Review Article

A review and evaluation of patient-reported outcome measures for spasticity in persons with spinal cord damage: Recommendations from the Ability Network – an international initiative

Per Ertzgaard [1], Anand Nene[2], Carlotte Kiekens [3,4], Anthony S. Burns [5]

[1]Rehabiliteringsmedicinska Kliniken, University Hospital, Linköping, Sweden, [2]Formerly, Roessingh Centre for Rehabilitation, Roessingh Research & Development, Enschede, The Netherlands, [3]Department of Development and Regeneration, KU Leuven – University of Leuven, Leuven, Belgium, [4]Physical and Rehabilitation Medicine, University Hospitals Leuven, Leuven, Belgium, [5]Division of Physiatry, Division of Medicine, University of Toronto, Toronto, Canada

Context: Patient-reported outcome measures (PROMs) are valuable for capturing the impact of spasticity on health-related quality of life (HRQoL) in persons with spinal cord damage (SCD) and evaluating the efficacy of interventions.

Objective: To provide practical guidance for measuring HRQoL in persons with spasticity following SCD.

Methods: Literature reviews identified measures of HRQoL and caregiver burden, utilized in studies addressing spasticity in SCD. Identified measures were evaluated for clinical relevance and practicality for use in clinical practice and research. The PRISM, SCI-SET, EQ-5D and SF-36 instruments were mapped to the International Classification of Functioning, Disability and Health (ICF). The PRISM and SCI-SET were evaluated using the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) checklist.

Results: Two spasticity-specific, five generic, and four preference-based measures were identified. ICF mapping and the COSMIN checklist supported the use of the PRISM and SCI-SET in SCD. The SF-36 is considered the most useful generic measure; disability-adapted versions may be more acceptable but further studies on psychometric properties are required. The SF-36 can be converted to a preference-based measure (SF-6D), or alternatively the EQ-5D can be used. While no measures specific to caregivers of people with SCD were identified, the Caregiver Burden Scale and the Zarit Burden Interview are considered suitable.

Conclusion: Recommended measures include the PRISM and SCI-SET (condition-specific), SF-36 (generic), and Caregiver Burden Scale and Zarit Burden Interview (caregiver burden). Consideration should be given to using condition-specific and generic measures in combination; the PRISM or SCI-SET combined with SF-36 is recommended.

Keywords: Muscle spasticity, Spinal cord diseases, Spinal cord injuries, Patient reported outcome measures, Health-related quality of life

Introduction

Spasticity is a common feature of many neurological conditions characterized by upper motor neuron pathology. Examples include stroke, multiple sclerosis, cerebral palsy, traumatic brain injury, and spinal cord damage (SCD). This report focuses on SCD. It summarizes the deliberations and findings of the Outcomes and Access working group of the Ability Network (AN), an international panel of clinical experts with the overarching goal of addressing challenges and barriers to optimizing the management of disabling spasticity in people with SCD.[1]

Correspondence to: Per Ertzgaard Rehabiliteringsmedicinska Kliniken, University Hospital, SE-582 85, Linköping, Sweden; phone +46 707955853. E-mail: Per.Ertzgaard@liu.se

While there have been many different definitions cited in the medical literature, the Ability Network (including the authors) endorses the definition previously proposed by Pandyan *et al.*: "disordered sensori-motor control, resulting from an upper motor neuron lesion, presenting as intermittent or sustained involuntary activation of muscles."[1,2] The severity of spasticity varies and while it is not always problematic, it can be disabling, with profound effects on functioning, wellbeing and health-related quality of life (HRQoL).[1,3–8] The Ability Network has defined disabling spasticity as "spasticity which is perceived by the individual or caregivers as hindering body function, activities, and/or participation."[1]

Following SCD, a routine part of the initial and ongoing assessment of spasticity is the determination of its clinical and functional impact. A wide range of clinical and functional measures have been utilized to assess spasticity in individuals with SCD. The utilization of these measures in clinical practice has recently been discussed by Nene *et al.*[9] To fully capture the impact of spasticity on the individual, it is increasingly recognized that patient-reported outcome measures (PROMs), particularly measures of HRQoL, are an important complement to clinical and functional measures. In addition, disabling spasticity can increase the burden on caregivers and adversely impact their wellbeing; therefore there is also value in capturing this aspect of spasticity.[10–12]

There can be a discrepancy between the individual with SCD's perception of spasticity and clinicians' findings on examination.[13,14] The use of PROMs measuring HRQoL, as well as measures of caregiver burden, can bridge this discrepancy and provide clinicians with additional insight into how spasticity impacts the daily lives and well-being of individuals with SCD and/or their caregivers. Similarly, PROMs and measures of caregiver burden are an important adjunct when evaluating the impact of treatment; if an intervention for spasticity is not resulting in changes that are meaningful to the individual with SCD and/or their caregivers, then this should be a signal to re-evaluate the approach. Finally, HRQoL is increasingly important to healthcare payers in many jurisdictions, and the use of PROMs for assessing HRQoL has become invaluable for determining the relative merits of new and existing treatments.

For all of the above reasons, the measurement of HRQoL and caregiver burden are highly relevant for spasticity following SCD. However, despite the clear rationale for their use, such measures have been underutilized in studies assessing the impact of interventions for spasticity in the context of SCD; probably due to a lack of awareness of the need to measure outcomes beyond body function.[15,16] Measures of HRQoL or health status can be divided into three types: generic measures of HRQoL, measures addressing disease- or condition-specific aspects of quality of life, and preference-based utility measures.

Generic measures describe health status and quality of life across many different health conditions. They typically capture this information using a numeric summary score encompassing a number of different dimensions of health, as well as summary scores for sub-scales. While generic measures capture broad aspects of HRQoL such as ability to participate fully in social and occupational roles, the nature of the questions may fail to capture condition-specific aspects of an experience. Furthermore, generic measures have both methodological and conceptual limitations when used with people with disabilities.[17]

Condition-specific measures are designed to address the limitations of generic measures by capturing the impact of particular clinical features. However, they may not provide a broad picture of social and occupational participation. Unlike generic measures, condition-specific measures do not allow for comparison of HRQoL between conditions.

Preference-based measures are the least specific to the patient's condition, and their ability to capture the patient experience using a limited range of items has been questioned.[18] However, they are central to decisions regarding funding of health care, both for health economic decision-making and prioritizing. They assign values to described health states, typically ranging between 1 (full health) and zero (death), and provide a means of assessing whether one overall health state is better (preferred) than another, without discriminating between the different dimensions influencing HRQoL (e.g. pain, mobility, mental health).[19] Preference-based measures can be used to calculate quality-adjusted life-years (QALY), or simply to assess changes in health state values.

There have been prior evaluations of HRQoL instruments in the context of SCD, notably by the Spinal Cord Injury Research Evidence Project (SCIRE).[20,21] SCIRE, however, did not focus specifically on the assessment of spasticity and its impact on individuals with SCD. In contrast, Ballioussis *et al.* performed a systematic review of instruments with potential utility for assessing HRQoL in the context of spasticity following SCD.[22] HRQoL instruments were identified and described, including relative strengths and limitations. In this report, we extend prior work by (1) expanding the discussion of preference-based measures of HRQoL and their potential role in economic analysis, (2) examining

the breadth of coverage of HRQoL instruments by mapping them to items impacted by spasticity in the International Classification of Functioning, Disability and Health (ICF) framework, (3) identifying and assessing measures with utility for determining caregiver burden, and (4) synthesizing findings into practical suggestions for measuring HRQoL and caregiver burden in people with spasticity and SCD, both in clinical practice and in research. Given their unique characteristics and complementary nature, the role and use of combinations of HRQOL instruments (generic, condition-specific, preference-based) is also discussed.

Methods
Identification of measures for evaluation
In order to identify candidate PROMs for evaluation, a literature review was performed to identify studies which both measured spasticity and included the use of PROMs measuring HRQoL. The Medline/PubMed and EMBASE databases were searched using the keywords spastic* AND (patient-reported outcomes measurements OR health-related quality of life)/ patient-reported outcomes assessment/quality of life/ caregiver quality of life/carer quality of life/institutional care quality of life/social quality of life/ patient outcomes/(health status OR functional status OR well-being). Eligible studies were those in English, French or German that administered HRQoL instruments to patients and measured spasticity occurring as a result of SCD, traumatic brain injury, stroke or multiple sclerosis. As additional assurance that all relevant measures were captured, the results of existing systematic literature reviews assessing quality of life instruments in an SCD population were also consulted.[21–25]

We also undertook a literature review to identify measures of caregiver burden that have been previously used in studies of the caregivers of people with SCD. The databases searched were Medline, EMBASE, The University of Oxford Patient Reported Outcomes Database (UO-PROD), and the Patient-reported Outcome and Quality of Life Instrument Database (PROQOLID). Search terms were carer OR caregiver; index OR scale OR survey OR questionnaire OR checklist OR screen OR inventory OR instrument; outcome OR outcomes OR impact OR "quality of life" OR reaction OR strain OR stress OR hassle OR hassles OR experience OR experiences OR distress OR burden. Studies that did not address the relevant health conditions were excluded. There were no language limitations.

Identified citations were screened for eligibility, and HRQoL and caregiver burden measures were extracted along with details of study design and population. As the reviews were primarily intended for the identification of outcome measures, no further data were extracted or analyzed. Identified measures were then evaluated for their utility for assessing the impact of spasticity following SCD.

Evaluation methods
PROMs should demonstrate adequate performance for three key measurement properties: reliability, validity, and responsiveness (ability to detect change).[26] Widely used generic and preference-based HRQoL measures are not specific to SCD and have been evaluated extensively against these criteria; we therefore did not repeat this exercise. Rather, we assessed generic and preference-based measures in terms of their clinical relevance to people with SCD and spasticity, as well as the practicality of their use for clinical practice and research.

In comparison to generic and preference-based measures, there has been less evaluation of the spasticity-specific measures. We therefore undertook an assessment of these using the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) checklist.[27] The COSMIN is a methodological framework for assessing the quality of health measurement instruments using a four-level rating system (excellent, good, fair, poor). As with the generic and preference-based measures, we evaluated the relevance and practicality of spasticity-specific measures for people with SCD.

We restricted our evaluation of caregiver burden measures to instruments that measure multiple constructs and have a substantial focus on the negative impacts (i.e. burden) of caregiving. Scales that focus on single issues such as stress were not included. For the identified measures, we searched for validation reports in the literature.

Mapping to the ICF framework
The International Classification of Functioning, Disability and Health (ICF) provides a standardized framework for the description of functioning and disability. It is organized into three domains: body function and structure, activity, and participation. Each domain contains subdomains comprised of multiple items.[28] Outcome measure instruments can be evaluated using the ICF framework by determining the level of coverage for individual items in accompanying subdomains, a process known as mapping. Using guidelines by Cieza et al.,[29,30] mapping was performed for two spasticity-specific measures (the Patient Reported Impact of Spasticity Measure [PRISM] and the Spinal Cord

Injury-Spasticity Evaluation Tool [SCI-SET]), the EuroQoL 5-dimension (EQ-5D), and the SF-36. The aim was to assess the breadth of coverage of ICF items relevant to spasticity.

Results

Identification of outcome measures

The literature review identified 56 studies that met inclusion criteria. In addition, one systematic review was identified (Balioussis *et al.*[22]) that specifically addressed measuring the impact of spasticity in SCD on HRQoL. The measures identified by Balioussis *et al.* were the Health Utilities Index (HUI), Life Situation Questionnaire Revised (LSQ-R), PRISM, Quality of Life Index SCI version (QLI-SCI), Sickness Impact Profile (SIP68), Short-form 36 (SF-36), SCI-SET, and two questionnaires specifically related to mood and social support. Our search did not identify any additional instruments. Although it was not identified in the literature review limited to studies of spasticity in SCD, we added the World Health Organization Quality of Life-BREF scale (WHOQOL-BREF) to the set of measures evaluated, the rationale being that it is used frequently in SCD, has multiple translations and is free of charge.

Neither our search nor the Balioussis review identified studies using preference-based measures other than the HUI. To our knowledge, no review has focused on the use of preference-based measures to assess the impact of spasticity in people with SCD. However, Whitehurst *et al.* (2012) systematically reviewed the use of generic preference-based measures of HRQoL in SCD.[25] They found only two studies using preference-based measures in the context of SCD; one used the Quality of Wellbeing Self-Administered (QWB-SA)[31] and one the SF-6D.[32] The two most widely recognized preference-based measures are SF-6D and the EQ-5D.[33] We therefore evaluated both for their appropriateness for clinical use in spasticity following SCD.

Spasticity-specific measures of HRQoL

Two spasticity-specific PROMS were identified: the PRISM[34] and the SCI-SET.[35] The PRISM consists of 41 items that describe impacts of spasticity, each of which is rated on a scale of 0–4 from "never true for me" to "very often true for me". Items are divided into seven subscales (social avoidance/anxiety; psychological agitation; social embarrassment; positive impact; daily activities; need for assistance/repositioning; need for intervention). Scoring is separate for each subscale; there is no total score. The SCI-SET consists of 35 items reflecting the impact of spasticity. Items are scored using a 7-point response scale, ranging from -3 (extremely problematic) to $+3$ (extremely helpful). The SCI-SET is not divided into domains or subscales, and not all items require a response. The item scores are averaged to create an overall SCI-SET score.

In their literature review, Balioussis *et al.* called for further research to establish the psychometric properties of the PRISM and SCI-SET.[22] In order to compare the PRISM and the SCI-SET and illuminate existing knowledge gaps regarding psychometric properties, we used the COSMIN framework (Table 1). COSMIN is a rigorous evaluation measure that uses a "worst score counts" approach. For example, if five items in a measure are assessed and four are "good" and one is "fair", the overall score for that measure is "fair". Furthermore, items may be rated "poor" if their development is insufficiently detailed, as opposed to poor methodology *per se*. Items in both questionnaires had to be downgraded over relatively minor issues.

The PRISM scored higher than the SCI-SET, particularly in the assessment of validity. However, the scientific quality of the SCI-SET is not necessarily inadequate. The methodology behind their development was judged to be sound for both measures, and both have been developed from the perspective of the person with SCD.

Table 1   Summary of COSMIN evaluation of spasticity-specific patient reported outcome measures: PRISM and SCI-SET.

Area of assessment	Measurement property	PRISM[34]	SCI-SET[35]
Reliability	Internal consistency	Good	Fair (to "Poor")
	Reliability (test/retest)	Fair	Fair (to "Poor")
	Measurement error	No evidence reported	No evidence reported
Validity	Content validity	Good	Fair (to Poor)
	Construct validity (structural validity)	Good	No evidence reported
	Construct validity (hypotheses testing)	Fair	Fair (to Poor)
	Construct validity (cross-cultural validity)	Not relevant/appropriate	Not relevant/Not appropriate
	Construct validity (criterion validity)	Not relevant/appropriate	Not relevant/Not appropriate
Responsiveness		No evidence reported	No evidence reported
Interpretability		Concerns	Concerns
Generalizability		Adequate	Adequate/Concerns

Table 2   Overall comparison of PRISM and SCI-SET from a clinical practice perspective.

Parameter	PRISM[34]	SCI-SET[35]
Scientific quality based on COSMIN evaluation	Good or fair	Generally fair, some aspects poor; but probably adequate scientific quality
ICF coverage (number of items)	*Body function*: focus primarily mental (mental function 13; genitourinary & reproductive 1) *Activity & participation*: wider coverage of self-care and interpersonal domains (general tasks 2; mobility 8; self-care 10; interpersonal & relationships 11; community, social & civic 1)	*Body function*: wider coverage (mental function 9 items; sensory & pain 1 item; genitourinary & reproductive 1 item; neuromusculoskeletal & movement 3 items) *Activity & participation*: wider mobility coverage (general tasks 2; mobility 12; self-care 5; domestic life 2; interpersonal 2; community 1)
Clinical relevance in SCD	Adequate but may be less relevant than SCI-SET in the context of SCD Similar to a conventional HRQoL measure; less coverage of symptoms; does not assess sleep; less sensitive to positive aspects of spasticity	More intuitive and clinically useful than PRISM Better coverage of symptoms; assesses sleep; better sensitivity to positive aspects
Clarity of questions and ease of interpretation	Questions more specific; answers clearer to interpret; has subscales	Some ambiguity, e.g. difference between "not applicable" and "no effect" could confuse; no subscales
Generalizability to all spasticity versus specificity for SCD	More generalizable to spasticity across conditions	Developed specifically for SCD
Sensitivity, responsiveness and validity	Respondents experiencing no spasticity could answer "never or rarely" Validity of PRISM subscales has been reported using a clinically assessed sample of patients. PRISM subscale scores corresponded to statistically significant mean rank scores, indicating the more severe the spasticity, the greater the PRISM score (greater the impact of spasticity)[34] Minimal clinically important difference (MCID) has not been assessed for PRISM to date	Not apparent how answers would change if a person no longer had spasticity. Some intervention studies have found no change in SCI-SET score despite changes in clinical measures of spasticity.[35,61] This is not surprising as the SCI-SET is unbiased towards positive and negative effects of spasticity SCI-SET has shown moderate to strong correlations with self-assessed spasticity severity and spasticity impact[35]
Languages available	English Serbian[57]	English Persian[58] Turkish[59,60]

Methodological quality is just one aspect, albeit important, to be taken into account when evaluating the utility of a PROM. Broader criteria relating to clinical relevance and ease of use are also important. The observations of the working group from a clinical practice perspective are summarized in Table 2. The ICF mapping (Table 2) suggests that, while there is much overlap between the measures, the PRISM has the greater emphasis on mental functions and interpersonal interaction, whereas the SCI-SET places greater emphasis on neuromusculoskeletal function and mobility. This might be a factor in the choice of instrument, depending on the relative importance of these two aspects for specific patients or studies.

Evidence of "responsiveness" and "minimal important differences" is currently lacking for both the PRISM and the SCI-SET. The minimal important difference is the smallest difference in scores that equates to a clinically meaningful difference from the patient's perspective. Responsiveness and minimal important difference can often only be evaluated after a measure has been used in several studies. Given that

neither measure has been extensively used, a body of evidence on how changes in score correlate to changes in clinical state has yet to be developed. It is therefore an open question whether either measure is adequate for detecting clinical changes.

Our conclusion is that both the PRISM and the SCI-SET can be recommended as spasticity-specific PROMs in SCD. Our findings concur with those of Balioussis *et al*. (2014),[22] who also concluded that the SCI-SET and the PRISM were the most promising tools for assessing the impact of spasticity on HRQoL after SCD. Each has its relative advantages and disadvantages, which should be weighed when deciding which measure best fits specific circumstances and objectives. Both are feasible for use in routine practice. However, the SCI-SET is currently only available in English, Turkish and Persian, and the PRISM only in English and Serbian; this is a major obstacle to their use in other language regions. A formal linguistic validation process needs to be undertaken for a PROM before validity can be assumed to apply to a translated version.[36]

Generic measures of HRQoL

The generic measures evaluated for use in spasticity following SCD were the SF-36 and its modifications, the LSQ-R, the QLI-SCI, the SIP-68 and the WHOQOL-BREF. Each measure has relative strengths and weaknesses, which are outlined below. More detailed reviews can be found in Balioussis *et al.*[22] and on the SCIRE project website,[37] although the latter focuses on their use in SCD in general rather than specifically for spasticity.

SF-36 and modifications

The SF-36[38] is one of the most widely used and validated generic health status measures and is available in over 170 languages. It provides a self-reported health status profile consisting of eight dimensions: general health; bodily pain; physical functioning; role-physical; mental health; vitality; social functioning; and role-emotional. These eight dimensions can be used to generate a physical and mental health summary score, after which health status may be compared against "norms" for the general population. The mapping exercise showed that the SF-36 has wide coverage of relevant items in the activity and participation ICF domains. In addition, SF-36 scores can be used to derive a preference-based measure, the Short Form 6D (SF-6D), without the need to administer an additional questionnaire.[32] A license fee is payable in order to use SF-36.

A disadvantage of the SF-36 in the context of SCD is that it contains questions on climbing stairs and walking. Some wheelchair-dependent people with SCD find this irrelevant, offensive or upsetting and may refuse to answer the questionnaire. A number of modifications have been proposed to remedy this, and these are reviewed by Whitehurst *et al.*[39] The SF-36 veterans version (SF36 V)[40] was created for this purpose but refers only to "wheeling". We did not consider this as an improvement on the standard version in the context of SCD, because not all respondents will be wheelchair users. The Enabled SF-36 (SF-36 E)[41] contains the enabled physical functioning items proposed by Meyers and Andresen,[42] and refers to activities "using your normal assistive devices (wheelchair/cane/prosthetic)". The SF-36 Walk-Wheel (SF-36 WW) has the standard SF-36 questions and three additional modified questions that use the term "wheel" in place of "walk".[43]

In their comprehensive review of SF-36 modifications in SCI, Whitehurst *et al.* note that the well-established validity of the original version cannot be assumed to apply to modified versions or to the use of the instrument with some items removed.[39] They note that there

is a trade-off between validity and clinical relevance, but conclude that validity is paramount. We do not recommend removing questions if using the SF-36 in research, because the results could be questioned by regulatory or health technology assessment authorities on the grounds that the validity of the instrument is compromised.

In our experience, the standard SF-36 can be successfully used in people with SCD if respondents are warned in advance that some of the questions are aimed at the general population and might not apply to them. However, the question remains as to whether they should leave the irrelevant items blank or answer them as though they referred to assisted mobility. Both of these options will affect the scoring in ways that may not accurately reflect the individual's HRQoL.[39] Another problem is that people using a wheelchair are scored unfairly low with the standard SF-36.[43] Use of a more disability-unbiased version of SF-36 is pertinent, but further studies on psychometric characteristics are needed before such adapted versions can be generally recommended.

WHOQOL-BREF

The WHOQOL-BREF is comprised of 26 items addressing the domains of physical health, psychological health, social relationships, and environment, and is available in over 20 languages.[44,45] It is a potentially relevant and useful measure for determining how SCD with spasticity impacts HRQoL, particularly as it uses the non-specific term "getting around" in relation to mobility. In a review of the use of HRQoL instruments in SCD (without a focus on spasticity), Hill *et al.*[23] concluded that the WHOQOL-BREF was the most acceptable and established instrument to assess HRQoL after SCD. However, it is arguably less suitable than the SF-36 for evaluating the impact of health interventions, because it includes aspects of quality of life that go beyond those related to health (e.g. questions about the person's economic situation).

QLI-SCI

The Quality of Life Index (SCI version; QLI-SCI) measures quality of life in terms of satisfaction within four domains: health and functioning, psychological/spiritual, social and economic, and family. The SCI version contains 37 items, and an evaluation is available on the SCIRE website.[37] It has not been widely used in spasticity-related studies, and Balioussis *et al.*[22] judged that further work is needed in order to determine whether or not it is sensitive to spasticity. It is available free of charge.

Sickness impact profile (SIP)

The Sickness Impact Profile (SIP)[46] has been used in a small number of spasticity-related studies in SCI and found to be sensitive to spasticity.[22] The SIP assesses physical, and psychosocial dimensions of health-related functioning and behavior using 12 categories. It is comprised of a list of statements addressing the impact of sickness, to which respondents indicate which items describe their health status.[37] The full SIP has 136 items and the time required for completion could affect feasibility in clinical practice. Even a shortened version, with 68 items (SIP 68), is lengthy and takes 15–20 min to complete. The nature of the questions may also be perceived as negative. It includes references to walking, although a scoring modification for wheelchair users has been proposed.[37] The SIP is available in English, Spanish and several other languages, and is free of charge for individual and academic use.

Life situation questionnaire-revised (LSQ-R)

The Life Situation Questionnaire-Revised (LSQ-R) is part of the larger Life Situation Questionnaire (LSQ).[47] It is designed specifically for use in people with SCD, and spasticity has been found to be negatively correlated with life satisfaction scores derived from the LSQ-R.[4] However, much of the LSQ-R addresses aspects of life that are not directly health-related (e.g. living arrangements, accomplishments, access to transportation), so we do not recommend its use as a measure for assessing the impact of health interventions.

Recommendation for generic measure

Our recommended generic measure is the SF-36, on the basis of its well-validated psychometric properties, its widespread use (which facilitates comparisons across studies) and its ability to be converted to a preference-based measure if required. The other measures may be useful in specific circumstances, for example if an instrument that is free of charge is required, or if they are a better fit to the particular research question being asked.

Preference-based measures of HRQoL

In their review, Whitehurst *et al.* noted "a distinct lack of conceptual or empirical research regarding the appropriateness of alternative preference-based HRQoL measures for SCD populations," and highlighted the need for more research in this area to enable interventions in SCD to be evaluated for cost-effectiveness. Another potential problem with preference-based measures is that the preferences are drawn from the general population and do not necessarily reflect the preferences of people with disabilities.[43] However, the

use of preference-based measures may be required in some situations, particularly if the study is to form the basis of an economic evaluation.

We concluded that both the SF-6D and EQ-5D are suitable for use in studies evaluating spasticity in people with SCD. Their relative advantages and disadvantages are discussed below.

The EQ-5D[33] is regarded as the measure of choice (albeit with limitations) by a number of reimbursement/health technology assessment authorities. The EQ-5D has five dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression), each having three levels of severity (EQ-5D-3L). A new version is now available with five levels of severity, ranging from "no problems" to "extreme problems" (EQ-5D-5L). This version addresses some of the shortcomings of the three-level questionnaire and is likely to be more appropriate in SCD; in particular, the mobility dimension where the previous wording of "confined to bed" has been replaced with "unable to walk about", although this is still far from ideal for wheelchair users. Coverage of ICF items is much less with the EQ-5D compared to the SF-36. The primary advantages of the EQ-5D are that it is widely recognized and is short and simple, thereby minimizing time requirements (burden) for respondents and staff.

The SF-6D is also widely accepted and can be derived from the responses to the SF-36, making it unnecessary to administer a second questionnaire. The primary disadvantage of the SF-6D is the floor effect where the lowest possible QALY value is 0.3. In contrast, the EQ-5D can capture even negative ("worse than death") health states. Thus, the SF-6D is likely to be less sensitive to severe health states. Work is currently under way to create a new value set for the SF-6D intended to address these problems. SF-6D's validity in SCI (but not specifically in the context of spasticity) was explored by Engel *et al.*, who found good practicality and discriminative validity but low responsiveness for detecting important health changes over time.[48] They also noted that its comparative performance against other preference-based measures in SCI is unknown.

The two other major preference-based measures are the Health Utilities Index (HUI)[49] and the Quality of Wellbeing Self-Administered (QWB-SA).[31] The HUI has not been widely used in studies of spasticity in SCD,[22] and its overall use is less widespread than the EQ-5D and SF-6D. It requires a substantial license fee and may be less familiar to health technology assessment agencies than the EQ-5D and SF-6D. The QWB-SA has been used in the setting of SCD but not

for spasticity specifically. An evaluation of the QWB-SA for use in SCD is available from the SCIRE project.[37]

Whitehurst *et al.*[50] published an interesting insight into people with SCI's perceptions of the available preference-based measures. All provoked a mixed reaction when qualitatively assessed by focus groups. The QWB-SA was the least preferred. Many participants were frustrated by their inability to adequately describe their health state within the framework of the four major measures. All participants ($n = 15$) perceived the most relevant instrument to be the Assessment of Quality of Life-8D (AQoL-8D), an Australian measure that has been little used outside of that country. This measure was considered to be comprehensive, wheelchair inclusive, and to use applicable items. The authors also note that it has a strong theoretical underpinning. Significant disadvantages are its lack of widespread use and the consequent difficulty in making comparisons across studies or between health conditions.

Measures of caregiver burden

Spasticity can have significant consequences for caregivers, in both the familial and the institutional setting. Generic HRQoL measures can be used with caregivers. For example, we found four studies that used the SF-36 with caregivers of people with SCI, and the SF-36 has also been used in a large number of other caregiver groups. In view of its widespread use and its well-validated psychometric properties, it can be regarded as a valid choice in this setting. However, while generic instruments measure a broad perspective of burden, they do not provide insight into caregiver-specific problems. To gain maximum insight it is preferable to use a caregiver-specific instrument.

We identified four studies measuring caregiver burden in SCI that met our inclusion criteria. Of these, three used the Caregiver Burden Scale and one used the Zarit Burden Interview. A review of the literature revealed that both measures had adequate psychometric properties. The Caregiver Burden Scale[51–54] is a 22-item instrument originally developed for use with dementia carers. It covers the domains of general strain, isolation, disappointment, emotional involvement and environment. The original language is Swedish but English and Portuguese (for Brazil) translations are also available. The Zarit Burden Interview,[55,56] developed for carers of the elderly, has 22 items and covers the domains of health, negative affect, social support, stress and coping, personal strain, and role strain. The original language was American English, but French, Portuguese and Japanese translations are available. Given the lack of condition-specific measures, both

can be considered suitable for use in SCD caregiver populations. We found no measures specifically designed for caregivers of people with SCD. This is a relevant area for further research, particularly given the increasing interest in the impact of treatments on indirect and societal burden.

Discussion

A multidimensional evaluation of spasticity is needed to adequately capture its impact on all aspects of affected persons' lived experience, and to facilitate comprehensive assessment of the effects of treatment on both clinical outcome measures and on quality of life. In order to be meaningful, measures must reflect the full impact of spasticity on people with SCD, who face distinct challenges that are not typical of the general population. Furthermore, measures must be sensitive to change, valid and reliable.

In this report we identified patient-reported HRQoL measures suitable for assessing the impact of spasticity following SCD. The advantages and disadvantages are outlined, but rigid recommendations have not been made. In all cases, the decision on which measure to use should take into account specific circumstances and preferences. The choice of instruments depends on the research question as well as other factors such as whether the work is likely to be used in an economic evaluation, available translations and costs for using the instrument.

Condition-specific outcome measures for the assessment of spasticity (PRISM and SCI-SET) have been used in relatively few studies. In carrying out a COSMIN assessment and ICF mapping of these measures we have added to the previous work published by Balioussis *et al.*[22] As experience with their use grows, it will be possible to more accurately evaluate aspects of their validity, as well as their responsiveness to change. Ongoing work by the Ability Network to develop a Spasticity Set within the ICF framework will provide further insight into any gaps in the coverage of currently available measures, and may suggest directions for further development. A major impediment to the use of PRISM and SCI-SET is their availability in English only (plus Serbian[57] for PRISM and Persian[58] and Turkish[59,60] for SCI-SET). There is therefore an urgent need for validated translated versions for use in other language regions.

The preference-based measures used for health-economic evaluations are rarely used in studies addressing SCD and were not used at all in the studies we identified that addressed spasticity in the context of SCD. We conclude that both the EQ-5D and SF-6D are suitable for

assessing the impact of spasticity after SCD, albeit with the understanding that responsiveness might be diminished[48] and some questions could be controversial for persons with disabilities.[50] Few studies used measures of caregiver burden following SCD, and none of these addressed spasticity after SCD. No measures of caregiver burden have been specifically developed for either spasticity or SCD. The Caregiver Burden Scale and the Zarit Burden Interview have been used, but further research is needed to establish their utility in these groups.

Choice of combinations of measures

Where possible, consideration should be given to using both condition-specific and generic HRQoL measures. This provides additional assurance that the patient/respondent's perspective is comprehensively captured. In many situations a preference-based measure will also be important if the study is to be used to inform health technology assessment or payer decision-making.

The choice of condition-specific and generic measures should not be made in isolation, and it is important to consider how the chosen measures work together, in terms of both coverage and administrative burden to respondents and staff. For example, there is a degree of overlap between the SF-36 and the PRISM. The selection of appropriate measures should also be based on the specific circumstances in which they are to be used. In research, the choice of instruments primarily depends on the research question that needs to be answered. In contrast, in clinical practice there is usually a compromise between the need to evaluate patients and perhaps report data for real-world evidence projects, and the necessity to be time-efficient.

It is not necessarily advantageous to use all three types of measures (condition-specific, generic, preference-based), and a choice of two that provide complementary coverage will often be sufficient. In the case where condition-specific, generic and preference-based measures are all desired, a combination of SF-36 and SCI-SET could be one possibility, as a preference-based measure (SF-6D) can be derived from the SF-36, thereby alleviating the need to administer a third questionnaire.

Limitations

It is important to acknowledge limitations of our work. We did not carry out a full systematic review of the literature relating to all aspects of HRQoL measures that have been used in spasticity in SCD. However, the Hill[23] SCIRE,[21] Balioussis,[22] and Whitehurst[39]

reviews, together with the targeted searches performed for this evaluation, provide broad coverage of the literature. Although the COSMIN exercise and ICF mapping were incorporated to support objective evaluation, a formal Delphi process with a larger group was not conducted to reach formal recommendations. The evaluation was limited to adults with SCD, so findings cannot be extended to children with SCD or to other etiologies of spasticity. Complementary evaluations will need to be performed for these groups. Furthermore, the evaluation of generic HRQoL measures was restricted to those that have already been used in spasticity with SCD. Other measures that have been used in SCD more broadly[23,37] might also be suitable for use in spasticity; this would be another area for further research.

Conclusion

PROMs are indispensable for determining HRQoL and are a valuable adjunct to clinical and functional measures for the assessment and evaluation of persons with spasticity following SCD. They play a valuable role in evaluating the effectiveness of interventions for individual patients, and are needed to promote enhanced access and funding for treatments that demonstrate clear benefit. PROMs that should be given particular consideration include the PRISM or SCI-SET to capture spasticity-specific aspects, the SF-36 as a generic measure of HRQoL, and the SF-6D or EQ-5D when a preference-based measure is required. Assessment can be further enhanced by using a combination of complementary HRQoL instruments, such as the PRISM or SCI-SET combined with the SF-36.

Disclaimer statements

ORCID

Per Ertzgaard http://orcid.org/0000-0001-6464-9130
Carlotte Kiekens http://orcid.org/0000-0001-8500-7751
Anthony S. Burns http://orcid.org/0000-0002-8679-3958

References

1 Burns AS, Lanig I, Grabljevec K, New PW, Bensmail D, Ertzgaard P, *et al.* Optimizing the management of disabling spasticity following spinal cord damage: the ability network-an international initiative. Arch Phys Med Rehabil 2016;97(12):2222–8.

2 Pandyan AD, Gregoric M, Barnes MP, Wood D, Van Wijck F, Burridge J, *et al.* Spasticity: clinical perceptions, neurological realities and meaningful measurement. Disabil Rehabil 2005;27(1-2): 2–6.

3 Adriaansen JJ, Ruijs LE, van Koppenhagen CF, van Asbeck FW, Snoek GJ, van Kuppevelt D, *et al.* Secondary health conditions and quality of life in persons living with spinal cord injury for at least ten years. J Rehabil Med 2016;48(10):853–60.

4 Westerkam D, Saunders LL, Krause JS. Association of spasticity and life satisfaction after spinal cord injury. Spinal Cord 2011;49 (9):990–4.

5 Flachenecker P, Henze T, Zettl UK. Spasticity in patients with multiple sclerosis - clinical characteristics, treatment and quality of life. Acta Neurol Scand 2014;129(3):154–62.

6 Gillard PJ, Sucharew H, Kleindorfer D, Belagaje S, Varon S, Alwell K, *et al.* The negative impact of spasticity on the health-related quality of life of stroke survivors: a longitudinal cohort study. Health Qual Life Outcomes 2015;13:159.

7 Skold C, Levi R, Seiger A. Spasticity after traumatic spinal cord injury: nature, severity, and location. Arch Phys Med Rehabil 1999;80(12):1548–57.

8 Levi R, Hultling C, Nash MS, Seiger A. The Stockholm spinal cord injury study: 1. Medical problems in a regional SCI population. Paraplegia 1995;33(6):308–15.

9 Nene AV, Rainha Campos A, Grabljevec K, Lopes A, Skoog B, Burns AS. Clinical assessment of spasticity in people with spinal cord damage: recommendations from the ability network, an international initiative. Arch Phys Med Rehabil 2018;99(9):1917–26.

10 Doan QV, Brashear A, Gillard PJ, Varon SF, Vandenburgh AM, Turkel CC, *et al.* Relationship between disability and health-related quality of life and caregiver burden in patients with upper limb poststroke spasticity. PM & R: J Injury, Funct & Rehabil 2012;4(1):4–10.

11 Zorowitz RD, Gillard PJ, Brainin M. Poststroke spasticity: sequelae and burden on stroke survivors and caregivers. Neurology 2013; 80(3 Suppl 2):S45–52.

12 Clark PC, Aycock DM, Reiss A, Tanner D, Shenvi NV, Easley KA, *et al.* Potential benefits for caregivers of stroke survivors receiving BTX-A and exercise for upper extremity spasticity. Rehabil Nurs: Off J Assoc Rehabil Nurses 2015;40(3):188–96.

13 Lechner HE, Frotzler A, Eser P. Relationship between self- and clinically rated spasticity in spinal cord injury. Arch Phys Med Rehabil 2006;87(1):15–9.

14 Voerman GE, Fleuren JF, Kallenberg LA, Rietman JS, Snoek GJ, Hermens HJ. Patient ratings of spasticity during daily activities are only marginally associated with long-term surface electromyography. J Neurol Neurosurg Psychiatry 2009;80(2):175–81.

15 Thavaneswaran S, Stauble F. An assessment of health-related quality of life patient-reported outcomes for spasticity in spinal cord injury. Presented at the international spinal cord society (ISCOS) 52nd annual scientific meeting; Istanbul, Turkey, October 2013.

16 Priebe MM, Sherwood AM, Thornby JI, Kharas NF, Markowski J. Clinical assessment of spasticity in spinal cord injury: a multidimensional problem. Arch Phys Med Rehabil 1996;77(7):713–6.

17 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.

18 Devlin NJ, Krabbe PF. The development of new research methods for the valuation of EQ-5D-5L. Eur J Health Econ 2013;14(Suppl 1):1–3.

19 Neumann PJ, Goldie SJ, Weinstein MC. Preference-based measures in economic evaluation in health care. Annu Rev Public Health 2000;21:587–611.

20 Spinal Cord Injury Research Evidence (SCIRE). Outcome measures toolkit. [cited July 2016]. Available from https://www.scireproject.com/outcome-measures-new/toolkit-development.

21 Miller WC, Noonan VN, Sakakibara BM, Tawashy AE, Aubut JL, Curt A, *et al.* Outcome Measures. 2013. In: Eng JJ *et al.* Spinal Cord Injury Rehabilitation Evidence Version 4.0 [Internet]. Vancouver: Spinal Cord Injury Rehabilitation Evidence; 28.1–.366. Available from http://www.scireproject.com/sites/default/files/SCIRE4-OM-Full-Ch.pdf.

22 Balioussis C, Hitzig SL, Flett H, Noreau L, Craven BC. Identifying and classifying quality of life tools for assessing spasticity after spinal cord injury. Top Spinal Cord Injury Rehabil 2014; 20(3):208–24.

23 Hill MR, Noonan VK, Sakakibara BM, Miller WC, Team SR. Quality of life instruments and definitions in individuals with spinal cord injury: a systematic review. Spinal Cord 2010;48(6): 438–50.

24 Ravenek KE, Ravenek MJ, Hitzig SL, Wolfe DL. Assessing quality of life in relation to physical activity participation in persons with spinal cord injury: a systematic review. Disabil Health J 2012;5(4):213–23.

25 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.

26 FDA. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. U.S. FDA. December 2009.

27 Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012;21(4):651–7.

28 World Health Organization. Towards a common language for functioning, disability and health: ICF The International Classification of functioning, disability and health. Geneva: WHO; 2002.

29 Cieza A, Brockow T, Ewert T, Amman E, Kollerits B, Chatterji S, *et al.* Linking health-status measurements to the international classification of functioning, disability and health. J Rehabil Med 2002;34(5):205–10.

30 Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustun B, Stucki G. ICF linking rules: an update based on lessons learned. J Rehabil Med 2005;37(4):212–8.

31 UCSD Health Services Research Center. Quality of well-being scale – self administered (QWB-SA). Available from https://hoap.ucsd.edu/qwb-info/.

32 Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 health survey. J Clin Epidemiol 1998;51(11):1115–28.

33 EUROQOL. EQ-5D. Available from http://www.euroqol.org/home.html.

34 Cook KF, Teal CR, Engebretson JC, Hart KA, Mahoney JS, Robinson-Whelen S, *et al.* Development and validation of patient reported impact of spasticity measure (PRISM). J Rehabil Res Dev 2007;44(3):363–71.

35 Adams MM, Ginis KA, Hicks AL. The spinal cord injury spasticity evaluation tool: development and evaluation. Arch Phys Med Rehabil 2007;88(9):1185–92.

36 Wild D, Eremenco S, Mear I, Martin M, Houchin C, Gawlicki M, *et al.* Multinational trials—recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: the ISPOR patient-reported outcomes translation and linguistic validation good research practices task force report. Value Health 2009;12(4):430–40.

37 SCIRE Project. Outcome measures: quality of life and health status. [cited October 2016]. Available from https://www.scireproject.com/outcome-measures/list-sci.

38 Ware JE, Snoww K, Kosinski M, Gandek B. SF-36® health survey manual and interpretation guide. Boston, MA: New England Medical Center, The Health Institute; 1993.

39 Whitehurst DG, Engel L, Bryan S. Short form health surveys and related variants in spinal cord injury research: a systematic review. J Spinal Cord Med 2014;37(2):128–38.

40 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.

41 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.

42 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.

43 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.

44 Skevington SM, Lotfy M, O'Connell KA, Group W. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004;13(2):299–310.

45 WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. Psychol Med 1998;28(3):551–8.

46 Bergner M, Bobbitt RA, Carter WB, Gilson BS. The sickness impact profile: development and final revision of a health status measure. Med Care 1981;19(8):787–805.

47 Krause JS, Crewe NM. Long term prediction of self-reported problems following spinal cord injury. Paraplegia 1990;28(3): 186–202.

48 Engel L, Bryan S, Evers SM, Dirksen CD, Noonan VK, Whitehurst DG. Exploring psychometric properties of the SF-6D, a preference-based health-related quality of life measure, in the context of spinal cord injury. Qual Life Res 2014;23(8): 2383–93.

49 Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, *et al.* Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care 2002;40(2):113–28.

50 Whitehurst DG, Suryaprakash N, Engel L, Mittmann N, Noonan VK, Dvorak MF, *et al.* Perceptions of individuals living with spinal cord injury toward preference-based quality of life instruments: a qualitative exploration. Health Qual Life Outcomes 2014;12(1):1–9.

51 Macera CA, Eaker ED, Jannarone RJ, Davis DR, Stoskopf CH. A measure of perceived burden among caregivers. Eval Health Prof 1993;16(2):204–11.

52 Simpson G, Jones K. How important is resilience among family members supporting relatives with traumatic brain injury or spinal cord injury? Clin Rehabil 2013;27(4):367–77.

53 Graca A, Nascimento MA, Lavado EL, Garanhani MR. Quality of life of primary caregivers of spinal cord injury survivors. Rev Bras Enferm 2013;66(1):79–84.

54 Blanes L, Carmagnani MI, Ferreira LM. Health-related quality of life of primary caregivers of persons with paraplegia. Spinal Cord 2007;45(6):399–403.

55 Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. Gerontologist 1980;20 (6):649–55.

56 Rodakowski J, Skidmore ER, Rogers JC, Schulz R. Role of social support in predicting caregiver burden. Arch Phys Med Rehabil 2012;93(12):2229–36.

57 Knezevic T, Konstantinovic L, Rodic S, Foti C, Drulovic J, Dackovic J, *et al.* Validity and reliability of the Serbian version of patient-reported impact of spasticity measure in multiple sclerosis. Int J Rehabil Res 2015;38(3):199–205.

58 Ansari NN, Kashi M, Naghdi S. The spinal cord injury spasticity evaluation tool: a Persian adaptation and validation study. J Spinal Cord Med 2017;40(4):380–8.

59 Akpinar P, Atici A, Ozkan FU, Aktas I, Kulcu DG, Kurt KN. Reliability of the spinal cord assessment tool for spastic reflexes. Arch Phys Med Rehabil 2017;98(6):1113–8.

60 Akpinar P, Atici A, Kurt KN, Ozkan FU, Aktas I, Kulcu DG. Reliability and cross-cultural adaptation of the Turkish version of the spinal cord injury spasticity evaluation tool. Int J Rehabil Res 2017;40(2):152–7.

61 Boutilier G, Sawatzky BJ, Grant C, Wiefelspuett S, Finlayson H. Spasticity changes in SCI following a dynamic standing program using the Segway. Spinal Cord 2012;50(8):595–8.