

Review

Short Form health surveys and related variants in spinal cord injury research: A systematic review

David G. T. Whitehurst^{1-3*}, Lidia Engel^{2,4}, Stirling Bryan^{2,5}

¹Faculty of Health Sciences, Simon Fraser University, Burnaby, British Columbia, Canada, ²Centre for Clinical Epidemiology and Evaluation, Vancouver Coastal Health Research Institute, Vancouver, British Columbia, Canada, ³International Collaboration on Repair Discoveries (ICORD), Vancouver, British Columbia, Canada, ⁴Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, Limburg, The Netherlands, ⁵School of Population and Public Health, University of British Columbia, Vancouver, British Columbia, Canada

Context: 'Short Form' health surveys – such as the SF-36 and SF-12 – are widely used in medical research. Spinal cord injury (SCI) is no exception, despite oft-cited concerns regarding measurement properties for populations with physical impairment.

Objective: To provide a comprehensive overview of the use of Short Form health surveys and their variants within the SCI literature.

Methods: Papers published between database inception and September 2012 were identified from 11 electronic databases; a supplementary reference list search was also conducted. Data extraction focused on details regarding the range of different Short Form surveys and variants used in SCI research, the respective frequency of use, the nature of reporting (complete versus partial reporting) and the method of survey administration.

Results: One hundred seventy-four papers were identified. Thirty-six-item Short Form health surveys were frequently administered as complete instruments ($n = 82$); in 69 of these 82 studies (84%), it was not clearly stated which 36-item version had been used (e.g. SF-36v1, SF-36v2, RAND-36). Data for individual items and domains were often reported (29% of identified studies), indicating significant partial use of standardized measures. Modified variants of standardized health surveys were administered in 12 studies.

Conclusion: Although standardized Short Form health surveys are common within SCI research, attempts to add, delete, or modify items have resulted in a number of variants, often with minimal supportive psychometric evidence. Using established, generic outcome measures is appealing for a number of reasons. However, validity is paramount and requires further explicit consideration within the SCI research community.

Keywords: Quality of life, Health surveys, Spinal cord injuries, Health services research

Introduction

There is growing interest in better understanding physical, psychological, and social well-being factors that affect the quality of life of individuals with spinal cord injury (SCI).¹⁻³ Two driving factors behind this interest are the continued pursuit of progress in SCI medical and

rehabilitative care, and the need for accountability in light of ever-increasing demands on limited health care resources. Quality-of-life measurement encompasses numerous distinctions and approaches to conceptualization, such as subjective and objective measurement, health-related and non-health-related considerations, condition-specific and generic outcomes, impairment versus participation, etc.¹ The importance of health-related quality-of-life (HRQoL) measurement is based on the need for metrics on performance and processes in health care, acknowledging that the perspectives of

*Department and institution to which the work can be attributed.
Correspondence to: David G. T. Whitehurst, Faculty of Health Sciences, Blusson Hall 10504, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, V5A 1S6, Canada.
Email: david_whitehurst@sfu.ca

consumers of health care resources (i.e. patients and individuals living with long-term conditions) are highly relevant to efforts to improve the quality and effectiveness of care. Accordingly, HRQoL outcome measures are often fundamental components of surveys, registries, and randomized controlled trials.

With regard to the distinction between condition-specific and generic outcomes, each measurement approach has respective strengths and weaknesses depending on the underlying research question and analytic objective.^{3,4} Generic measures are purposely designed for use across the complete spectrum of diseases, for all types of health care interventions. Condition-specific measures are considered to be more sensitive to subtle changes in health because of the focus on dimensions of health specifically tailored to the clinical area of interest. An example within the context of SCI is the Qualiveen, an outcome that provides a perspective on quality of life in SCI for urinary disorders.⁵ A disadvantage of using condition-specific measures is the restricted generalizability, both within clinical areas (there may be multiple condition-specific measures to choose from) and across conditions.

The 'Family' of Short Form health surveys

One of the most widely used generic HRQoL instruments in SCI research is the Medical Outcomes Study 36-item Short Form Health Survey, more commonly known as the SF-36.^{6,7} This outcome measure comprises 36 items across eight health domains: physical functioning, role limitations caused by physical health problems, bodily pain, social functioning, mental health, role limitations caused by emotional problems, vitality, and general health perceptions.^{2,6} Scores can be derived for each of the eight health domains, as well as psychometrically based physical component summary (PCS) and mental component summary (MCS) scores. The measurement tool has undergone extensive psychometric validation across diverse patient groups,⁸ including SCI.^{3,9–11}

A number of different Short Form instruments exist, such as the 12-item SF-12 and two generic health outcome surveys developed as part of the Veterans Health Study (the Veterans RAND 12-Item Health Survey (VR-12) and the Veterans RAND 36-Item Health Survey (VR-36)).^{12–14} Furthermore, 8-item (SF-8) and 20-item (SF-20) health surveys have been developed, although the use of these instruments in the general medical literature is markedly less frequent than the 36-item and 12-item surveys.^{15,16} For the purpose of conducting economic evaluations within a cost-utility framework using Short Form health survey

data, scoring algorithms are available to convert SF-36, SF-12, VR-12, and VR-36 data into health state valuations (or 'utility' scores).^{17–19} It is important to note that 'Short Form', when used in this paper, refers only to health surveys that can be traced back to the Medical Outcomes Study and, therefore, does not refer to any and all short form instruments in the health outcomes literature. Also, 'health survey' is a term that has become synonymous with outcome measures emerging from the Medical Outcome Study. Although 'survey' is often used in research to describe a questionnaire that comprises multiple outcome measures, in this paper, from this point, the term refers to a HRQoL instrument.

Formal amendments to standardized outcome measures are consequences of gaining experience regarding the merits of an instrument. Accordingly, it is common for new versions of outcome measures – both generic and condition-specific – to be developed over time,²⁰ such as the SF-36 version 1 (SF-36v1) and SF-36 version 2 (SF-36v2), and the SF-12 version 1 (SF-12v1) and SF-12 version 2 (SF-12v2). An additional layer of variation regarding Short Form measures relates to the availability of different scoring procedures.²¹ It is beyond the scope of the current paper to fully describe the differences between the 36-item (i.e. SF-36v1, SF-36v2, RAND-36, VR-36) and 12-item (i.e. SF-12v1, SF-12v2, RAND-12, VR-12) instruments.

For simplicity in this paper, we use the collective terms 'SF-36' and 'SF-12' to refer to the various 36-item and 12-item versions of Short Form health surveys, respectively. The VR-36 and VR-12 instruments are not included in these collective terms. Where there is a substantive issue raised, distinctions across different versions are made.

Objective and evaluative framework

The topic of SCI and quality of life has been the subject of numerous reviews, with objectives including the identification and critical review of commonly used measures,^{3,7,20,22,23} and comparison of data with non-SCI populations.²⁴ The objective of the current review is to present a systematic exploration and description of how Short Form health surveys have been used in peer-reviewed SCI studies, providing a unique contribution to the quality-of-life literature. Extending beyond establishing the frequency with which standardized health surveys have been used, the evaluative framework of the review addresses two concerns that have been raised by the SCI research community, namely (i) non-standardized use of Short Form health surveys and (ii) survey administration methodology. Prior to the Methods section, which provides specific details of

the analytic considerations and data extraction processes, the remainder of this section explores these two concerns.

Despite the widespread use of 36-item Short Form health surveys in studies of SCI populations, and the existence of SCI-specific validation studies,^{3,9-11} the outcome measure is not without problems. In particular, concerns have been raised within the SCI and broader disability literature about the appropriateness of the physical functioning subscale for individuals with significant physical impairment. For example, items referring to an individual's capacity to walk or climb stairs may be deemed, at best, irrelevant, or, at worst, insulting,^{9,25} which has led some commentators to suggest that the SF-36 may assess a person's gross motor deficits, rather than functional health, due to the failure to document rehabilitation measures and adaptive behaviour.²⁶ More generally, the content validity of multiple SF-36 domains (physical functioning, role-physical, social functioning, and role-emotional) has been shown to be compromised in SCI due to the irrelevance of items and/or response options.^{27,28}

In response to the perceived shortcomings of the SF-36 for SCI populations, a number of different modifications have been tested, such as changing 'walk' and 'climb' to 'go',²⁹ and changing 'walk' to 'wheel'.²⁸ It is important to acknowledge that, in general, amending standardized health surveys is discouraged because the resultant data is no longer associated with established psychometric evidence; summary scores are, effectively, invalid. Any amendment, however trivial it may seem, would require the modified instrument to be subjected to a new evaluation of measurement properties.²⁹ Accordingly, this review seeks to provide a comprehensive listing of modifications that have been used in research settings.

The second indication of concern – methods of survey administration – stems from the common observation that individuals report higher levels of quality of life for interviewer-administered modes of data collection compared to self-administered modes. This finding has been observed in disease-specific groups, including SCI, as well as generic groups.²⁹⁻³² Given the likelihood that systematic differences in HRQoL could be attributable to the method of data collection alone, an explicit objective of the current review is to establish the state-of-play for alternative modes of survey administration in the context of SCI.

Due to fundamental differences in measurement objectives between profile measures of HRQoL (such as the SF-36 and SF-12) and utility instruments (e.g. SF-6D), the review focuses solely on outcome measures

that yield health profiles. The use of utility instruments in SCI research has been the topic of a recent systematic review.³³

Methods

Data sources and search strategy

A systematic search was undertaken for articles published from database inception to September 2012 in the following databases: Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Methodology Register (CMR), Cumulative Index to Nursing and Allied Health Literature (CINAHL), Excerpta Medica Database (EMBASE), Health Technology Assessment (HTA; part of the Cochrane Library), Health and Psychosocial Instruments (HaPI), Medline, National Health Service Economic Evaluation Database (NHSEED), PsycINFO, PubMed, and SPORTDiscus. The search strategy comprised SCI-specific search terms (e.g. paraplegia, tetraplegia, and quadriplegia) and the names and abbreviations of outcome measures within the suite of Short Form instruments. The database search strategy is provided in Appendix A; the strategy follows the approaches utilized in other SCI-related quality-of-life reviews.^{3,33} To enhance the comprehensiveness of the search strategy, we conducted a bibliographic search of reference lists for all database-identified papers included in the review. From the reference list searches, papers deemed potentially to be relevant, based on title alone, were retrieved and subjected to the study selection process.

Selection criteria

A two-stage approach was used to identify eligible papers. Within the first stage, titles and abstracts of database-identified papers were screened to establish likely relevance to a study population of individuals with SCI and the presence, or possible presence, of a Short Form health survey instrument. As this description suggests, stage one comprised a liberal application of inclusion criterion, primarily focused on the identification of definite exclusions. For example, reference to 'quality of life' as an outcome measure, without specifying an instrument, was sufficient for inclusion at stage one. Requirements for papers to be written in English and published in a peer-reviewed journal were applied at this stage. Full text papers were obtained for all stage one inclusions.

The purpose of the second stage was to verify the key analytic features of the review; namely, that a Short Form health survey had been administered to a study population including individuals with SCI. Inclusion

did not require studies to report summary statistics for complete instruments, such as summary scores for the SF-36, because partial use of instruments was a specific analytic consideration (see below). Neither were studies required to include individuals with SCI only. However, for studies that considered broader patient populations (e.g. trauma or injury), inclusion for this review required explicit reporting of relevant data for the respective SCI subgroup.

Both stages of the study selection process were performed, independently, by two of the authors (D.G.T.W. and L.E.); a colleague provided assistance with regard to clinical queries. Disagreements between reviewers were resolved through discussion. Reasons for exclusion were documented at each stage. The bibliographic search was conducted by a single author only (L.E.).

Data extraction

Data extraction focused on the extent to which standardized Short Form health surveys and their variants have been used in published SCI studies. In addition to identifying the range of measures used in SCI research, specific considerations were also (i) to determine the frequency of use for different Short Form surveys and variants, (ii) to establish the degree to which instruments have been administered and/or reported as partial measures (reflecting the selective use of items or domains rather than a complete survey instrument), and (iii) to explore variations in survey administration methodology. A single author (L.E.) extracted information from the included studies into a data extraction form specifically designed for this review. All data extraction queries regarding the identification or classification of information pertinent to the review were dealt with through discussion among the study team.

Results

A total of 463 studies were identified from database ($n = 456$) and reference list ($n = 7$) searches; 174 met the two-stage selection criteria. A breakdown of the review process, including reasons for exclusion at each stage, is reported as a flow chart in Appendix B. A complete reference list is available from the corresponding author on request.

'Complete' and 'Partial' use of standardized instruments

Across the 174 studies, three different standardized Short Form instruments were identified; SF-36 (the collective term), SF-12 (the collective term), and VR-36. There were no studies reporting use of the SF-20, SF-

8, or VR-12. Summary data for complete, standardized instruments were reported in 111 (64%) studies. In addition to 12 (7%) studies that had adopted modified variants (discussed later), evidence of non-standardized use due to the partial selection of items was observed in the remaining 51 (29%) papers (Table 1).

Table 2 provides a breakdown of 'partial use' categories for the 51 studies where partial selection of items was identified. The most commonly reported 'partial use' domains were mental health ($n = 7$) and bodily pain ($n = 7$). Isolated use of questions regarding

Table 1 Classification of Short Form instrument use within the identified studies*

Classification	Frequency (%)
'SF-36' (collective term)	82
SF-36v1	7
SF-36v2	5
RAND-36	1
Ambiguous/not stated	69
'SF-12' (collective term)	26
SF-12v1	1
Ambiguous/not stated	25
VR-36	3
Partial use	51
Modifications	12

*The five categories ('SF-36', 'SF-12', VR-36, Partial use, and Modifications) are mutually exclusive and collectively exhaustive with regard to the 174 identified studies. For each study included in the review, determination of the appropriate category was based solely on information provided in the body-text of the manuscript. Using the reported search strategy and inclusion criteria, there were no studies that had used the VR-12, SF-8, or SF-20.

Table 2 Breakdown of the partial use of individual items and domains of standardized Short Form health surveys ($n = 51$)

Description of partial use	SF-36	SF-12
Bodily pain domain	4	3
General health domain	1	–
Mental health domain	7	–
Physical functioning domain	1	–
Vitality domain	1	–
Item 1*	7	1
Items 1 and 2 of the 36-item health survey**	4	n/a
Physical component summary score	2	–
Deletion of items or domains	17	2
Deletion and addition of items	–	1
Total	44	7

*Item 1 (self-perceived general health; 'In general, would you say your health is:') is identical in the 36-item and 12-item Short Form health surveys.

**The self-perceived general health transition question ('Compared to one year ago, how would you rate your health in general now?') is only present in 36-item Short Form health surveys.

self-perceived general health (item 1 of the SF-36 and SF-12) and general health transition (item 2 of the SF-36) was also observed on multiple occasions ($n = 12$). Two studies explored either the PCS or MCS only, and administered only those questions which were required for creating the respective summary component scores.^{34,35}

Modifications to standardized Short Form health surveys

Standardized outcome measures had been modified (items amended/added and included in the conventional scoring algorithm) in 12 (7%) studies; in all instances, modifications relate to the physical functioning subscale of the respective instrument.^{28,36–46} Details of

Table 3 Details of modifications made to standardized Short Form health surveys

Modification (with reference(s))	Description
'Enabled' SF-36 (Froehlich-Grobe <i>et al.</i> , 2008 ³⁶ ; Nanda <i>et al.</i> , 2003 ³⁷ ; Rowell and Connelly, 2008 ³⁸ ; Rowell and Connelly, 2010 ³⁹ ; Unalan <i>et al.</i> , 2007 ⁴⁰)	Froehlich-Grobe and colleagues used the enabled physical functioning items, as proposed by Meyers and Andresen, and also made additional changes. The instructions for the physical functioning section were modified in a manner that guided respondents to report on their physical functioning while using assistive devices; ' <i>The following items are about activities you might do during a typical day using your normal assistive devices (wheelchair/cane/prosthetic).</i> ' A final modification changed the activities listed under item 3 (vigorous activities) Nanda and colleagues administered the enabled version as described by Meyers and Andresen Rowell and Connelly (2008, 2010), ^{38,39} and Unalan <i>et al.</i> (2007) ⁴⁰ administered the conventional SF-36 and the five enabled physical functioning items. No changes were made to the ordering of items to reflect increasing distances
SF-36 Walk-Wheel (Lee <i>et al.</i> , 2009 ²⁸ ; De Wolf <i>et al.</i> , 2012 ⁴¹)	Lee and colleagues developed the SF-36 Walk-Wheel (SF-36WW), a health survey administered as a 39-item instrument, comprising the 36 standardized items and three additional questions. The three additional items are exactly the same as items 9, 10 and 11, except with the replacement of the word 'walk' with the word 'wheel' De Wolf and colleagues used the SF-36WW in their psychometric evaluation of the World Health Organization Disability Assessment Scale II (WHO-DAS II) in the context of SCI
Modification 3 (Dudley-Javoroski and Shields, 2006 ⁴² ; Mueller <i>et al.</i> , 2012 ⁴³)	Dudley-Javoroski and Shields adapted both physical functioning items to ' <i>improve the sensitivity and appropriateness of the SF-12 for a population with complete SCI</i> '; the standard 12 items and the amended two items were administered (i.e. 14 items in total). The purpose of these changes was to gauge respondents' ability to perform wheelchair-specific functional tasks. Both changes related to the description of activities; ' <i>...moving a table, pushing a vacuum cleaner, bowling, or playing golf</i> ' was changed to ' <i>...using your wheelchair around your home</i> ', and ' <i>Climbing several flights of stairs</i> ' was changed to ' <i>Going rapidly in your wheelchair for several blocks</i> ' Mueller and colleagues used the same amended version in a randomized controlled trial
Modification 4 (Litchke <i>et al.</i> , 2012 ⁴⁴)	Litchke and colleagues amended text in the physical functioning items of the SF-36; these changes were different to those of Dudley-Javoroski and Shields. Although the exact changes are not discernible from the paper alone, the authors describe them as, ' <i>For example 'climbing stairs or walking more than a mile' was changed to 'propelling up a steep ramp or pushing more than a mile.'</i> '
Modification 5 (Luther <i>et al.</i> , 2006 ⁴⁵)	The objective of this study was to revise items of the VR-36 physical functioning scale to be more appropriate for use with people with SCI. The result was a set of 8 items that were considerably different from the original items
Modification 6 (Kalpakjian <i>et al.</i> , 2007 ⁴⁶)	Text for both items of the SF-12 physical functioning scale were amended; ' <i>...moving a table, pushing a vacuum cleaner, bowling, or playing golf</i> ' was changed to ' <i>...pushing a vacuum cleaner, or climbing 1 flight of stairs/ramp</i> ', and ' <i>Climbing several flights of stairs</i> ' was changed to ' <i>Climbing several flights of stairs/ramps</i> '

modifications in each of these studies are provided in Table 3. The most commonly identified modified Short Form health survey was an ‘enabled’ version of the SF-36, developed by Meyers and Andresen,²⁹ which has been used in a total of five studies. Another notable modification is the SF-36 Walk-Wheel (SF-36WW), which comprises the addition of three questions to the standard SF-36 (i.e. 39 items in total) and where the word ‘walk’ is exchanged for the word ‘wheel’ in items 9, 10, and 11.²⁸

In the enabled version of the SF-36 developed by Meyers and Andresen²⁹ – sometimes referred to as the SF-36E – the word ‘go’ is substituted for ‘walk’ (items 9, 10, and 11) and ‘climb’ (items 6 and 7). Furthermore, the question ordering is different: questions about shorter distances are asked prior to questions about longer distances and, subsequently, the longer distance questions are conditional on the shorter distance responses. For example, if someone does not accomplish an easier mobility question (i.e. one hundred metres, or one flight of stairs), the respondent would skip the rest of the series that asks about more demanding activities (e.g. several hundred metres, or several flights of stairs). The description of Short Form health survey modifications provided in Table 3 highlights the regularity with which a second level of study-specific amendments occurs, even when a modified survey has already been selected. Of five identified studies that purport to use the enabled version of the SF-36, only one study – Nanda *et al.*³⁷ – appears to administer it in the exact manner prescribed by Meyers and Andresen.

Although not a modification according to our definition, an alternative use of the standard SF-36 by Tate *et al.* is noteworthy.⁴⁷ In addition to the eight standard subscales of the SF-36, the authors derived a ‘Mobility’ subscale from items 9–11 of the SF-36 plus three new items that asked similar questions about mobility with the use of a wheelchair. The purpose of the new subscale was to measure the degree of an individual’s mobility and ambulation; it provided supplementary data to the standard SF-36 (rather than a modification).

Methods of survey administration

A range of administration modes were used to capture Short Form health survey responses. The 174 papers can be categorized into four mutually exclusive administration method groups: self-administration (the individual living with SCI completes the health survey on his/her own; $n = 42$), interview-based ($n = 41$), a

combination of multiple methods (including proxy-completion or assisted-completion to accommodate individuals who were unable to provide responses on their own; $n = 37$), or unclear ($n = 54$). In the studies where data collection comprised a combination of alternative methods, examples include self-administered postal or on-line surveys in addition to telephone interviews,⁴⁸ or a mixture of face-to-face and telephone interviews (i.e. no single approach).^{49,50}

Discussion

Key findings

The unique contribution of the current study is to provide the first systematic summary of how Short Form quality-of-life measures have been administered and reported in SCI research. The research track-record of the SF-36 provides it ‘gold standard’ status in many clinical areas,⁵¹ and so it is not surprising that this review identified the 36-item surveys as the most frequently used Short Form HRQoL measures in SCI research. However, findings from this review illustrate a number of interesting and important issues pertinent to the SCI research community. The starkest observation was the frequency of non-standard use or adaptation of the standardized instruments when used in SCI clinical research studies. This finding manifested itself in multiple ways, such as the regularity of selective use (or selective reporting) of individual items or domains, and the addition, deletion, or modification of items in an attempt to make the instrument better suited to the characteristics of study participants.

A further observation was the regularity with which the reporting of survey administration methods was ambiguous or unclear – this was an issue in 31% of identified studies. This particular analytic focus originated from the established phenomenon that the mode of administration can affect HRQoL outcomes. If there is potential for systematic differences between Short Form health survey responses attributable to the data collection method, it is imperative that the mode of administration is clearly reported. Undoubtedly, tight word restrictions in some peer-reviewed journals require authors to give careful consideration to the content of their manuscript. However, for any empirical study, the integrity of the data is vital and a clear explanation of the data collection process is a necessity.

Selective use and modifications

Concerns about the appropriateness of standardized Short Form health surveys for individuals with SCI

are well-founded. Unsurprisingly, items within the physical functioning domain are of particular concern; asking questions about a wheelchair user's ability to walk several hundred metres or climb several flights of stairs are of questionable merit as components of a physical functioning health domain. Faced with this scenario, a researcher has options: (i) use a different outcome measure that does not exhibit the same problems (e.g. a SCI-specific instrument), (ii) be selective in the administration of items from a standardized instrument, or (iii) modify an existing generic measure in a manner that appears to improve the face validity and content validity for the SCI context. The latter two options may be appealing because of the perceived advantages of generic outcome measurement, i.e. the ability to use the same outcome, or domains of an outcome, across all clinical areas.

Examples of case-specific selective use of items/domains can be seen in the studies by Tan *et al.* (SF-12 plus the pain subscale of the SF-36),⁵² and Hultling *et al.* (10 items from the SF-12, plus the health transition question from the SF-36).⁵³ Tan and colleagues' decision to use the SF-12 may have been because of concerns regarding the SF-36 physical functioning scale – a decision that has been highlighted by other SCI researchers.⁵⁴ However, the SF-36 pain scale (two items) was preferred to the SF-12 pain scale (one item) and, therefore, the additional SF-36 pain item was included in their study survey. In the second case, Hultling *et al.* stated that items 3 and 4 of the SF-12, the two items of the SF-12 physical functioning scale, 'were not applicable to patients with SCI and therefore not completed', but provided no further justification. Other researchers have made the decision to drop the SF-36 physical functioning scale because of supposed irrelevance, preferring to administer only seven of the eight subscales.^{55–57} Herein lies a problem; it is not the mere fact that an outcome is generic that permits comparability across studies (and conditions); the outcome measure needs to be generic and valid.

One-off modifications of standardized health surveys are inappropriate if undertaken without any assessment of measurement properties – an issue acknowledged by some authors who have used modified instruments.^{28,29,45} Even if a modified health survey demonstrates good psychometric performance in a particular setting, the value of such a measure is negligible if it is not used in subsequent studies and does not undergo further validation testing. Similarly, evidence for the validity of individual domains or items within SCI

populations is necessary to justify the chosen analytic approach. Examples of such work include the examinations of the mental health subscale and the pain interference item.^{58,59} Unquestionably, the rigour of a study is enhanced if the outcome measure(s) being used has evidence of strong psychometric properties.

Areas for further research

It is important to acknowledge that the extensive evidence base to support the psychometric properties of the standard versions of the SF-12 or SF-36 cannot be assumed to apply for the validity of modified variants or the partial use of items. The SF-36 has been subjected to SCI-related modifications since 1997 and a number of 'modifications of modifications' have followed. Accordingly, it seems timely to formulate a consensus position about the role of Short Form health surveys for SCI research, and clarify a research agenda for modified variants and comprehensive psychometric evaluation.

Another area of study that would provide a highly valuable contribution to the SCI research community is in relation to data collection methodology. Logistical and financial considerations ensure that a trade-off between pragmatism and data quality is common in health research,⁶⁰ and so we more often see self-administered rather than interviewer-led surveys. Reflecting on the widespread use of the SF-36 despite concerns regarding the physical functioning scale, there are important questions that have received very little attention in SCI research. For example, if an on-line or interviewer-led survey is used to collect outcome data, should respondents be able to leave items blank if they choose to? If not, what are the consequences of requiring an individual to respond to an item that they do not relate to? Also, given the subjective nature of many items in Short Form HRQoL instruments, there is a need to better understand the appropriateness of alternative sources of proxy information. At a time when innovative data collection methods are being explored for health research,⁶¹ it would be judicious to consider the implications of alternative approaches for the administration of health surveys, with the long-term goal of devising authoritative guidelines for SCI research.

Study limitations

Search strategies to identify the use of specific outcome measures (or families of outcome measures) have potential for relevant articles to be overlooked for a number of reasons. For example, the subjective reporting styles of

authors may mean that generic terms are used in an abstract instead of naming the instrument(s), or HRQoL could be a secondary outcome that isn't deemed to be of sufficient importance for the abstract. It is also acknowledged that systematic reviews regarding HRQoL instruments can be problematic due to the absence of formal standards for indexing 'outcome measure keywords' in databases. Improving the quality of abstract content to aid retrieval and screening of relevant papers has been highlighted in the SCI literature.⁶² In this study, a comprehensive search strategy and a two-stage approach to identifying eligible papers were used to address these potential problems.

A second potential limitation relates to the nature of the evaluative framework. Throughout all analyses, working assumptions were that all necessary data (i.e. the outcome measure, changes to standardized layouts, and the mode of administration) were to be extracted from the manuscripts only. We did not contact authors to get supplemental information. Accordingly, the use of Short Form health surveys described in this review is contingent on the information reported in the identified studies. We do not believe this limitation compromises the broad implications of our observations; after all, the value of research is primarily determined by the manner in which results are reported.

Conclusion

This review has highlighted the evident trade-off that researchers face when using a generic outcome measure that poorly aligns with characteristics of a clinical condition. Researchers may decide to include the standardized measure (and accept the pitfalls), select only certain items and domains (i.e. partial use of an instrument), or make modifications that appear to better reflect the SCI-context. In the case of modified versions, there is an apparent lack of direction and commitment to a single approach in SCI. For any outcome measure, validity is paramount. Given that both use and scepticism of Short Form health surveys are widespread among SCI researchers, this review provides a valuable basis from which to consider future research priorities.

Acknowledgements

We would like to thank the two anonymous reviewers for their helpful comments, and Dr Vanessa Noonan

(Rick Hansen Institute, Canada) for her clinical support during the screening of identified papers.

No specific funding has supported this work and none of the authors report a conflict of interest.

Appendix A: EMBASE search strategy, searched via OvidSP^a

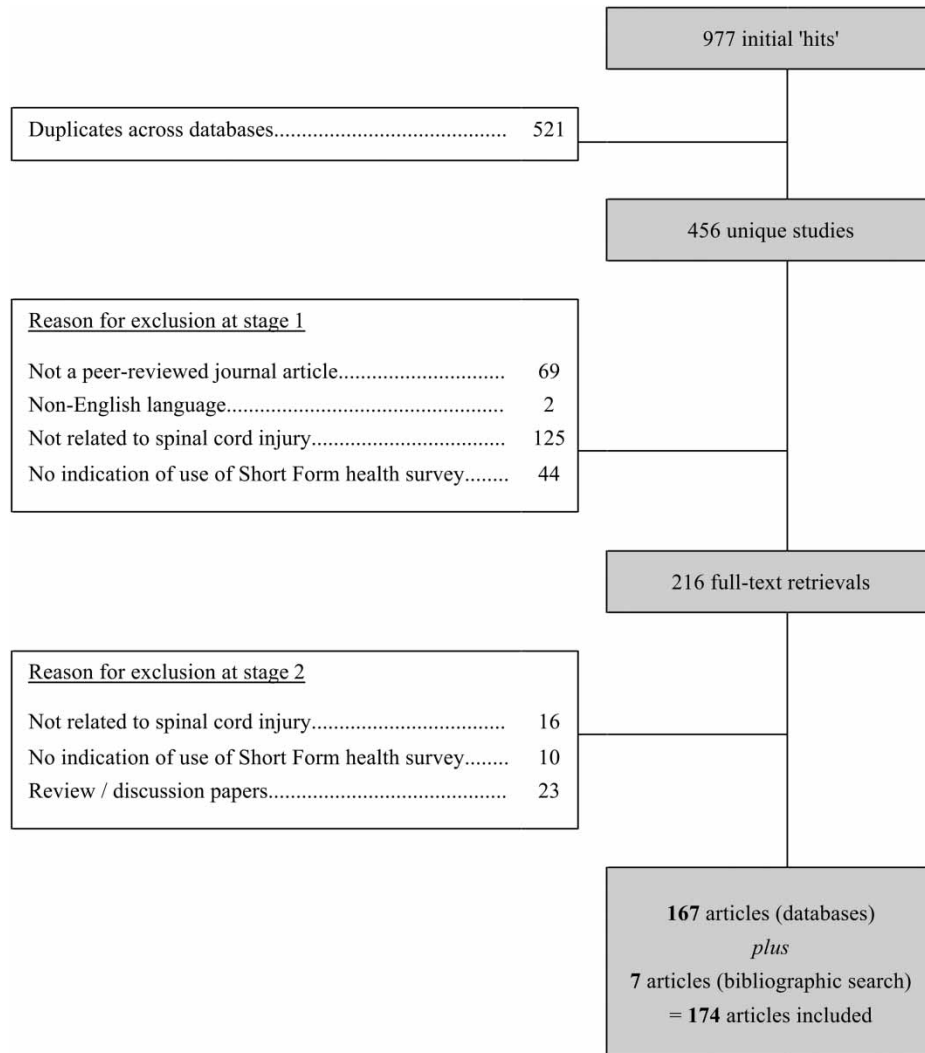
Search term

- 1 Spinal Cord Injury/
 - 2 Spinal Cord Injur*.mp.
 - 3 SCI.mp.
 - 4 Paraplegia/
 - 5 Parapleg*.mp.
 - 6 Quadriplegia/
 - 7 Quadripleg*.mp.
 - 8 Tetrapleg*.mp.
 - 9 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
 - 10 Short Form 36/
 - 11 SF36*.mp.
 - 12 SF-36*.mp.
 - 13 MOS36*.mp.
 - 14 MOS-36*.mp.
 - 15 RAND-36.mp.
 - 16 RAND36.mp.
 - 17 VR36.mp.
 - 18 VR-36.mp.
 - 19 Vetera* RAND*.mp.
 - 20 Vetera* SF*.mp.
 - 21 Short Form 12/
 - 22 SF-12*.mp.
 - 23 SF12*.mp.
 - 24 MOS-12*.mp.
 - 25 MOS12*.mp.
 - 26 RAND-12.mp.
 - 27 RAND12.mp.
 - 28 VR-12.mp.
 - 29 VR12.mp.
 - 30 SF-6D.mp.^b
 - 31 SF6D.mp.^b
 - 32 SF-8.mp.
 - 33 SF8.mp.
 - 34 SF-20.mp.
 - 35 SF20.mp.
 - 36 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20
or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30
or 31 or 32 or 33 or 34 or 35
 - 37 9 and 36
 - 38 Limit 37 to English language
-

^aThis strategy was modified for the requirements of each respective database. An asterisk (*) indicates the use of the truncation facility to account for variations of the same term, e.g. 'injury' and 'injuries'. Hyphenated search terms (for example, SF-36) also retrieve non-hyphenated terms (for example, SF 36). The abbreviation 'mp.' indicates a keyword search term within OvidSP, while the forward slash '/' represents a subject heading search term.

^bSearch terms for the SF-6D were included in the interests of completeness, due to the inextricable link to Short Form health surveys.

Appendix B: Flow chart of the review process, including reasons for exclusion at each study selection stage^a



^aThe same reasons for exclusion appear in stages 1 and 2 because of the different levels of scrutiny.

References

- 1 Dijkers MPJM. Quality of life of individuals with spinal cord injury: a review of conceptualization, measurement, and research findings. *J Rehabil Res Dev* 2005;42(3 Suppl 1):87–110.
- 2 Hays RD, Hahn H, Marshall G. Use of the SF-36 and other health-related quality of life measures to assess persons with disabilities. *Arch Phys Med Rehabil* 2002;83(12 Suppl 2):S4–9.
- 3 Hill MR, Noonan VK, Sakakibara BM, Miller WC. Quality of life instruments and definitions in individuals with spinal cord injury: a systematic review. *Spinal Cord* 2010;48(6):438–50.
- 4 Guyatt GH, Feeny DH, Patrick DL. Measuring health related quality of life. *Ann Internal Med* 1993;118(8):622–9.
- 5 Costa P, Perrouin-Verbe B, Colvez A, Didier J, Marquis P, Marrel A, et al. Quality of life in spinal cord injury patients with urinary difficulties. Development and validation of qualiveen. *Eur Urol* 2001;39:107–13.
- 6 Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30(6):473–83.
- 7 Wilson JR, Fehlings MG. Spinal cord injury and quality of life: a systematic review of outcome measures. *Evid Based Spine Care J* 2011;2(1):37–44.
- 8 McHorney CA, Ware JE, Jr, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994;32(1):40–66.
- 9 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;80:877–84.
- 10 Forchheimer M, McAweeney M, Tate DG. Use of the SF-36 among persons with spinal cord injury. *Am J Phys Med Rehabil* 2004;83:390–5.
- 11 Lin MR, Hwang HF, Chen CY, Chiu WT. Comparisons of the brief form of the World Health Organization Quality of Life and Short Form-36 for persons with spinal cord injuries. *Am J Phys Med Rehabil* 2007;86:104–13.
- 12 Iqbal SU, Rogers W, Selim A, Qian S, Lee A, Ren XS, et al. The Veterans RAND 12-Item Health Survey: what it is and how it is used [accessed 2013 August 17]. Available from: <http://www.va.gov/chqoer/VR12.pdf>.

- 13 Kazis LE, Miller DR, Skinner KM, Lee A, Ren XS, Clark JA, *et al.* Applications of methodologies of the Veterans Health Study in the VA healthcare system: conclusions and summary. *J Ambul Care Manage* 2006;29(2):182–8.
- 14 Ware J, Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34(3):220–33.
- 15 Ware JE, Kosinski M, Dewey J, Gandek B. How to score and interpret single-item health status measures: a manual for users of the SF-8 health survey. Boston: QualityMetric; 2001.
- 16 Ware JE, Sherbourne CD, Davies AR. Developing and testing the MOS 20-item short-form health survey: a general population application. In: Stewart AL, Ware JE, (eds.) *Measuring functioning and well-being: the Medical Outcomes Study approach*. Durham, NC: Duke University Press; 1992. p. 277–90.
- 17 Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;21: 271–92.
- 18 Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Med Care* 2004;42:851–9.
- 19 Selim AJ, Rogers W, Qian SX, Brazier J, Kazis LE. A preference-based measure of health: the VR-6D derived from the veterans RAND 12-Item Health Survey. *Qual Life Res* 2011;20(8):1337–47.
- 20 Miller WC, Sakakibara BM, Noonan VK, Tawashy AE, Aubut JL, Connolly SJ, *et al.* Outcome measures. In: Eng JJ, Teasell RW, Miller WC, Wolfe DL, Townson AF, Hsieh JTC, *et al.* (eds.). *Spinal cord injury rehabilitation evidence* (v3.0). Vancouver; p. 1–147 [accessed 2013 August 17]. Available from: http://www.scriproject.com/sites/default/files/outcome_measures.pdf.
- 21 Cunningham WE, Nakazono TT, Tsai KL, Hays RD. Do differences in methods for constructing SF-36 physical and mental health summary measures change their associations with chronic medical conditions and utilization? *Qual Life Res* 2003;12(8): 1029–35.
- 22 Andresen EM, Meyers AR. Health-related quality of life outcomes measures. *Arch Phys Med Rehabil* 2000;81(12 Suppl 2):S30–45.
- 23 Ku JH. Health-related quality of life in patients with spinal cord injury: review of the short form 36-health questionnaire survey. *Yonsei Med J* 2007;48(3):360–70.
- 24 Boakye M, Leigh BC, Skelly AC. Quality of life in persons with spinal cord injury: comparisons with other populations. *J Neurosurg Spine* 2012;17(1 Suppl):29–37.
- 25 Tate DG, Kalpakjian CZ, Forchheimer MB. Quality of life in individuals with spinal cord injury. *Arch Phys Med Rehabil* 2002;83(12 Suppl 2):S18–25.
- 26 Leduc BE, Lepage Y. Health-related quality of life after spinal cord injury. *Disabil Rehabil* 2002;24(4):196–202.
- 27 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 cord injury. Centre for Health Economics Research and Evaluation (CHERE) Working Paper 2007/4 [accessed 2013 August 17]. Available from: http://www.chere.uts.edu.au/pdf/wp2007_4.pdf.
- 28 Lee BB, Simpson JM, King MT, Haran MJ, Marial O. 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. *Spinal Cord* 2009;47(1):50–5.
- 29 Meyers AR, Andresen EM. Enabling our instruments: accommodation, universal design, and access to participation in research. *Arch Phys Med Rehabil* 2000;81(12 Suppl 2):S5–9.
- 30 Cauter A, Northcott S, Clarkson L, Pring T, Hilari K. Does mode of administration affect health-related quality-of-life outcomes after stroke? *Int J Speech Lang Pathol* 2012;14(4):329–37.
- 31 Erhart M, Wetzel RM, Krügel A, Ravens-Sieberer U. Effects of phone versus mail survey methods on the measurement of health-related quality of life and emotional and behavioural problems in adolescents. *BMC Public Health* 2009;30(9):491–502.
- 32 Hanmer J, Hays R, Fryback DG. Mode of administration is important in U.S. national estimates of health-related quality of life. *Med Care* 2007;45(12):1171–9.
- 33 Whitehurst DG, Noonan VK, Dvorak MF, Bryan S. A review of preference-based health-related quality of life questionnaires in spinal cord injury research. *Spinal Cord* 2012;50(9):646–54.
- 34 Ghogawala Z, Martin B, Benzel EC, Dziura J, Magge SN, Abbed KM, *et al.* Comparative effectiveness of ventral vs dorsal surgery for cervical spondylotic myelopathy. *Neurosurgery* 2011;68(3): 622–30.
- 35 Wu M, Landry JM, Schmit BD, Hornby TG, Yen SC. Robotic resistance treadmill training improves locomotor function in human spinal cord injury: a pilot study. *Arch Phys Med Rehabil* 2012;93(5):782–9.
- 36 Froehlich-Grobe K, Andresen EM, Caburnay C, White GW. Measuring health-related quality of life for persons with mobility impairments: an enabled version of the short-form 36 (SF-36E). *Qual Life Res* 2008;17(5):751–70.
- 37 Nanda U, McLendon PM, Andresen EM, Armbrecht E. The SIP68: an abbreviated sickness impact profile for disability outcomes research. *Qual Life Res* 2003;12(5):583–95.
- 38 Rowell D, Connelly LB. Personal assistance, income and employment: the spinal injuries survey instrument (SISI) and its application in a sample of people with quadriplegia. *Spinal Cord* 2008;46(6):417–24.
- 39 Rowell D, Connelly L. Labour market outcomes for people with a spinal cord injury. *Econ Hum Biol* 2010;8(2):223–32.
- 40 Unalan H, Celik B, Sahin A, Caglar N, Esen S, Karamehmetoglu S. Quality of life after spinal cord injury: the comparison of the SF-36 health survey and its spinal cord injury-modified version in assessing the health status of people with spinal cord injury. *Neurosurg Q* 2007;17(3):175–9.
- 41 De Wolf AC, Tate RL, Lannin NA, Middleton J, Lane-Brown A, Cameron ID. The World Health Organization Disability Assessment Scale, WHODAS II: reliability and validity in the measurement of activity and participation in a spinal cord injury population. *J Rehabil Med* 2012;44(9):747–55.
- 42 Dudley-Javoroski S, Shields RK. Assessment of physical function and secondary complications after complete spinal cord injury. *Disabil Rehabil* 2006;28(2):103–10.
- 43 Mueller G, Hopman M, Perret C. Comparison of respiratory muscle training methods in individuals with motor complete tetraplegia. *Top Spinal Cord Inj Rehabil* 2012;18(2):118–21.
- 44 Litchke LG, Lloyd LK, Schmidt EA, Russian CJ, Reardon RF. Effects of concurrent respiratory resistance training on health-related quality of life in wheelchair rugby athletes: a pilot study. *Top Spinal Cord Inj Rehabil* 2012;18(3):264–72.
- 45 Luther SL, Kromrey J, Powell-Cope G, Rosenberg D, Nelson A, Ahmed S, *et al.* A pilot study to modify the SF-36V physical functioning scale for use with veterans with spinal cord injury. *Arch Phys Med Rehabil* 2006;87(8):1059–66.
- 46 Kalpakjian CZ, Scelza WM, Forchheimer MB, Toussaint LL. Preliminary reliability and validity of a Spinal Cord Injury Secondary Conditions Scale. *J Spinal Cord Med* 2007; 30(2):131–9.
- 47 Tate DG, Riley BB, Perna R, Roller S. Quality of life issues among women with physical disabilities or breast cancer. *Arch Phys Med Rehabil* 1997;78(12 Suppl 5):S18–25.
- 48 Edwards L, Krassioukov A, Fehlings MG. Importance of access to research information among individuals with spinal cord injury: results of an evidenced-based questionnaire. *Spinal Cord* 2002; 40(10):529–35.
- 49 Romero FJ, Gambarrutta C, Garcia-Forcada A, Marin MA, Diaz de la Lastra E, Paz F, *et al.* Long-term evaluation of phrenic nerve pacing for respiratory failure due to high cervical spinal cord injury. *Spinal Cord* 2012;50(12):895–8.
- 50 Wilson MW, Richards JS, Klapow JC, DeVivo MJ, Greene P. Cluster analysis and chronic pain: an empirical classification of pain subgroups in a spinal cord injury sample. *Rehabil Psychol* 2005;50(4):381–8.
- 51 Meyers AR, Andresen EM, Hagglund KJ. A model of outcomes research: spinal cord injury. *Arch Phys Med Rehabil* 2000;81(12 Suppl 2):S81–90.
- 52 Tan G, Rintala DH, Jensen MP, Richards JS, Holmes SA, Parachuri R, *et al.* Efficacy of cranial electrotherapy stimulation for neuropathic pain following spinal cord injury: a multi-site randomized controlled trial with a secondary 6-month open-label phase. *J Spinal Cord Med* 2011;34(3):285–96.
- 53 Hultling C, Giuliano F, Quirk F, Peña B, Mishra A, Smith MD. Quality of life in patients with spinal cord injury receiving Viagra

- (sildenafil citrate) for the treatment of erectile dysfunction. *Spinal Cord* 2000;38(6):363–70.
- 54 Tate DG, Forchheimer M. Health-related quality of life and life satisfaction for women with spinal cord injury. *Top Spinal Cord Inj Rehabil* 2001;7(1):1–15.
- 55 Crawford A, Hollingsworth HH, Morgan K, Gray DB. People with mobility impairments: physical activity and quality of participation. *Disabil Health J* 2008;1(1):7–13.
- 56 Martens FM, den Hollander PP, Snoek GJ, Koldewijn EL, van Kerrebroeck PE, Heesakkers JP. Quality of life in complete spinal cord injury patients with a Brindley bladder stimulator compared to a matched control group. *NeuroUrol Urodyn* 2011;30(4):551–5.
- 57 Mattson-Prince J. A rational approach to long term care: comparing the independent living model with agency-based care for persons with high spinal cord injuries. *Spinal Cord* 1997;35(5):326–31.
- 58 van Leeuwen CM, van der Woude LH, Post MW. Validity of the mental health subscale of the SF-36 in persons with spinal cord injury. *Spinal Cord* 2012;50(9):707–10.
- 59 Bryce TN, Budh CN, Cardenas DD, Dijkers M, Felix ER, Finnerup NB, *et al*. Pain after spinal cord injury: an evidence-based review for clinical practice and research. Report of the National Institute on Disability and Rehabilitation Research Spinal Cord Injury Measures meeting. *J Spinal Cord Med* 2007;30(5):421–40.
- 60 Puhan MA, Ahuja A, Van Natta ML, Ackatz LE, Meinert C; Studies of Ocular Complications of AIDS Research Group. Interviewer versus self-administered health-related quality of life questionnaires – does it matter? *Health Qual Life Outcomes* 2011;9:30.
- 61 Walker DM. The internet as a medium for health service research. *Part 1 Nurse Res* 2013;20(4):18–21.
- 62 Dijkers MP. Searching the literature for information on traumatic spinal cord injury: the usefulness of abstracts. *Spinal Cord* 2003;41:76–84.