation. Participants using the SF-36ww now have a relevant additional response to all of the questions about walking. While these three additional SF-36ww items appeared to enhance the acceptability of the questionnaire, a weakness of this study is that our retrospective analysis of content validity is likely to have

Table 4 Sample size (of patients who develop UTI) necessary to detect changes in health status associated with UTI: comparison of standard SF-36 and modified SF-36ww

| Domain | Overall | Paraplegic | Tetraplegic |
|--------------------|---------|------------|-------------|
| | n = 138 | n = 61 | n = 77 |
| PF ^a | 60 | 13 | 611 |
| PFww ^b | 17 | 11 | 24 |
| PCS ^a | 23 | 21 | 26 |
| PCSww ^b | 15 | 16 | 14 |
| MCS ^a | 16 | 10 | 21 |
| MCSww ^b | 18 | 11 | 27 |

Abbreviations: MCS, mental composite score; PCS, physical composite score; PF, physical function; ww, walk-wheel.

^aStandard SF-36 domains: PF, PCS and MCS.

^bDomains modified with the ww modifiers: PFww, PCSww and MCSww.

underestimated the actual number of participants who experienced problems during its administration. Respondents had to feel strongly enough about a question to complain as this involved recording a comment on a datasheet, with or without assistance. As a result, additional studies are necessary to clarify content validity issues related to the modification. However, overall such minor modifications are likely to be as acceptable and feasible in application as existing standard versions of the SF-36. The additional three questions did not appreciably increase the completion time of the questionnaire.

On the basis of retrospective review of recorded comments, participants completing the SF-36ww questionnaire should have the following additional information made clear in a preface:

- (1) That the wheelchair questions are to be completed by the main mode of wheelchair used by the participants (for example, if the patient uses both an electric and manual wheelchair, they should score the chair they use the most at the time of assessment), and;
- (2) That patients in situations such as complete bed rest should score based on their current restrictions.

Overall, the SF-36ww modification is quick to implement and is attractive in that it goes a good way toward addressing the content validity problems of the standard SF-36 in the SCI population. Including both the standard and the modified versions of the three walking items retains comparability with disease groups external to SCI by the ability to code to standard SF-36 PF, PCS and MCS values as required. This also allows for powerful interpretations of the underlying SF-36 to be maintained such as utility estimates through SF-6D transformation.^{6,18}

Reassuringly, the discriminant validity between tetraplegic and paraplegic subgroups for the SF-36ww was almost identical to that of published validation assessments using the standard SF-36.³ The slightly better self-reported mental health in the tetraplegic vs paraplegic group has been reported previously and reflects the negative weighting given to the PF domain in calculation of the MCS.¹ These

problems were not rectified by the SF-36ww modification, so there is no advantage or disadvantage in the area of internal discriminant validity. Likewise, the internal consistency of the PFww domain, as demonstrated by Cronbach's α , remained similar to that of the PF domain in the standard SF-36.

In addition to improved content validity, the major benefit of the SF-36ww modification over its predecessor when used in SCI populations is the impact on the responsiveness of the health status measure to incident disease states. The standard PF domain scores were poorly responsive and most heavily influenced by the floor effect in the tetraplegic group.³ The standard SF-36 PF domain scores should not be expected to detect change in disease states over time, particularly where a significant proportion of a sample are tetraplegic patients who predominantly utilize a wheelchair for locomotion. The SF-36ww (walk-wheel) modification significantly improved the floor effect in the tetraplegic group, thereby enabling it to be a useful tool to detect within-group clinical change over time.

Accordingly, the SF-36ww health status measure is likely to be useful in studies and clinical management of medical conditions associated with profound physical disability where a significant proportion of the sample are likely to utilize a wheelchair for some or all of their locomotion and where disease-state change is of interest. Further validation studies will be required in populations without SCI, such as those with latter stage neuromuscular disorders.

We have provided a guide to the sample size calculations required in clinical trials and practise where health status change over time is of key concern. The SF-36ww modification demonstrates a clear advantage in study power, particularly for the tetraplegic subgroup. To determine actual sample size estimates from our figures it is necessary to divide the figures in Table 4 by the proportion expected to get a UTI (or other condition). Given the differences demonstrated between paraplegic and tetraplegic populations, if the proportion of each is likely to differ from our sample (55% tetraplegic), it would be necessary to find the sample size for paraplegics and tetraplegics separately and calculate the actual sample size required after estimating the proportion of each likely to be enrolled.

Conclusion

The SF-36ww is a simple modification, which substantially addresses the known problems of acceptability, content validity and floor effects of the standard SF-36 physical domains within populations with SCI while retaining discriminant validity and internal consistency. We demonstrated improved responsiveness for disease-state change within a sample with SCI that will enhance the power of future studies to assess the effect of disease progression, treatment and prevention. The application of the SF-36ww to other populations with profound physical disabilities warrants investigation.

Acknowledgements

This Study was sponsored by New South Wales Motor Accidents Authority (NSW MAA).

SF-36 Health Survey (c) 1988, 2002 by Medical Outcomes Trust and QualityMetric Incorporated—All rights reserved. SF-36 is a registered trademark of the Medical Outcomes Trust.

References

- Ware JE, Snow KK, Kosinski M, Gandek B. *SF-36 Health Survey—Manual and Interpretation Guide*. The Health Institute, New England Medical Centre Boston: Boston, MA, 1993.
- Haran MJ, Lee BB, King MT, Marial O, Stockler MR. Health status rated with the medical outcomes study 36-item short-form health survey after spinal cord injury. *Arch Phys Med Rehabil* 2005; **86**: 2290–2295.
- Haran MJ, King MT, Stockler MR, Marial O, Lee BB. Validity of the SF-36 Health Survey as an outcome measure for trials in people with spinal injury. *CHERE Working Paper 2007/4, CHERE, Sydney* 2007. Available at http://www.chere.uts.edu.au/pdf/wp2007_4.pdf.
- Andresen EM, Fouts BS, Romeis JC, Brownson CA. Performance of health-related quality-of-life instruments in a spinal cord injured population. *Arch Phys Med Rehabil* 1999; **80**: 877–884.
- Dijkers M. Quality of life of individuals with spinal cord injury: a review of conceptualisation, measurement, and research findings. *J Rehabil Res Dev* 2005; **42**: 87–110.
- Lee BB, King MT, Simpson JM, Haran MJ, Stockler MR, Marial O *et al*. Validity, Responsiveness and Minimal Important Difference for the SF-6D Health Rating Scale in a Spinal Cord Injured Population. Value in Health (accepted 1 August 2007, published online 11 January 2008, doi:10.1111/j.1524-4733.2007.00311.x.).
- Tate DG, Kalpakjian CZ, Forchheimer MB. Quality of life issues in individuals with spinal cord injury. *Arch Phys Med Rehabil* 2002; **83**: S18–S25.
- Meyers AR, Andresen EM. Enabling our instruments: accommodation, universal design, and access to participation in research. *Arch Phys Med Rehabil* 2000; **81**: S5–S9.
- Lee BB, Haran MJ, Hunt LM, Simpson JM, Marial O, Rutkowski SB *et al*. Spinal-injured neuropathic bladder antiseptis (SINBA) trial. *Spinal Cord* 2007; **45**: 542–550.
- Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989; **27**: S178–S189.
- Bland JM, Altman DG. Statistics notes: Cronbach's alpha. *BMJ* 1997; **314**: 572.
- Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; **16**: 297–334.
- Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol* 2000; **53**: 459–468.
- Wittes J. Sample size calculations for randomised controlled trials. *Epidemiol Rev* 2002; **24**: 39–53.
- Walters SJ. Sample size and power estimation for studies with health related quality of life outcomes: a comparison of four methods using the SF-36. *Health Qual Life Outcomes* 2004; **2**: <http://www.hqlo.com/content/2/1/26>.
- Wood-Dauphinee S, Exner G, Bostanci B, Exner G, Glass C, Jochheim KA *et al*. Quality of life in patients with spinal cord injury—basic issues, assessment, and recommendations. *Restor Neurol Neurosci* 2002; **20**: 135–149.
- Andresen EM, Meyers AR. Health-related quality of life outcomes measures. *Arch Phys Med Rehabil* 2000; **81**: S30–S45.
- Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the sf-36. *J Health Econ* 2002; **21**: 271–292.
- Maynard Jr FM, Bracken MB, Creasey G, Ditunno Jr JF, Donovan WH, Ducker TB *et al*. International standards for neurological and functional classification of spinal cord injury. American spinal injury association. *Spinal Cord* 1997; **35**: 266–274.