#1

From the 2006 NIDRR SCI Measures Meeting
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

Pain Committee:

Thomas N. Bryce, MD[1]; Cecilia Norrbrink Budh, RPT, PhD[2,3]; Diana D. Cardenas, MD[4]; Marcel Dijkers, PhD[1]; Elizabeth R. Felix, PhD[5]; Nanna B. Finnerup, MD[6]; Paul Kennedy, DPhil[7]; Thomas Lundeberg, MD, PhD[3]; J. Scott Richards, PhD[8]; Diana H. Rintala, PhD[9]; Philip Siddall, MBBS, PhD[10]; Eva Widerstrom-Noga, DDS, PhD[5]

Received March 6, 2007; accepted April 2, 2007

Abstract

Background/Objectives: To examine the reliability, validity, sensitivity, and practicality of various outcome measures for pain after spinal cord injury (SCI), and to provide recommendations for specific measures for use in clinical trials.

Data Sources: Relevant articles were obtained through a search of MEDLINE, EMBASE, CINAHL, and PubMed databases from inception through 2006.

Study Selection: The authors performed literature searches to find articles containing data relevant to the reliability and validity of each pain outcome measure in SCI and selected non-SCI populations.

Data Extraction: After reviewing the articles, an investigator extracted information utilizing a standard template. A second investigator reviewed the chosen articles and the extracted pertinent information to confirm the findings of the first investigator.

Data Synthesis: Taking into consideration both the quantity and quality of the studies analyzed, judgments on reliability and validity of the measures were made by the two investigators. Based upon these judgments, recommendations were formulated for use of specific measures in future clinical trials. In addition, for a subset of measures a voting process by a larger group of SCI experts allowed formulation of recommendations including determining which measures should be incorporated into a minimal dataset of measures for clinical trials and which ones need revision and further validity and reliability testing before use.

Conclusions: A 0–10 Point Numerical Rating Scale (NRS) is recommended as the outcome measure for pain intensity after SCI, while the 7-Point Guy/Farrar Patient Global Impression of Change (PGIC) scale is recommended as the outcome measure for global improvement in pain. The SF-36 single pain interference question and the Multidimensional Pain Inventory (MPI) or Brief Pain Inventory (BPI) pain interference items are recommended as the outcome measures for pain interference after SCI. Brush or cotton wool and at least one high-threshold von Frey filament are recommended to test mechanical allodynia/hyperalgesia while a Peltier-type thermotester is recommended to test thermal allodynia/hyperalgesia. The International Association for the Study of Pain (IASP) or Bryce-Ragnarsson pain taxonomies are recommended for classification of pain after SCI, while the Neuropathic Pain Scale (NPS) is recommended for measuring change in neuropathic pain and the Leeds Assessment of Neuropathic Symptoms and Signs (LANSS) for quantitating neuropathic and nociceptive pain discrimination.

J Spinal Cord Med. 2007;30:421–440

Key Words: Pain, chronic, classification, neuropathic, non-neuropathic; Treatment outcome; Spinal cord injuries; Pain scales; Reproducibility of results

INTRODUCTION

In clinical trials of the treatment of pain after spinal cord injury (SCI) there is a lack of consensus, both as to which constructs (defined as non-directly observable entities that one wishes to measure) need to be incorporated and which specific instruments or scales (defined as the tools whose primary function is to measure these specific entities) should be used. If different investigators measure different constructs or measure the same constructs using different instruments, it is difficult to compare the results

of the studies and to accumulate knowledge using systematic review.

For chronic pain in general, the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) resulted in the formation of an ad hoc committee of experts in pain representing academia, the pharmaceutical industry, and governmental agencies. This pain committee developed 6 specific core outcome domains which the committee recommended be assessed in all clinical trials concerned with the treatment of chronic pain (1). These 6 domains include pain itself in its sensory, affective, and evaluative components; physical functioning; participant ratings of global improvement and satisfaction with treatment; emotional functioning; symptoms and adverse events; and participant disposition. In addition, the group has recommended that specific measures within these domains be used in all clinical trials (2). Similar types of recommendations for neuropathic pain research have been made by the European Federation of Neurological Societies (EFNS) (3).

For research on individuals with SCI, the motor and sensory impairments of SCI provide circumstances which need to be taken into consideration when choosing an appropriate instrument to evaluate a given construct. For example, for a person with high-level tetraplegia and impaired hand function, the standard version of a visual analog scale may not be the most accurate instrument for measurement of pain intensity. Also, asking a subject with SCI a question about pain interference with walking, a common item in many pain interference measures, is not likely to be useful for evaluating pain interference for someone who uses a wheelchair on a regular basis. These are just two examples of problems with the applicability of specific instruments in evaluating outcomes for persons with SCI.

The purpose of this article is to review briefly the concepts and techniques used in the evaluation of the quality of measurement instruments, to present the results of a systematic review of outcome measures relevant to chronic pain after SCI, and to offer recommendations regarding the appropriateness and readiness of select outcome measures for use in clinical trials relating to pain after SCI. The outcome measures chosen for evaluation represent constructs thought necessary to be evaluated in clinical trials of pain after SCI. It should be noted, however, that the list of constructs is limited to pain proper and does not include those which are not directly related to pain such as quality of life, emotional functioning, and physical functioning, which the authors recognize as important for evaluation in a clinical trial of pain after SCI. These are beyond the scope of this article.

BACKGROUND

Basic Psychometric Concepts

Error is inherent in measurement. The error may be small (as might, for instance, occur in recording a subject's age) or large (as might occur when quantifying an abstract concept such as the affective impact of pain). The paragraphs in this section summarize a few of the many existing concepts and techniques used in evaluating the quality of measurement instruments: reliability, validity, sensitivity, and practicality. For a more in-depth treatment, the reader is referred to introductory (4–7) and more advanced (8–13) journal articles as well as textbooks (14, 15).

"Validity" refers to the question: is this instrument measuring what it purports to measure? If it is targeting characteristic X, do the numbers that result from the measurement operation actually reflect X rather than characteristic Y. "Reliability" refers to the question: how reproducible is this measurement if the thing being measured does not change? If the measurement operation is repeated with the same or an equivalent "ruler," would we get the same result? An instrument can be very reliable without being valid. If it is not reliable at all, by definition it cannot be valid. The goal we are aiming for is instruments that are both valid (they measure what we want to measure) and reliable (they give results that are reproducible). Finally, it should be kept in mind that there is only one validity and one reliability, which are estimated using different techniques, some of which are described below.

Reliability

All methods of estimating the reliability of measures are based on some form of repeat measurement. If two

*In this paper, experts in the area of spinal cord injury chronic pain critically evaluated the quality of a number of instruments used to quantify SCI pain and its experience, based on the published literature. In many instances, one or more authors developed the measures evaluated or contributed to the literature used. In writing this review, conflicts of interest were avoided to the degree possible. At a minimum, an author without conflict of interest contributed to and critically reviewed each summary of a specific measure or instrument.

[1]Rehabilitation Medicine, The Mount Sinai Medical Center, New York, New York; [2]Spinalis SCI Unit, Karolinska University Hospital, [3]Dandyryds Hospital Karolinska Institutet, Stockholm, Sweden; [4]Department of Rehabilitation Medicine, University of Miami, Miami, Florida; [5]Miami VA Medical Center and The Miami Project to Cure Paralysis, University of Miami, Miami, Florida; [6]Danish Research Pain Center, Aarhus University Hospital, Denmark; [7]University of Oxford, Oxford, UK; [8]Department of Physical Medicine and Rehabilitation, University of Alabama, Birmingham, Alabama; [9]VA Medical Center, Houston, Texas; [10]Pain Management Research Institute, University of Sydney, Royal North Shore Hospital, Sydney, Australia.

Please address correspondence to Thomas N. Bryce, MD, Department of Rehabilitation Medicine, The Mount Sinai Medical Center, 5 East 98th Street, 6th floor, Box 1240B, New York, NY 10021; phone 212.241.6321; fax: 212.369.6389 (e-mail: Thomas.bryce@mssm.edu).

clinicians at the same time rate the pain behavior of patient X using instrument Y, they should come up with the same number. If they do not, one or both are wrong. A statistical formula such as coefficient kappa can be used to express the agreement between the two. These formulas are constructed in such a way that the result, the reliability coefficient, varies between 0.00 (no reliability whatsoever) and 1.00 (perfect reliability). "Interrater reliability" can be estimated by having two (or more) raters assess the same group of subjects. If the same subjects are rated by the same clinician twice, we similarly can calculate "intra-rater reliability," the degree to which the clinician agrees with her earlier ratings. "Test-retest reliability" can be calculated when the same instrument is administered twice to the same subjects, before the characteristic of interest has changed.

Instruments designed to measure an abstract entity (a "construct," in psychometrical parlance), such as pain interference with functioning, typically consist of multiple indicators (items), each of which is assumed to represent the construct to some degree. The scores on the items are combined to adequately operationalize the theoretical definition we may have; this has the additional advantage of offsetting any random measurement error involved with quantifying any one item. Because each item in the instrument is a repeat measurement of the construct, we can calculate the agreement between items as yet another estimate of reliability. A number of formulas to estimate this "internal consistency reliability" exist, the most frequently used of which is (Cronbach's) coefficient alpha. "Split-half" and "parallel forms" reliability are related formulas. All of them take values between 0.00 and 1.00.

The minimal reliability a measure needs to have depends on how the results of the measurement are to be used. A minimum of 0.90 for situations where decisions on an individual patient need to be made is often quoted, while 0.70 or 0.80 is a minimum typically required for group applications, such as in program evaluation and research.

Validity

Validity cannot be estimated as simply as reliability, except in one unusual situation: there is an existing instrument that we are certain is perfectly valid. In that case, we can administer the old instrument and the new one to a sample, calculate the correlation between the two scores, and use that correlation as the estimate of the validity of the new measure. Developing an instrument that is shorter than the "gold standard" may be the only situation in which this occurs, and where we can quantify "criterion validity." Less powerful methods are used in the more common situation: there is no existing gold standard measure, or the existing ones are problematic in themselves. Such terms (and procedures) as "face," "concept," "criterion," and "predictive" validity may come into play.

"Face validity" is (in the eyes of some authorities) not a form of validity determination, but an answer to the question: does the instrument "on the face of it" measure what those completing it expect to see? Does a measure of trait X actually have questions about X that subjects recognize as such? The closely related term "content validity" refers to a measure covering the *entire* construct the developer is targeting, in the eyes of experts.

"Predictive validity" concerns the ability of a measure to predict a future state or event that is inherently linked to the characteristic being measured. A college entrance examination is said to have predictive validity if it can be used to accurately predict who in 4 (5, 6) years will graduate. There are no hard and fast rules as to what should be the minimum level of success in prediction. "Known group validity" or "discriminant validity" is based on differences in scores between two groups that are known to differ in the characteristic the instrument used aims to measure. People with severe pain should report higher pain interference levels than those with mild pain. If the data do not parallel these expectations, the pain interference instrument is probably not measuring what we think it is. Alternatively, significant systematic error (bias) is reflected in the data.

"Construct validity" concerns the relationships between the measurement data for a (highly abstract) construct and data for other constructs. We may have a basis in theory to predict that construct K should be strongly related to (yet not identical with) construct L, and be independent of construct M. If the data are consistent with this prediction, the operationalization of K should be valid (and similarly the operationalizations of L and M). If the predicted association between K and L is minimal or absent, however, we do not know if the problem is with the theory, the operationalization of K, or the measurement of L.

The terms "clinical validity," "prescriptive validity," and "consequential validity" have been used to designate the ability of an instrument's data to change the management of patients: the extent to which the scores really mean something that affects decisions about care.

Estimating the validity of instruments or, more properly, the validity of the data produced by instruments, is always less straightforward than quantifying their reliability. All methods of validity estimation are roundabout, and in practice it is necessary to use all possible methods of estimating validity, and "patch together" multiple findings supporting validity.

Sensitivity and Responsiveness

If a pain severity measure has just two categories: "no pain" and "pain," it lacks sensitivity: it cannot reflect fine distinctions in experienced pain severity, and it cannot be used to record minor but clinically significant changes in pain levels. Sensitivity refers to the ability of an instrument to capture distinctions that are clinically relevant or small enough to be of importance in research,

across the full range of the cases to be measured. When sensitivity is discussed in relation to change, the term "responsiveness" is frequently used.

Floor effects and ceiling effects are one issue in sensitivity. If the range of subjects' levels on characteristic X is wider than can be captured by an instrument developed to measure X, those falling outside the measured range are perforce assigned to either the highest or lowest measured level. In other situations, the sensitivity problem may be one of too few intermediate points on a scale. A variety of indices are used to determine responsiveness, including effect sizes (the mean change between time 1 and time 2 divided by the standard deviation at time 1), the standardized response mean (the mean change between time 1 and time 2, divided by the standard deviation of change scores), receiver operating characteristic (ROC) analysis, and many others. These indices are mostly useful for comparing the responsiveness of one measurement instrument with that of another, allowing one to select the most responsive one.

Practicality

Beyond validity, reliability and sensitivity, there are a number of other characteristics of measurement instruments that are relevant to their use in clinical, program and research applications, most of which have to do with practicality. These include language; training, time, and equipment requirements; disability adjustment; and norms.

Language is especially relevant to self-administered instruments, but may also be an issue with observational and other measures. Both the reading level and translation in a language the user is familiar with are of concern. Translations should be done using back-translation, and should be shown to have linguistic and functional equivalence with the original (an issue of validity).

With regards to training, time, and equipment requirements, many observational instruments and test-type measures require the user to be trained, and sometimes certified, in order to produce reliable data. Measures that take an inordinate time on the part of the subjects or the administrator, or that use special equipment, may not be suitable outside research applications.

With regards to disability issues, some existing instruments developed for the non-disabled population have shortcomings when used with individuals with SCI, which may affect feasibility (eg, people with tetraplegia cannot draw the line required for VAS assessment of pain severity) or validity (eg, in answering pain interference questions, can subjects distinguish interference due to pain from interference due to paralysis?).

Finally, if norms are available (either for the population at large, or for people with SCI) it is easy to determine how far removed from "average" or "typical" a particular subject's score is, taking such characteristics as age, gender or injury level into account, if appropriate.

METHODS

Under the auspices of the National Institute on Disability and Rehabilitation Research (NIDDR) in the Office of Special Education and Rehabilitation Services, US Department of Education, a committee consisting of experts on the subject of outcome measures for clinical trials related to SCI, including psychometricians and other methodologists, and employees of NIDRR was formed with the goal of evaluating the state of the science with regard to the currently available outcome measures and their readiness for use in clinical trials related to SCI. Several areas of study were identified by this group, including motor and sensory function, functional outcomes, imaging, quality of life, and pain. For each of these areas subcommittees were formed, composed of experts from around the world, including within each subcommittee a methodologist. In addition a committee of methodology experts was formed to develop a common template which could be used during the evaluation of the instruments by all sub-groups. There were two templates developed, one for evaluation of psychosocial measures and another for evaluation of physical measures.

The committee on pain measurement consisted of the 12 authors who are academically based pain researchers from the United States, United Kingdom, Sweden, Denmark, and Australia. The committee first determined by consensus the constructs thought necessary to be evaluated in clinical trials evaluating the treatment of pain after SCI. These constructs were then assigned to subcommittees who searched the literature using the MEDLINE, EMBASE, CINAHL, and PubMed databases, identifying those measures of the constructs that had been used by at least two different research groups for the evaluation of persons with SCI during the preceding 5 years. The quality of these measures was further evaluated by examination of reliability and validity data that had been published since each measure was first introduced. The reliability and validity data were extracted primarily from studies on the SCI population but also from selected research studies on non-SCI populations when data for SCI samples were lacking. The main criteria used in evaluating the outcome measures selected were: (1) availability of a disability adapted version, if needed, (2) burden of administration, (3) internal consistency, (4) reliability, (5) inter-rater, test-retest, and other reliability/reproducibility characteristics, (6) bias, (7) sensitivity to change, (8) ceiling/ floor effects, (9) clinical utility, (10) applicability in SCI (vs other groups), (11) extent of use in SCI, (12) predictive (and discriminant) validity, and (13) concurrent validity.

The following constructs (and relevant measures) were included: (1) Pain intensity: Visual Analogue Scale (VAS), Verbal Rating Scale (VRS), Numerical Rating Scale (NRS); (2) Global improvement of pain: Patient Global

Table 1. Subject Demographics

Characteristic	% (n)
Continent of origin	100 (60)
US	45 (27)
Other North America	8 (5)
Europe	35 (21)
Asia	3 (2)
Australia	5 (3)
Africa	3 (2)
Profession	100 (59)
Physician	53 (31)
Psychologist	10 (6)
Physical therapist	5 (3)
Occupational therapist	7 (4)
Nurse	5 (3)
Other	20 (12)
Years in profession	100 (63)
0 to 5	8 (5)
6 to 10	16 (100
11 to 20	30 (19)
$\geq$ 21	46 (29)
Published on pain and SCI	100 (63)
Yes	25 (16)
No	75 (47)

Impression of Change (PGIC); (3) Meaningful reduction in pain; (4) Sensory, affective, and evaluative aspects of pain as assessed in multidimensional pain measures: McGill Pain Questionnaire (MPQ); (5) Cognitive, affective, social, and behavioral responses to pain: West Haven-Yale Multidimensional Pain Inventory (WHYMPI/MPI); (6) Punctate (or pinprick) hyperalgesia and allodynia: Von-Frey hair/monofilaments; (7) Dynamic mechanical (tactile) allodynia: Brush; (8) Thermal allodynia and hyperalgesia: Hot pain and cold pain threshold and supra-threshold ratings; (9) Classification of pain after SCI: International Association for the Study of Pain (IASP) SCI Classification, Bryce/Ragnarsson SCI Pain Taxonomy (BR-SCI-PT), and the Cardenas SCI Pain Classification; (10) Discriminating between neuropathic and non-neuropathic pain: Douleur neuropathique 4 questions (DN-4), Leeds Assessment of Neuropathic Symptoms and Signs (LANSS), Neuropathic Pain Questionnaire (NPQ); (11) Pain qualities associated with neuropathic pain: Neuropathic Pain Scale (NPS), Pain Quality Assessment Scale (PQAS), and Neuropathic Pain Symptom Inventory (NPSI); (12) Interference of pain with activities, mood, relationships, and life: Graded Chronic Pain Scale (GCPS), SF-36 Pain interference item, MPI Life Interference subscale, and Brief Pain Inventory (BPI) Pain Interference subscale; and (13) Cognitive and behavioral pain coping strategies: Coping Strategies Questionnaire (CSQ).

After reviewing the articles, at least one investigator extracted pertinent information utilizing the appropriate template guide, which included all of the criteria noted above (reliability, validity, etc.). A second investigator reviewed the template completed by the first one and the reference materials noted, in order to confirm the findings of the first investigator. The investigators discussed discrepancies (if any) and came to consensus with regard to each measure. Taking into consideration the extent of evaluation in persons with SCI and both the quantity and quality of the studies analyzed, judgments on reliability, validity, and applicability of the measures were made by the investigators. Based upon these judgments, recommendations were formulated for the use of specific measures in future clinical trials.

A precourse of the 2006 combined American Spinal Injury Association (ASIA)/International Spinal Cord Society (ISCoS) scientific meeting in Boston on the topic outcome measures for clinical trials sponsored by NIDDR was held on June 24th, 2006. During this precourse, which attracted approximately 500 participants, a subset of the outcome measures were presented along with available reliability and validity data. The measures that were presented included measures of the severity of pain (VAS, VRS, NRS), a measure of the global improvement of pain (PGIC), multidimensional measures of pain (MPQ and MPI), classifications of pain (IASP, BR-SCI-PT, Cardenas), measures of pain interference (GCPS, SF-36, MPI, BPI), and a measure of coping (CSQ). Following these presentations, a series of questions was posed to the audience. One question referred to whether or not a specific measure should be part of a recommended dataset for clinical trials related to pain after SCI. The conference participants could vote on multiple choice answers using an anonymous electronic audience response system. Table 1 lists the demographics of the conference attendees who voted on the various measures. The choices included are listed along the column headings of Tables 2 and 3. In addition, questions were posed to the participants regarding ranking of choices for minimum datasets as seen in Table 4.

RESULTS

Unidimensional Measures

Visual Analogue Scales. Visual analogue scales (VAS) most often consist of a 10-cm long line, but also have been reported in a 15-cm version (16, 17). Anchor labels vary from study to study but typically one end represents '*no pain*' and the other '*worst possible pain.*' Visual analogue scales have been reported not only for measuring pain intensity but also for the affective-motivational component of pain (ie, pain unpleasantness) (16, 17). Studies have reported that a VAS is a valid scale for measuring both pain intensity and pain affect when compared to other unidimensional pain scales in correlational analyses (18–21). Visual analogue scales have been reported to be sensitive to change in pain

Table 2. Audience Voting on Instrument Validity and Usefulness

Instrument and number of experts voting	It is a valid measure and should be part of a **minimum dataset** for clinical trials related to pain after SCI (ie, suggested for all clinical trials) % (n)	It is a valid measure but should be part of an **expanded dataset** only % (n)	It needs **further study** to establish its **reliability and validity** before being part of any recommended dataset for SCI % (n)	It needs **revision** followed by **further study** of **reliability and validity** before being part of any recommended dataset for SCI % (n)	It is not valid or relevant for use in SCI research % (n)
Pain Intensity					
VRS (n = 51)	25 (13)	27 (14)	41 (21)	N/A	6 (3)
NRS (n = 50)	64 (32)	14 (7)	20 (10)	N/A	2 (1)
VAS (n = 58)	14 (8)	38 (22)	45 (26)	N/A	3 (2)
Subject impression of change					
PGIC (n = 57)	23 (13)	14 (8)	48 (33)	N/A	5 (3)
Multidimensional Measures					
LF-MPQ (n = 53)	6 (3)	19 (10)	25 (13)	43 (23)	8 (4)
SF-MPQ (n = 49)	2 (1)	2 (1)	29 (14)	61 (30)	6 (3)
MPI (n = 50)	6 (3)	20 (10)	12 (6)	48 (24)	14 (7)
MPI-SCI (n = 48)	19 (9)	19 (9)	48 (23)	15 (7)	0 (0)
Pain Interference					
BPI (n = 44)	0 (0)	7 (3)	41 (18)	41 (18)	11 (5)
SF-36 (n = 52)	12 (6)	12 (6)	37 (19)	27 (14)	13 (7)
GCPS (n = 44)	5 (2)	7 (3)	20 (9)	45 (20)	23 (10)
MPI (n = 47)	9 (4)	4 (2)	36 (17)	45 (21)	6 (3)
Pain Coping					
CSQ (n = 52)	4 (2)	15 (8)	44 (23)	29 (15)	8 (4)

Table 3. Audience Voting on Pain Classification Validity and Usefulness

Instrument and number of experts voting	Valid and useful % (n)	Useful but requires more validation % (n)	Useful but requires changes/improvement then further validation % (n)	Not useful or valid for research in SCI % (n)
Pain Classification				
IASP (n = 59)	19 (11)	47 (28)	31 (18)	3 (2)
BR-SCI-PT (n = 59)	14 (8)	42 (25)	36 (21)	8 (5)
Cardenas (n = 56)	4 (2)	20 (11)	52 (29)	25 (14)

intensity (16, 17, 22, 23) and in pain affect (16, 17). Sensitivity to change has been higher for a VAS compared to a 4-point VRS (18, 22, 23) and similar to that of an 11 point NRS (18). However, a 7-point VRS has been reported to be more sensitive than the VAS (24).

Reproducibility of pain intensity and pain affect ratings have been reported to be high in persons with chronic pain and myofascial pain dysfunction (16, 17, 25). However, low test-retest agreement has been reported after SCI (26). Visual Analogue Scales have been reported to have high failure rates, higher than for NRS and VRS. Failure rates from 5.3 to 22% have been reported (20, 21, 24, 27). This has been related to increased age (20, 27) although when validity and reliability were assessed in healthy younger vs older adults, age was not found to impact failure (24). Instead, motor impairment was the only statistically significant variable associated with scale failure. Another study found opioid intake to be related to scale failure (21).

Finally, persons with tetraplegia and decreased hand function can have difficulties using the standard version of VAS properly, although this has not been studied thoroughly and visual inspection with a point or tell approach could theoretically be used during a face-to-face query.

Verbal Rating Scales. There is a wide range of different verbal rating scales (VRS), ordered categorical scales, and descriptor scales described in the literature. In this review we have included scales with 4 to 7 categories. One of them is adopted from the McGill Pain Questionnaire, but most are independent scales not linked to other measures. Verbal rating scales have been reported to have good validity compared to other unidimensional pain intensity rating scales in correlational analysis' both in acute pain and in chronic pain (18–21, 24, 26). Internal consistency was reported to be good in an experimental design for a 7-point VRS which was similar to that of NRS and VAS (24). Also, good construct validity has been reported (19, 24). One study has assessed the relation between a VRS and a VAS in persons with SCI-related neuropathic pain (26) and found a level of concordance of 0.88. In the same study test-retest reliability was 100%.

Seven-point VRS have been reported to be more sensitive to change than both VAS and NRS (24) but 4-point VRS have been shown to be less sensitive than VAS (18, 22, 23) and NRS (18). Failure rates are reported to be null (20, 21, 24) or low (27). The number of categories used seems important when using VRS. Four-point VRS have shown less sensitivity than VAS and NRS that have more response categories. However more categories are not necessarily associated with higher sensitivity (27).

Another important issue is that of language and translation. Verbal rating scales are dependent on an interpreter's translation into each language to a greater degree than VAS or NRS. Use of a standard version with fixed labels can minimize bias, although persons not fluent in the actual language can have difficulties understanding the differences in the labels.

Numerical Rating Scales. Numerical rating scales (NRS) or Likert scales can have a range of 0–10, 0–20

Table 4. Audience Voting on Preference for Inclusion within Dataset

	1st choice for minimum dataset % (n)	2nd choice for minimum dataset % (n)
Pain Intensity	100 (57)	100 (53)
VRS	16 (9)	55 (29)
NRS	79 (45)	8 (4)
VAS	5 (3)	38 (20)
Subject impression of change	100 (53)	
Original Guy/Farrar PGIC	10 (5)	N/A
Modified Guy/Farrar PGIC I	24 (12)	N/A
Modified Guy/Farrar PGIC II	67 (34)	N/A
Pain Interference	100 (49)	100 (41)
BPI	16 (8)	20 (8)
SF-36	39 (19)	29 (12)
GCPS	6 (3)	7 (3)
MPI	31 (15)	22 (9)
Other measure not listed	2 (1)	5 (2)
None should be included	6 (3)	17 (7)

and 0–100. The instrument can either be used verbally or as a paper and pen version, the latter commonly as a box-scale. Anchors vary from scale to scale although typically one of the anchors is labeled '*no pain*' and the other anchor is labeled '*worst possible pain.*' If pain affect is measured, '*no unpleasantness*' and '*worst possible unpleasantness*' are typical anchor labels. The lack of uniform anchors limits comparisons between trials.

Validity has been established for NRS (18–21, 24) as well as sensitivity to change (18, 24). Construct validity has been reported to be good both in an experimental design and in persons with rheumatic disease (19, 24). Failure rates have been reported to be low on the NRS, varying between 0 and 5.3% (20, 21, 24, 27).

Recommendations. Numerical rating scales are recommended over the other unidimensional pain measures by the committee and were the first choice for inclusion in a minimum dataset for clinical trials of pain after SCI by the participants of the ASIA/ISCOS Measurement precourse. We recommend using a verbal scale based on the premise that persons with tetraplegia can have difficulties with a paper and pen version and a verbal instrument would facilitate telephone follow-ups/surveys. The instrument needs fixed anchor labels and we endorse the IMMPACT recommendations (2) (ie, '*0 = no pain*' and '*10 = pain as bad as you can imagine*') accompanied by the instructions "*please rate your pain by indicating the number that best describes your pain on average in the last 24 hours.*"

We also recommend that '*pain at its least*' and '*pain at its worst*' accompanied by the instructions "*please rate your pain by indicating the number that best describes your pain as its least/its worst during the last week*" be considered as adjunct measures. Furthermore we recommend rating pain unpleasantness, using '*0 = no unpleasantness*' and '*10 = unpleasantness as bad as you can imagine*' accompanied by the instructions "*please rate your pain unpleasantness by indicating the number that best describes your unpleasantness on average in the last 24 hours*" be considered as an adjunct measure. We also wish to emphasize that NRS pain ratings produce ordinal data and should be treated as such.

Patient Global Impression of Change Scales

Jensen et al (23) stated that "pain relief is something more than just change in pain intensity" and pain relief has been considered to "assess the patients' overall status since starting the study, integrating the effect of treatment, side-effects, and patient expectations" (28). Therefore, a scale for measuring the global treatment effect can be used as a complement to the unidimensional pain intensity scales.

There are several scales that have been referred to as (Patient) Global Impression of Change scales. Amongst these are a 5-grade scale from '*marked improvement* through *marked worsening*' (29), a 5-grade scale with the end-points '*much better and much worse*' (30), a 7-grade scale using a range from '*no change to a great deal better*' (31), and a 7-grade scale using '*very much improved–very much worse*', also referred to here as the original Guy/Farrar-PGIC (28, 32).

No formal validity or reliability testing have been carried out but some of the PGIC scales have been compared to pain intensity NRS and VAS changes in order to determine clinical significance of change. A clinically significant change (much better or much or very much improved) is in these studies equal to a decrease of 2 units or more on the NRS (28, 30) and to a 3-cm reduction on a 10-cm VAS (33). Farrar et al (28) concluded that the baseline pain scores also affected the level of PGIC achieved such that higher initial scores needed larger raw score changes to result in a global "improvement" rating.

Transitional PGIC scales have been reported to moderately correlate to VAS change, Spearman $r = 0.67$ (34), with a high discordance and large standard deviations for the mean change in VAS scores. Another study reported a correlation coefficient of $r = 0.77$ for NRS change (30).

Recommendation. When 3 different versions of the Guy/Farrar PGIC (Table 5), the original 7-grade Guy/Farrar PGIC and two modified versions, the Guy/Farrar-PGIC I and II, were presented to the participants at the ASIA/ISCOS Measurement precourse, they preferred the second modification to that of the original Guy/Farrar 7-point PGIC and to the modified 9-point PGIC, by a significant margin (Table 4). The reason for suggesting modifications of the PGIC is based on reports from persons with SCI who have neuropathic pain that there is too large a gap between rating pain '*minimally*' vs '*much improved*' (Norrbrink Budh and Lundeberg personal experience) and that 4-point (in one direction) verbal ratings scales have demonstrated low sensitivity (18, 22, 23). However, since neither of the modified scales has been validated in any population and the original Guy/Farrar PGIC has been extensively used and shown sensitivity to change, the committee recommends the original Guy/Farrar-PGIC for use in clinical trials.

Multidimensional Pain Measures

McGill Pain Questionnaire. The McGill Pain Questionnaire has been used extensively, in a long-form (LF-MPQ) (35) and in a short-form (SF-MPQ) (36), to measure multidimensional aspects of the pain experience. The LF-MPQ consists of 78 descriptors and accompanying ratings, grouped into sensory, affective, and evaluative categories (which can be scored separately or together), and a Present Pain Intensity (PPI) measure consisting of a 6-point scale from "no pain" to "excruciating." The SF-MPQ consists of the 15 most-frequently used descriptors (sensory and affective) from the LF-MPQ and corresponding ratings of the descriptors, as well as the PPI and a VAS. Both versions of the MPQ are administered by the researcher or physician or can be self-

Table 5. Versions of Patient Global Impression of Change

Original Guy/Farrar-PGIC	Modified Guy/Farrar-PGIC I	Modified Guy/Farrar-PGIC II
Very much improved	Very much improved	Much improved
Much improved	Much improved	Moderately improved
Minimally improved	Moderately improved	Minimally improved
	Minimally improved	
No change		No change
	No change	
Minimally worse		Minimally worse
Much worse	Minimally worse	Moderately worse
Very much worse	Moderately worse	Much worse
	Much worse	
	Very much worse	

administered. They require a relatively limited amount of time (5–20 minutes), and have been translated into many languages (37). The LF-MPQ and SF-MPQ are comparable with relatively high correlations (0.68 to 0.93) between corresponding sub-scales (36, 38). Use of the SF-MPQ in the SCI population has prompted the addition of some descriptors appropriate for SCI pain (39, 39–41).

Internal consistency for the SF-MPQ (Cronbach's alpha between 0.70 and 0.73) (42, 43) and test-retest reliability (intraclass correlations between 0.70 and 0.97 for repeat administration between 5 and 15 days) (44, 45) are satisfactory in persons with non-SCI pain. Several studies indicate adequate sensitivity of the MPQ to change in pain experience when surgical or pharmacological treatments have been introduced (36, 45–47). Three studies have examined the ability of the MPQ to differentiate between pain types in persons with SCI. One of these studies found significant differences in some subscores of the MPQ between persons with SCI and allodynia compared to those without allodynia (48). However, two other studies suggested limited usefulness of the MPQ to distinguish between specific pain types within the SCI population (40, 41).

Although the MPQ has been used extensively and may serve a purpose in general comparisons between SCI pain and pain in other populations, the usefulness of the present versions of the LF-MPQ or the SF-MPQ specifically for use in evaluating pain in the SCI population is limited. It is suggested that SCI-appropriate revisions to this assessment be developed and the reliability and validity of this revised MPQ be assessed before any recommendation for its use in the target population can be made.

The Multidimensional Pain Inventory. The Multidimensional Pain Inventory (MPI) (49) was designed to assess pain and a range of self-reported behavioral and psychosocial factors associated with the impact of chronic pain syndromes on physical functioning, emotional functioning, and responses from significant others to the presence of pain. The MPI has been shown to have excellent psychometric properties and the factor structure has been confirmed in several studies (50–52). The MPI is widely used in clinical practice and research and is one of the core outcome measures recommended by the IMMPACT group for clinical pain trials (2). The MPI consist of 3 sections: Pain Impact, Responses by Significant Others, and Activities. Eight of the 12 subscales measure cognitive, affective, social, and behavioral responses: Pain Severity (PS), Life Interference (LI), Life Control (LC), Affective Distress (AD), Support (S), Negative Responses (NR), Solicitous Responses (SR), and Distracting Responses (DR). The other subscales assess the degree of participation in various types of daily activities (household, away from home, social, and outdoor). These are frequently combined into a single General Activity scale (GA).

Although the MPI was originally developed for heterogeneous chronic pain populations, it was later adapted for use with the SCI chronic pain population, resulting in an instrument designated the MPI-SCI (53). In the MPI-SCI the following modifications were made in order to improve the fit of the factor structure: (1) removal of 3 items from the LI and one item from the LC; (2) removal of two items in the Responses by Significant Others section; (3) adding one question per item in the GA scale addressing whether level of activity was decreased due to pain as distinct from restrictions of activity due to other aspects of SCI. The MPI-SCI consists of 50 items and takes about 15 to 20 minutes to administer. Its psychometric properties were recently examined in 161 individuals with SCI and chronic pain (54). All MPI-SCI subscales had moderate to substantial internal consistency reliability (Cronbach's alpha coefficients = 0.61–0.94) except for Affective Distress, which had fair reliability (Cronbach's alpha = 0.60). The test-retest reliability of the MPI-SCI in general was adequate with test-retest reliability in the range of

moderate to substantial, except for the Support and Life Control subscales.

The MPI-SCI may be a useful measure for the evaluation of chronic pain impact following SCI. The strengths of the MPI are its brevity, ease of administration, subject acceptance, and demonstrated utility in multiple clinical and research investigations. Future research should continue to refine the psychometric properties of the Affective Distress, Life Control, and Support subscales of the MPI-SCI. The MPI-SCI has not yet been widely used in the SCI population. However, several international studies using this measure are underway.

Recommendations. Use of the MPI-SCI is recommended over other measures by the committee and by the participants of the ASIA/ISCOS Measurement precourse. While 38% of the participants of the precourse considered the MPI-SCI to be a valid measure for use in the SCI population, 48% felt that the MPI-SCI needed more study to further establish reliability and validity. In comparison, the current versions of the LF-MPQ and the SF-MPQ were viewed as valid for use in SCI by 25% and 4% of participants, respectively.

Psychophysiological Measures

In contrast to self-report measures of spontaneous pain, psychophysiological measures rely on a report by subjects on sensations experienced (or not experienced) as the result of stimulation of a particular kind by means of an instrument wielded by the investigator. The subject response can be of the yes/no type (yes vs not detected) or on a NRS or similar measure of the strength of the experience (eg, pain severity). Generally, for psychophysiological measures the construct measured (eg, allodynia) is yoked with the instrument used to evoke the experience (eg, a brush), and the psychometric information reported here focuses on the combination (the instrument and unidimensional measure) because the stimulus instrument plays such a crucial role. These types of measures are essential for evaluating and quantifying pain which is not spontaneous but evoked. Evoked pains are not uncommonly seen after SCI (55, 56). There are several types of evoked pain. During sensory examination, if pain is elicited with a stimulus that does not normally provoke pain, this evoked pain is often defined as allodynia. This evoked sensation or pain can also be qualified in relation to how the stimulus is applied, (ie, dynamic; a monofilament brushed along the skin), (or static; a monofilament touched once at a single location on the skin). In addition, a touch or pressure stimulus can be qualified as mechanical, while a hot or cold stimulus can be qualified as thermal. The pain threshold is defined as the lowest intensity of a stimulus at which a subject experiences pain. Hypoalgesia is an increased pain threshold, compared to the average person or compared to other areas of a patient's body. Hyperalgesia is a decreased pain threshold (an "exaggerated" painful response to a pain provoking stimulus).

Von Frey hair/monofilaments for pinprick hyperalgesia and allodynia. Von Frey filaments are used to determine touch sensibility and have in recent years also been used in pain research. When originally introduced in 1896 by von Frey, the filaments were made of human or animal hair (57). Filaments used today are made of nylon (eg, Semmes Weinstein Monofilaments, Stoelting Co IL USA or Aesthesiometer, Somedic Sweden) or optic glass (eg, Optihair, Marstock Nervtest, Germany). The force produced when bending a von Frey filament depends on stiffness (and therefore diameter) and length and is relatively independent of the degree of bending. The filaments are calibrated, the force being proportional to the diameter. The filament identification number represents the logarithm of 10 times the buckling force in mg (calculated force). The filaments are applied perpendicularly to the skin by the examiner who applies a force until the filament bows. The filaments are simple to use and relatively inexpensive. A single high-threshold filament can be used to detect pinprick hyperalgesia, while a complete kit of von Frey filaments can be used for obtaining sensory detection and pain detection thresholds and for stimulus-response functions.

Thresholds and pain intensity vary with age and site and over time. A normative sample or, preferably, a normal reference site in the same individual is necessary to diagnose pinprick hyperalgesia or allodynia. Lack of a normal contralateral side, as is often the case in persons with SCI, makes the test less sensitive. Von Frey filaments produce relatively reliable designated forces (58, 59), but in nylon filaments, the environmental humidity and temperature may influence the force required to buckle the filament (60). This is not the case for optic-glass fibers. Some nylon filament kits have a built-in thermo-hygrometer. If the tip is flat, the applied pressure may change during bending. Reliability and reproducibility studies of pain detection threshold and pain scores using von Frey filaments do not exist, and there is a lack of standardization for use in humans (61).

Brush for dynamic mechanical allodynia. Tactile allodynia is examined by lightly stroking the skin with a brush or cotton wool or another similar material. Few standardized tools are available (eg, Somedic brush).

A brush is handheld, which can cause variability in brushing force and stroking velocity. Because the method is not standardized, there is considerable variation from study to study. Increased length of area brushed and number of strokes significantly increase the total brush-evoked pain intensity, while the brush width does not seem to influence brush-evoked pain (62). Inter-rater reliability tested in one study on neuropathic and non-neuropathic pain ($n = 160$) showed an acceptable Cohen Kappa coefficient of 0.71 (63). The method has been shown to be sensitive to change in several randomized trials in SCI (64–67). The validity of the method is supported by

significant correlations between a person's verbal report and using a brush to evaluate tactile allodynia (68).

Thermal nociceptive thresholds and suprathreshold measures. Hyper- or hypo-sensitivity to thermal stimuli has generally been assessed using Peltier-type devices: the TSA II Neurosensory Analyzer (Medoc, Ltd; Israel) or the Modular Sensory Analyzer Thermotest (Somedic; Sweden). Generally, thresholds for pain are measured using the method of limits (gradual heating or cooling of the thermode from baseline until the subject detects hot pain or cold pain, respectively) and comparisons are made between affected sites and unaffected sites in the same individual, or similar sites in a normative sample. Thresholds greater than two standard deviations outside normative values are generally labeled abnormal (hypo- or hypersensitivity (allodynia), depending on the direction of deviation) (69). The time to administer tests depends on the number of test sites being evaluated, the attentiveness of the subject, and the number of trials (generally 2 to 4 measurements are averaged to determine the threshold at each test site).

Test-retest reliability for thermal pain thresholds has been demonstrated in healthy, able-bodied subjects (69, 70) and in a sample of persons with SCI (71). The sensitivity to change of thermal sensory measures is questionable (65, 67, 72): treatments that have had measurable effects on ratings of subjects' spontaneous pain are not consistently mimicked by changes in these evoked pain measures (although low power due to small sample sizes may be responsible). Further testing of reliability and validity in persons with SCI is necessary for thermal nociceptive measures before they are recommended as standard outcome measures in SCI clinical trials.

Recommendations. To evaluate mechanical allodynia/hyperalgesia, it is recommended to use a brush or cotton wool and at least one high-threshold von Frey filament (3). Validity, reliability, and standardization studies are needed in those with SCI, but these methods are simple and provide important information and, therefore, increased use is suggested. Thermal testing using a Peltier-type thermotester is recommended as a measure of decreased and/or increased thermal sensitivity. Further research on validity and reliability and standardization is needed. Clinical trials utilizing these measures will be helpful in studying the pathophysiology of pain in SCI.

Classification of Pain after SCI

International Association for the Study of Pain Taxonomy. The International Association for the Study of Pain (IASP) appointed a distinguished work group to propose what it hoped would become a standardized classification scheme for SCI pain. Building on the preliminary published work of one of its members (73), the outcome of their deliberations was first published in 2000 (74). They proposed a 3-tiered system with Tier I being composed of nociceptive and neuropathic subcategories. Once someone is classified at the Tier I level, a second Tier of subclassification occurs. Within the nociceptive category are two subtypes: musculoskeletal and visceral, and within the neuropathic category, 3 subtypes: above, at and below level pain. Further classification within each of those subcategories occurs, when possible, at the Tier III level, based on specific underlying structural causes of pathology. The authors point out that given our current state of diagnostic ability, it will not always be possible to demonstrate underlying pathology at the Tier-III level. Siddall and colleagues have begun using this classification scheme in publications documenting the development of SCI pain over time (75) and in proposed chronic SCI pain management approaches (76). There has only been one study examining inter-rater reliability of the IASP scheme (77). In this study, the authors used pairs of raters who independently rated pain subtype to the Tier II level using videotaped descriptions of pains. Agreement across pairs of raters (two physicians and a psychologist) ranged from 61% to 78%. When considering the two Tier-I subtypes: neuropathic and nociceptive, agreement across all three raters was 78%.

Bryce/Ragnarsson SCI Pain Taxonomy. Ragnarsson published an influential review of pain in SCI and its management in 1997 (78) and he and a colleague, Bryce, went on to publish essentially a variant of the IASP scheme in 2000, and have continued work on it since that time. The major difference between the Bryce/Ragnarsson SCI Pain Taxonomy (BR-SCI-PT) and the IASP scheme seems to lie in a reversal of tiers. Rather than first deciding whether the pain being evaluated is nociceptive or neuropathic, Bryce et al propose that the first distinction be whether the pain is above, at or below the level of the SCI lesion, and then further characterized as nociceptive or neuropathic as the second step (79). More detail, presumably based on the authors' clinical experience and the research literature, is provided about the various subtypes of pain. Some more specificity is added to the IASP definitions as well; for example defining "at level" as within two dermatomes above and below the neurologic level of injury (80). In a recent study, 39 physicians with SCI expertise classified 135 vignettes describing 179 separate pain sites for persons with SCI pain (80). These had been developed by Bryce and colleagues. Results of the physician raters were compared with the judgments of the authors which therefore represented a "gold standard" for comparison purposes. Agreement for the 15 subtypes of pain was 72%, with that percentage rising to 83% when the authors were convinced by the raters to reconsider their own initial ratings. Kappa values averaged 0.70 for all raters across pain subtypes, which the authors argue is likely an underestimate since it was based on original, and not revised ratings. Considering Tier I decisions (above, at or below level pain), 84% of initial judgments were correct with that percentage rising to 93% when

revised standards were applied. For Tier II decisions (neuropathic or nociceptive), 86% were correct (increased to 90% when modified standards were applied).

Cardenas SCI Pain Taxonomy. Cardenas and her group developed their own SCI pain classification scheme (81) for a funded study, prior to the publication of the IASP and Bryce-Ragnarsson schemes, although those publications were cited by the time the results of the study by Cardenas et al were published (40). The Cardenas group proposed 2 major categories: neurologic and musculoskeletal. Neurologic pain is divided into 4 subcategories: SCI pain, transition zone pain, radicular pain, and visceral pain. Musculoskeletal pain is divided into mechanical spine pain and overuse pain. The authors provide some descriptive information about these subtypes. A questionnaire designed to generate the data needed to derive these subtypes was returned by 163 persons with SCI. All pain sites were categorized by one investigator with 41 subjects presenting with 68 pain sites rated independently by a second person. In a related study, 15 subjects with SCI pain were interviewed in person independently by two raters. The kappa value for the paired questionnaire ratings was 0.68 across the 6 subtypes of pain, and 0.66 for the in-persons interviews.

Recommendations. All 3 systems have early promise in that substantial inter-rater reliability has been demonstrated across pairs of raters. This is true for the 5 Tier-II subcategories of the IASP scheme, the 15 subcategories of the Bryce-Ragnarsson scheme, and the 6 subcategories of the Cardenas et al scheme. The view of the participants at the ASIA/ISCoS meeting who reviewed these 3 schemes was that the IASP and Bryce/Ragnarsson schemes were useful in their present form but required more validation, and that the Cardenas scheme required changes and/or improvements and then further validation. Furthermore, any of the 3 could be used for clinical research studies with the caveat that decisions about pain subtype should be made independently by two raters, and subjects potentially excluded from the study when agreement does not occur. All 3 systems need additional work to establish reliability, specifically test-retest reliability, and validity. All 3 would benefit by moving to the next step of developing and testing a structured questionnaire and/or decision tree that would help make the derivation of pain subtypes easier. Lastly, it would be important for the field if the authors of these 3 systems and other experts were to combine intellectual forces to propose a pain classification system that could combine the best features of each, and further develop it to the point that universal acceptance for clinical and research purposes could occur.

Neuropathic Pain Measures

Douleur Neuropathique 4 Questions (Neuropathic Pain 4 Questions). The Douleur Neuropathique 4 Questions (Neuropathic Pain 4 Questions) (DN-4) (63) is a French instrument designed to discriminate between neuropathic and non-neuropathic pain. An English version exists, but its psychometric properties have not been reported. The DN-4 consists of 10 items, which fall into 4 categories: (a) pain descriptors (7 self-reported items), (b) association of paresthesia/dysesthesia within painful area (1 clinician examination item), (c) sensory deficits (1 examination item), and (d) evoked pain (1 examination item). Each of the 10 items is scored on a binary scale (eg, the pain can be described as "burning" or not). A total score ranging from 0 to 10 is calculated by summing the "Yes" responses. A cut-off score of 4 has been identified by the developers as indicating neuropathic pain. A score for only the 7 self-reported items also can be calculated and a cut-off score of 3 has been identified for that subscale. In one study, inter-rater reliability for the 3 clinician-examination items ranged from 0.71 to 0.78. Test-retest reliability on the self-report items over a 3-day period ranged from 86% to 98% (Kappa 0.70–0.96). Using diagnoses agreed upon by two clinicians as a gold standard and a cut-off score of 4, the full 10-item DN-4 correctly classified 86% of the pains (sensitivity = 83%, specificity = 90%). Using a cut-off score of 3, the 7 self-reported items correctly classified 79% (sensitivity 78%, specificity, 81%).

The French version of the DN-4, which was tested in France, has good psychometric properties but, to date, there is only one study that used it. The English version is yet untested. The DN-4 was developed with 167 persons from a multidimensional pain center and included only 5 persons with SCI for whom there is no specific information. Further testing is needed if the DN-4 in English or any other language is to be used in the US SCI population.

Leeds Assessment of Neuropathic Symptoms and Signs and Self-completed Leeds Assessment of Neuropathic Symptoms and Signs. The Leeds Assessment of Neuropathic Symptoms and Signs (LANSS) (82) is a 7-item instrument developed in the United Kingdom to discriminate between neuropathic and non-neuropathic pain. It assesses 5 types of pain: (a) thermal, (b) dysesthesia, (c) paroxysmal, (d) evoked, and (e) autonomic dysfunction. Five items are self-report and two involve clinician examination. Response options are binary (eg, the pain feels like pins and needles or it doesn't), but the items have different weights assigned to "Yes" answers, ranging from 2 to 5 points. The maximum score is 24 and a score of 12 or higher is considered indicative of neuropathic pain. In one study that used clinician diagnosis as the gold standard, the LANSS correctly classified 82% of the pains. In two studies, specificity ranged from 80% to 87% and sensitivity ranged from 83% to 85%. The LANSS has been translated into Mexican Spanish (83) and Turkish (84). Good psychometric properties were reported for the

Turkish version (eg, 94% of pains correctly classified, sensitivity = 90%, specificity = 94%).

The self-completed version of the LANSS (S-LANSS) (85) has the subject rub and press on the painful area of the body and a non-painful area and compare the two with regard to the type of discomfort (eg, burning, numbness) in place of the clinician examination. The internal consistency (Cronbach's alpha) of the S-LANSS ranged from 0.72 to 0.81, with greater consistency when completed as an interview. Using a cut-off of 12, the S-LANSS correctly identified 75% of the pains (sensitivity = 74%, specificity = 75%).

The psychometric properties of the LANSS are good and those of the S-LANSS are acceptable, especially if the S-LANSS is done as an interview rather than self-completed. The LANSS has not been used with persons with SCI. Due to impaired sensation at and/or below the level of injury, the examination items may not be valid for this population. Persons with SCI may not be able to perform the required rubbing and pressing areas of the body for two of the items of the S-LANSS. However, another person such as a family member may be able to perform those functions.

Neuropathic Pain Questionnaire and Neuropathic Pain Questionnaire—Short-Form. The Neuropathic Pain Questionnaire (NPQ) (86) is a 12-item measure designed to (a) discriminate between neuropathic and non-neuropathic pain and (b) assess neuropathic pain symptoms and change in those symptoms. The NPQ has 8 sensory descriptors of pain: burning, overly sensitive to touch, shooting, numbness, electric, tingling, squeezing, and freezing; 2 affective descriptors: unpleasant, and overwhelming; and 2 items on the causes of increased pain: touch, and weather changes. Cronbach's alpha was 0.95 for the original 32-items from which the NPQ was derived by factor analysis. In two groups, 71% and 76% of pains were correctly classified (sensitivity = 67%–75%, specificity = 74%–77%). Two persons with SCI were included in one sample, but no specific information on them was reported.

The Neuropathic Pain Questionnaire—Short-Form (NPQ-SF) (87) is a 3-item version that was derived by conducting a stepwise discriminant analysis to classify pains into neuropathic or non-neuropathic categories. Only numbness, tingling pain, and increased pain due to touch were significant predictors (discriminators). Total accuracy was 73% (sensitivity = 65%, specificity = 79%).

Although the developers indicate that the NPQ was designed to assess change, there are no published reports regarding its sensitivity to change. Among the 3 sets of instruments designed to differentiate between neuropathic and non-neuropathic pain, the NPQ and NPQ-SF correctly classified the smallest percentage of pains (71%–76%) and had the lowest sensitivity (65%–75%).

Neuropathic Pain Scale, Neuropathic Pain Scale—Revised, and Pain Quality Assessment Scale. The Neuropathic Pain Scale (NPS) (88), the Neuropathic Pain Scale-Revised (NPS-R) (89), and the Pain Quality Assessment Scale (PQAS) (89) were designed to assess pain qualities associated with neuropathic pain. The PQAS additionally assesses 3 qualities commonly associated with non-neuropathic pain. All 3 measures are designed to assess treatment effects. The NPS consists of 10 pain quality items (eg, sharp, hot, and dull) and a temporal pattern item (intermittent, variable, or stable). The NPS-R includes all of the NPS items plus 7 additional pain quality items (eg, electrical, tingling, and radiating). The PQAS includes all the NPS-R items plus another 3 pain quality items (shooting, cramping, and heavy) common in non-neuropathic pain; the instructions for completing the temporal pattern item were revised for better clarity. Each item is rated on a 0-to-10 scale, except the temporal pattern question, which asks the respondent to select one of three patterns. Each item was designed to be used as an independent rating without summation into a total score. However, in a 2002 article by the developers (90), NPS composite scores were calculated using all or selected items excluding the temporal pattern item. Another study by Jensen et al. combined the 6 pain descriptor items (91). Cronbach's alpha for the 10 pain quality items of the NPS ranged from 0.86 to 0.92 in two non-SCI populations.

Tai et al used the NPS in a small randomized controlled trial of gabapentin in persons with SCI (92). The NPS indicated improvement in unpleasant feeling, pain intensity and burning sensation in the group receiving gabapentin but not in the placebo group (92). A number of other studies have found the NPS to be sensitive to change in a variety of populations (47, 90, 91, 93, 94).

Of the measures designed to assess status and change in neuropathic pain, the NPS has been used in the largest number of studies and it is the only one that has been tested specifically in an SCI population. The NPS has been translated into 42 languages other than the original English; however, studies which used these other language versions have not been published to date except for an Italian version (95). The NPS-R and the PQRS are yet to be evaluated for the SCI population and their psychometric properties are largely unknown.

Neuropathic Pain Symptom Inventory. The Neuropathic Pain Symptom Inventory (NPSI) (68) assesses pain qualities associated with neuropathic pain and is designed to assess treatment effects. It consists of 10 descriptors of different symptoms and 2 items regarding the duration of spontaneous ongoing pain and paroxysmal pain. A total score is calculated by summing the scores on the 10 descriptor items. From the 10 items, 5 subscale scores can be calculated: Evoked Pain (pain evoked from brushing, pressure or contact with cold), Pressive or Deep Pain (pressure and squeezing), Paroxysmal Pain (electric shock and stabbing), Abnormal Sensations (tingling and pins and

needles), and Spontaneous Ongoing Pain (burning). Total scores can range from 0 to 100 for the 10 pain descriptor questions. Fourteen persons in the initial study (8% of the sample) had pain resulting from spinal cord trauma; however, no specific information was provided about this sub-sample.

Subjects completed a single-item global impression of change measure at the second visit. The Spearman correlation of the change in the NPSI between the first and second visit and the patient's global impression of change was 0.67. The clinician also rated global change on a similar scale. The correlation of change in the NPSI and the clinician's global impression of change was 0.58. The correlation between total score of the NPSI and a global rating of pain intensity was 0.60. The scores on the three items related to evoked pain were compared to the magnitude of mechanical or thermal evoked pain estimated by the investigator. The correlations were 0.70 for brushing, 0.66 for pressure, and 0.73 for cold. The total and subscale scores of the NPSI and changes in those scores were not significantly related to measures of anxiety and depression.

The original French version of the NPSI appears to have been well developed and to have good psychometric properties. However, the English version is untested and it is unknown how appropriate either the French or English versions are for use with persons with SCI.

Recommendations. Two types of neuropathic pain measures were reviewed: (a) those designed to discriminate neuropathic from non-neuropathic pain (DN-4, LANSS, and NPQ) and (b) those designed to assess status of and change in neuropathic pain (NPQ, NPS, and NPSI). All but the NPSI have versions with fewer or more items. With the exception of one study evaluating the NPS, the psychometric properties and usefulness of these scales have not been explicitly tested in SCI. Among the measures designed to discriminate between neuropathic and non-neuropathic pain, currently the full LANSS is the best choice for use in English because it has better psychometric properties than the NPQ and the properties of the English version of the DN-4 are unknown. If only self-reported data can be obtained, the S-LANSS administered as an interview is the best choice. To assess status and change, the best current choice with English-speaking subjects is some version of the NPS (NPS, NPS-R, PQRS) because it has better psychometric properties than the NPQ and the English version of the NPSI has not been tested. Furthermore, the NPS has been use in one study of persons with SCI and has been used by a number of different research teams with a variety of populations. Clearly, there is a need for more studies evaluating the usefulness of all of these measures in the SCI population. It is highly likely that measures specifically designed for the SCI population are needed that can overcome the problem of lack of sensation if the pain is below the level of injury.

Pain Interference Measures

The Graded Chronic Pain Scale. The Graded Chronic Pain Scale (GCPS) (96) also known as the von Korff scale, consists of 7 items, 3 referring to pain intensity and 4 referring to pain interference. On the basis of subscores, 4 classes of subjects are distinguished, differentiated on the basis of pain intensity and interference with activities: I. Low interference, low pain intensity; II. Low interference, high intensity; III. Moderate interference; and IV. High interference. The GCPS is a self- or interviewer-administered questionnaire that takes a few minutes to complete. Extensive psychometric information is available for non-SCI pain populations, and has been summarized by Von Korff (97). In the SCI population, the GCPS and its two subscales have been used by two different groups (55, 98–101). Internal consistency was 0.93 (Cronbach alpha) for the 3-item interference subscale (101). The interference-subscale score correlated with average pain intensity (0.50) and with the 5-item SF-36 Mental Health subscale (101). The interference-subscale also correlated with age, the pain intensity subscale score, and the Catastrophizing subscale of the Coping Strategies Questionnaire (100).

Sensitivity to change may be limited due to the phrasing of the interference items for people with long-standing pain (101) and a lack of differentiation between interference due to SCI per se and interference due to chronic pain. For this reason, the GCPS's applicability to SCI is limited.

SF-36 interference item. The Short Form-36 health status measure has a 2-item bodily pain subscale, one of which has been used in a number of studies as a measure of pain interference. "During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)"? is answered selecting the appropriate category on a 5-point Likert scale ranging from ''not at all'' to "extremely." This item has been used in SCI pain research by two different groups (102–106); all of these investigations studied subsamples of cases in the National SCI Database, and likely the same cases were included in two or more analyses.

The pain interference item can be self- or interviewer administered, and takes one minute to complete. The SF-36, the parent instrument, has been translated into most western languages and many non-western ones. Limited psychometric information is available for the SCI population. In one study, the one-year test-retest reliability had a value of $C = 0.47$ (association coefficient) (104). More pain interference was reported by those with gunshot wound etiology (104). Pain interference is correlated with satisfaction with life (0.32), CHART mobility subscore (0.11), and both the physical health component (0.20) and the mental health component (0.36) of the SF-12 – all measured a year prior to pain interference (105). No studies have reported on the item's sensitivity to change.

Multidimensional Pain Inventory interference items. The Multidimensional Pain Inventory (MPI) (49) has a subscale reflecting interference of pain with life that has been used by various SCI investigators (53, 54, 107–111). The MPI Life Interference (LI) subscale taps into interference with major activities (employment, household chores, recreational and social, family-related); relationships (friendships, spouse and family); and the satisfaction or enjoyment of the activities using 11 questions, which can be self- or interviewer administered in 5 to 10 minutes. The MPI has been translated into Dutch, Spanish, Swedish, German, Italian, and French. As used by Widerström-Noga in the SCI-MPI, the LI subscale has 8 items: two items related to work and enjoyment of work, and one item related to interference with day-to-day activities were deleted to improve the factor structure. The various Widerström-Noga studies, which are the main source for the evidence presented, all used subsets of Miami Project subjects. Samples in the studies varied in their diversity.

Internal consistency reliability has been reported (Cronbach alpha) at 0.91 (53) and 0.90 (54), and test-retest reliability ("one to several weeks" interval) at 0.81 (54). The LI scores correlate with scores on the Pain Disability Index (PDI), a measure of pain interference with functioning: 0.61 (54). Individuals who use prescription medications for pain report higher interference than those who do not (effect size: 0.60) (108). The MPI LI scores correlated at about 0.30 with a number of factors reported to aggravate one's pain (109). The LI scores predict satisfaction with life (measured with the Satisfaction with Life Scale (SWLS)), even after controlling for perceived control over one's life, affective distress, general activity level, and solicitous responses from spouse/partner (54). Life Interference also is correlated with anger, vigor, depression, and trait anxiety, even after controlling for interference with activities and life enjoyment due to SCI per se (107).

Because its items explicitly refer to the impact of pain, the LI subscale of the MPI is suitable for the SCI population.

Brief Pain Inventory interference items. The Brief Pain Inventory (BPI) was developed by Cleeland, and has been used in a number of investigations of chronic pain (112). Its pain interference subscale has been used in several investigations of pain in SCI (101, 113–115). In its original version the subscale consisted of 7 items measuring interference with general activity, sleep, mood, relationships, etc.; Jensen has added 3 items of relevance to people with disabilities (self-care, recreation and social activities), and later added communication and learning new information (101). These versions are referred to as BPI-I7, BPI-I10, and BPI-I12, respectively.

The BPI pain interference subscale can be self- or interviewer administered in about 5 minutes; translations (of the BPI-I7) are available in Spanish, French, German, Italian, Chinese, and other languages. For uses with a SCI sample, in item 3 of BPI-I7, "walking ability" should be replaced by "ability to get around."

Cronbach alpha was 0.92 for BPI-I7, 0.95 for BPI-I10, and 0.96 for BPI-I12 (101), indicating good reliability. The 3 versions of the BPI-I correlated at 0.62 or higher with pain intensity, and 0.60 or higher with the SF-36 mental health scale (101). All individual items correlated with pain intensity (101, 114). BPI-I7 scores were higher, as expected, in adult-onset wheelchair users than in those who started using wheelchairs before their 16th birthday (Effect size: 0.68) (115). BPI-I7 scores correlated 0.35 with scores on the Wheelchair User's Shoulder Pain Index (115). As hypothesized, BPI-I10 scores correlated with Solicitous and with Negative Partner responses on the Multidimensional Pain Inventory (113).

Recommendations. The SF-36 pain interference item as a single-item pain interference measure and the MPI or BPI interference items as a more comprehensive measure of pain interference are recommended over the GCPS by the committee. They also are the first choice by the participants of the ASIA/ISCOS Measurement precourse as measures that could be included in a minimum dataset for clinical trials of pain after SCI. Further research is needed on whether the MPI is to be preferred over the BPI. The 10- and 12-item versions of the BPI seem unnecessary expansions beyond the 7-item version, which has adequate internal consistency. Norms for SCI samples have not been published for either of these instruments.

Pain Coping Measure

Coping Strategies Questionnaire (CSQ). While there is substantial literature documenting the significant prevalence of pain with SCI, little research has examined the specific cognitive and behavioral responses to the pain experience itself such as catastrophizing in adjustment to chronic pain (116). Some research has explored partial beliefs in coping and adjustment to SCI (117). These studies have been single center focused and have not been replicated. Turner et al (100) have commented on the absence of research that explores coping processes in chronic pain post SCI.

The CSQ is an instrument that is widely used in studies of coping with heterogeneous pain. It was designed to include both cognitive (diverting attention, reinterpreting pain sensations, coping self-statements, ignoring pain sensations, and prayer or hoping) and behavioral (increasing activity level) reactions to pain. Rosenstiel developed the CSQ with 50 items, defining 8 subscales with 6 items (plus 2 non-scale items rating the effectiveness of coping strategies) (118). Each item is rated on 0–6 scale: "Never do" to "always do" when in pain. There is a 24-item version developed by Harland & Georgieff and a 14- and 7-item version (119, 120). Dutch and German versions are available. Harland & Georgieff carried out a varimax-rotated component analysis and

Table 6. Recommendations for Outcome Measures

Construct	Recommended measures
Pain intensity	0–10 Point NRS*
Global improvement of pain	7 Point PGIC (Original Guy/Farrar)†
Pain interference	SF-36 single question and MPI or BPI pain interference items†
Neuropathic and nociceptive pain discrimination	LANSS‡
Change in neuropathic pain	NPS‡
Classification of pain after SCI	Proposed IASP Taxonomy or BR-SCI-PT‡
Mechanical allodynia/hyperalgesia	Brush or cotton wool and at least one high-threshold von Frey filament‡
Thermal allodynia/hyperalgesia	Peltier-type thermotester‡

*Established validity.
†Adequate validity for measuring pain after SCI.
‡Unknown validity for measuring pain after SCI.

identified a 4-factor solution which was similar to solutions obtained in previous research (119).

In one non-SCI sample, internal consistency Cronbach alpha coefficients ranged from 0.71–0.85 for all the sub-scales, except increasing pain behaviors, which had an alpha coefficient of 0.28 (118). Adequate test/retest reliability has been demonstrated (121). Evidence for sensitivity to change has been demonstrated in a heterogeneous group of people with chronic pain and a sample of people with SCI (100).

Concurrent validity has been extensively studied and supported. The scales have explained significant variance in adjustment and wellbeing (pain intensity and psychological distress). The CSQ has explained an average of 17% of the variance in pain and intensity (122). The catastrophizing sub-scale was found to be consistently and strongly associated with pain intensity, distress and pain related disability (100, 123).

Numerous pain studies have documented a robust relationship between catastrophizing in general and adverse pain related outcomes (124). The CSQ has recently been used in SCI and results would suggest that the catastrophizing subscale is of particular relevance to SCI clinicians. The full CSQ was used by Turner while Roth used only the Catastrophizing subscale (100, 123). The shorter versions of the CSQ would seem to be the most useful in SCI.

Recommendations. As 73% of the participants of the ASIA/ISCOS Measurement precourse noted that this measure needs further study of its reliability and validity before being part of any recommended dataset for SCI, it is not recommended for inclusion in a minimal dataset at this time.

CONCLUSIONS

From this systematic review of outcome measures relevant to chronic pain, it can be concluded that most of the measures need further development and determination of reliability and validity in SCI populations. That being said, it is the consensus of the committee that for each construct the available measures with the greatest validity, incomplete as their assessment may be, should be used in clinical trials while this development proceeds. Some measures should be used without hesitation as they have been shown to be valid and reliable in the study of pain after SCI or in other populations for which there would not be expected to be differences in the way pain is experienced from persons with SCI. Others, which may have been shown to be valid when used in a non-SCI population, should be used with more caution as their applicability without modification for use in evaluating persons with SCI may be questionable especially in regards to the presence of the underlying motor and sensory impairments intrinsic to SCI confounding the outcomes measured. Still others still should be used only in selected circumstances depending upon what is to be measured in a particular clinical trial.

Specific recommendations for constructs and for measures of these constructs are listed in Table 6. These recommendations are based upon a judgment of the committee of the validity and reliability of the measures and the applicability of these measures to pain after SCI. The recommendations have been further qualified by noting if the measures have established, adequate, or unknown validity for measuring pain after SCI. It should be emphasized that this list includes only those constructs for which there was adequate literature available to review and make judgments of the validity and reliability of the measures for these constructs. In addition, the list of constructs is limited to pain proper and does not include those subjective experiences which are not conceptually related to pain such as quality of life, emotional functioning, and physical functioning. The authors recognize these as important for describing the overall impact of pain on life after SCI, and often as useful in a clinical trial of treatments of pain

after SCI. However, they are beyond the scope of this article; some of the other papers describing results of the ASIA/ISCoS outcomes assessment project address these instruments.

SUPPORT

This work was supported by funds from the National Institute on Disability and Rehabilitation Research, Office of Special Education and Rehabilitation Services in the US Department of Education to Mount Sinai School of Medicine, The University of Alabama, Birmingham, and the University of Washington, as well as the New South Wales Government Spinal Cord Injury & Other Neurological Conditions Research Grants Program.

REFERENCES

1. Turk DC, Dworkin RH, Allen RR, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. *Pain.* 2003;106:337–345.
2. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain.* 2005;113:9–19.
3. Cruccu G, Anand P, Attal N, et al. EFNS guidelines on neuropathic pain assessment. *Eur J Neurol.* 2004;11:153–162.
4. Roberts P, Priest H. Reliability and validity in research. *Nurs Stand.* 2006;20:41–45.
5. Baxter H. Understanding research: 2. Ensuring reliability and validity. *J Wound Care.* 2001;10:329–331.
6. Davis DW. Validity and reliability: Part I. *Neonatal Netw.* 2004;23:54–56.
7. Davis DW. Reliability and validity: Part II. *Neonatal Netw.* 2004;23:74–76.
8. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. *Am J Med.* 2006;119:166.e7–e16.
9. Fava GA, Belaise C. A discussion on the role of clinimetrics and the misleading effects of psychometric theory. *J Clin Epidemiol.* 2005;58:753–756.
10. Borsboom D, Mellenbergh GJ, van Heerden J. The concept of validity. *Psychol Rev.* 2004;111:1061–1071.
11. Dekker J, Dallmeijer AJ, Lankhorst GJ. Clinimetrics in rehabilitation medicine: current issues in developing and applying measurement instruments 1. *J Rehabil Med.* 2005;37:193–201.
12. Streiner DL, Norman GR. "Precision" and "accuracy": two terms that are neither. *J Clin Epidemiol.* 2006;59:327–330.
13. Smith GT. On construct validity: issues of method and measurement. *Psychol Assess.* 2005;17:396–408.
14. Nunnally J, Bernstein I. *Psychometric Theory,* 3rd ed. New York, NY: McGraw-Hill; 1994.
15. Streiner D, Norman G. *Health Measurement Scales. A Practical Guide to their Development and Use,* 3rd ed. New York, NY: Oxford University Press; 2003.
16. Price D, Harkins S. Combined use of experimental pain and visual analogue scales in providing standardized measurement of clinical pain. *Clin J Pain.* 1987;3:1–8.
17. Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain.* 1983;17:45–56.
18. Breivik EK, Bjornsson GA, Skovlund E. A comparison of pain rating scales by sampling from clinical trial data. *Clin J Pain.* 2000;16:22–28.
19. Downie WW, Leatham PA, Rhind VM, Wright V, Branco JA, Anderson JA. Studies with pain rating scales. *Ann Rheum Dis.* 1978;37:378–381.
20. Kremer E, Atkinson JH, Ignelzi RJ. Measurement of pain: patient preference does not confound pain measurement. *Pain.* 1981;10:241–248.
21. Paice JA, Cohen FL. Validity of a verbally administered numeric rating scale to measure cancer pain intensity. *Cancer Nurs.* 1997;20:88–93.
22. Joyce CR, Zutshi DW, Hrubes V, Mason RM. Comparison of fixed interval and visual analogue scales for rating chronic pain. *Eur J Clin Pharmacol.* 1975;8:415–420.
23. Jensen MP, Chen C, Brugger AM. Postsurgical pain outcome assessment. *Pain.* 2002;99:101–109.
24. Herr KA, Spratt K, Mobily PR, Richardson G. Pain intensity assessment in older adults: use of experimental pain to compare psychometric properties and usability of selected pain scales with younger adults. *Clin J Pain.* 2004;20:207–219.
25. Gallagher EJ, Bijur PE, Latimer C, Silver W. Reliability and validity of a visual analog scale for acute abdominal pain in the ED. *Am J Emerg Med.* 2002;20:287–290.
26. Lund I, Lundeberg T, Sandberg L, Budh CN, Kowalski J, Svensson E. Lack of interchangeability between visual analogue and verbal rating pain scales: a cross sectional description of pain etiology groups. *BMC Med Res Methodol.* 2005;5:31.
27. Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. *Pain.* 1986;27:117–126.
28. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain.* 2001;94:149–158.
29. Buchbinder R, Bombardier C, Yeung M, Tugwell P. Which outcome measures should be used in rheumatoid arthritis clinical trials? Clinical and quality-of-life measures' responsiveness to treatment in a randomized controlled trial. *Arthritis Rheum.* 1995;38:1568–1580.
30. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain.* 2004;8:283–291.
31. Hurst H, Bolton J. Assessing the clinical significance of change scores recorded on subjective outcome measures. *J Manipulative Physiol Ther.* 2004;27:26–35.
32. Guy W. *ECDEU Assessement Manual for Psychopharmacology,* Revised 1976. US Government Printing Office; 1976. Rockville, Md.
33. Forouzanfar T, Weber WE, Kemler M, van Kleef M. What is a meaningful pain reduction in patients with complex regional pain syndrome type 1? *Clin J Pain.* 2003;19:281–285.
34. Fosnocht DE, Chapman CR, Swanson ER, Donaldson GW. Correlation of change in visual analog scale with pain relief in the ED. *Am J Emerg Med.* 2005;23:55–59.
35. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain.* 1975;1:277–299.

36. Melzack R. The Short-Form McGill Pain Questionnaire. *Pain.* 1987;30:191–197.
37. Melzack R, Katz J. Pain assessment in adult patients. In: *Wall and Melzack's Textbook of Pain.* McMahon S, Koltzenburg M, (eds). Edinburgh: Churchill Livingstone; 2005.
38. Dudgeon D, Raubertas RF, Rosenthal SN. The Short-Form McGill Pain Questionnaire in chronic cancer pain. *J Pain Symptom Manage.* 1993;8:191–195.
39. Dudgeon BJ, Ehde DM, Cardenas DD, Engel JM, Hoffman AJ, Jensen MP. Describing pain with physical disability: narrative interviews and the McGill Pain Questionnaire. *Arch Phys Med Rehabil.* 2005;86:109–115.
40. Cardenas DD, Turner JA, Warms CA, Marshall HM. Classification of chronic pain associated with spinal cord injuries. *Arch Phys Med Rehabil.* 2002;83:1708–1714.
41. Putzke JD, Richards JS, Hicken BL, Ness TJ, Kezar L, DeVivo M. Pain classification following spinal cord injury: the utility of verbal descriptors. *Spinal Cord.* 2002;40:118–127.
42. Beattie PF, Dowda M, Feuerstein M. Differentiating sensory and affective-sensory pain descriptions in patients undergoing magnetic resonance imaging for persistent low back pain. *Pain.* 2004;110:189–196.
43. Georgoudis G, Watson PJ, Oldham JA. The development and validation of a Greek version of the Short-Form McGill Pain Questionnaire. *Eur J Pain.* 2000;4:275–281.
44. Grafton KV, Foster NE, Wright CC. Test-retest reliability of the Short-Form McGill Pain Questionnaire: assessment of intraclass correlation coefficients and limits of agreement in patients with osteoarthritis. *Clin J Pain.* 2005;21:73–82.
45. Georgoudis G, Oldham JA, Watson PJ. Reliability and sensitivity measures of the Greek version of the Short Form of the McGill Pain Questionnaire. *Eur J Pain.* 2001;5: 109–118.
46. Dworkin RH, Corbin AE, Young JP Jr, et al. Pregabalin for the treatment of postherpetic neuralgia: a randomized, placebo-controlled trial. *Neurology.* 2003;60:1274–1283.
47. Lynch ME, Clark AJ, Sawynok J. Intravenous adenosine alleviates neuropathic pain: a double blind placebo controlled crossover trial using an enriched enrollment design. *Pain.* 2003;103:111–117.
48. Finnerup NB, Johannesen IL, Sindrup SH, Bach FW, Jensen TS. Pain and dysesthesia in patients with spinal cord injury: a postal survey. *Spinal Cord.* 2001;39:256–262.
49. Kerns RD, Turk DC, Rudy TE. The West Haven-Yale Multidimensional Pain Inventory (WHYMPI). *Pain.* 1985; 23:345–356.
50. Bergstrom KG, Jensen IB, Linton SJ, Nygren AL. A psychometric evaluation of the Swedish version of the Multidimensional Pain Inventory (MPI-S): a gender differentiated evaluation. *Eur J Pain.* 1999;3:261–273.
51. Lousberg R, Van Breukelen GJ, Groenman NH, Schmidt AJ, Arntz A, Winter FA. Psychometric properties of the Multidimensional Pain Inventory, Dutch language version (MPI-DLV). *Behav Res Ther.* 1999;37:167–182.
52. Riley JL 3rd, Zawacki TM, Robinson ME, Geisser ME. Empirical test of the factor structure of the West Haven-Yale Multidimensional Pain Inventory. *Clin J Pain.* 1999; 15:24–30.
53. Widerstrom-Noga EG, Duncan R, Felipe-Cuervo E, Turk DC. Assessment of the impact of pain and impairments associated with spinal cord injuries. *Arch Phys Med Rehabil.* 2002;83:395–404.
54. Widerstrom-Noga EG, Cruz-Almeida Y, Martinez-Arizala A, Turk DC. Internal consistency, stability, and validity of the spinal cord injury version of the Multidimensional Pain Inventory. *Arch Phys Med Rehabil.* 2006;87:516–523.
55. Barrett H, McClelland JM, Rutkowski SB, Siddall PJ. Pain characteristics in patients admitted to hospital with complications after spinal cord injury. *Arch Phys Med Rehabil.* 2003;84:789–795.
56. Siddall PJ, McClelland JM, Rutkowski SB, Cousins MJ. A longitudinal study of the prevalence and characteristics of pain in the first 5 years following spinal cord injury. *Pain.* 2003;103:249–257.
57. von Frey M. Ueber den Gebrauch von Reizhaaren. On the use of stimulation hairs. In: *Untersuchungen über die Sinnesfunctionen der menschlichen Haut. Erste Abhandlung: Druckempfindung und Schmerz. (Investigations on the sensory functioning of the human skin. Part one: Pressure sense and pain.). Abhandlungen der Mathematisch-Physischen Classe der Königlich Sachsischen Gesellschaft der Wissenschaften. (Archives of the Mathematical-Physical Chamber of the Royal Saxon Scientific Society)*. 1896;23:208–217.
58. Bell-Krotoski J, Tomancik E. The repeatability of testing with Semmes-Weinstein monofilaments. *J Hand Surg [Am].* 1987;12:155–161.
59. Bove GM, Robichaud DR, Grigg P. Three-dimensional load analysis of indentation stimulators. *J Neurosci Methods.* 2003;123:23–30.
60. Andrews K. The effect of changes in temperature and humidity on the accuracy of von Frey hairs. *J Neurosci Methods.* 1993;50:91–93.
61. Bove G. Mechanical sensory threshold testing using nylon monofilaments: The pain field's "tin standard". *Pain.* 2006;124:13–17.
62. Samuelsson M, Leffler AS, Hansson P. Dynamic mechanical allodynia: on the relationship between temporospatial stimulus parameters and evoked pain in patients with peripheral neuropathy. *Pain.* 2005;115:264–272.
63. Bouhassira D, Attal N, Alchaar H, et al. Comparison of pain syndromes associated with nervous or somatic lesions and development of a new neuropathic pain diagnostic questionnaire (DN4). *Pain.* 2005;114:29–36.
64. Attal N, Gaude V, Brasseur L, et al. Intravenous lidocaine in central pain: a double-blind, placebo-controlled, psychophysical study. *Neurology.* 2000;54:564–574.
65. Attal N, Guirimand F, Brasseur L, Gaude V, Chauvin M, Bouhassira D. Effects of IV morphine in central pain: a randomized placebo-controlled study. *Neurology.* 2002; 58:554–563.
66. Canavero S, Bonicalzi V. Intravenous subhypnotic propofol in central pain: a double-blind, placebo-controlled, crossover study. *Clin Neuropharmacol.* 2004;27:182–186.
67. Finnerup NB, Biering-Sorensen F, Johannesen IL, et al. Intravenous lidocaine relieves spinal cord injury pain: a randomized controlled trial. *Anesthesiology.* 2005;102: 1023–1030.
68. Bouhassira D, Attal N, Fermanian J, et al. Development and validation of the Neuropathic Pain Symptom Inventory. *Pain.* 2004;108:248–257.

69. Rolke R, Magerl W, Campbell KA, et al. Quantitative sensory testing: a comprehensive protocol for clinical trials. *Eur J Pain.* 2006;10:77–88.
70. Cathcart S, Pritchard D. Reliability of pain threshold measurement in young adults. *J Headache Pain.* 2006;7: 21–26.
71. Defrin R, Ohry A, Blumen N, Urca G. Characterization of chronic pain and somatosensory function in spinal cord injury subjects. *Pain.* 2001;89:253–263.
72. Jorum E, Warncke T, Stubhaug A. Cold allodynia and hyperalgesia in neuropathic pain: The effect of N-methyl-D-aspartate (NMDA) receptor antagonist ketamine–a double-blind, cross-over comparison with alfentanil and placebo. *Pain.* 2003;101:229–235.
73. Siddall PJ, Taylor DA, Cousins MJ. Classification of pain following spinal cord injury. *Spinal Cord.* 1997;35:69–75.
74. Siddall PJ, Yezierski RP, Loeser JD. Pain following spinal cord injury: clinical features, prevalence, and taxonomy. *IASP Newsletter.* 2000;3:3–7.
75. Siddall PJ, Taylor DA, McClelland JM, Rutkowski SB, Cousins MJ. Pain report and the relationship of pain to physical factors in the first 6 months following spinal cord injury. *Pain.* 1999;81:187–197.
76. Siddall PJ, Middleton JW. A proposed algorithm for the management of pain following spinal cord injury. *Spinal Cord.* 2006;44:67–77.
77. Putzke JD, Richards JS, Ness T, Kezar L. Interrater reliability of the International Association for the Study of Pain and Tunks' spinal cord injury pain classification schemes. *Am J Phys Med Rehabil.* 2003;82:437–440.
78. Ragnarsson KT. Management of pain in persons with spinal cord injury. *J Spinal Cord Med.* 1997;20:186–199.
79. Bryce TN, Ragnarsson KT. Pain management in persons with spinal cord disorders. In: *Spinal Cord Medicine: Principles and Practice.* Lin VW, Cardenas DD, Cutter NC, (Eds). New York: Demos Publishing; 2003:441–460.
80. Bryce TN, Dijkers MP, Ragnarsson KT, Stein AB, Chen B. Reliability of the Bryce/Ragnarsson spinal cord injury pain taxonomy. *J Spinal Cord Med.* 2006;29:118–132.
81. Cardenas D. Current concepts of rehabilitation of spinal cord injury patients. *Spine State Art Rev.* 1999;13:575–585.
82. Bennett M. The LANSS pain scale: the Leeds Assessment of Neuropathic Symptoms and Signs. *Pain.* 2001;92:147–157.
83. Martinez-Lavin M, Lopez S, Medina M, Nava A. Use of the Leeds Assessment of Neuropathic Symptoms and Signs questionnaire in patients with fibromyalgia. *Semin Arthritis Rheum.* 2003;32:407–411.
84. Yucel A, Senocak M, Kocasoy Orhan E, Cimen A, Ertas M. Results of the Leeds Assessment of Neuropathic Symptoms and Signs pain scale in Turkey: a validation study. *J Pain.* 2004;5:427–432.
85. Bennett MI, Smith BH, Torrance N, Potter J. The S-LANSS score for identifying pain of predominantly neuropathic origin: validation for use in clinical and postal research. *J Pain.* 2005;6:149–158.
86. Krause SJ, Backonja MM. Development of a neuropathic pain questionnaire. *Clin J Pain.* 2003;19:306–314.
87. Backonja MM, Krause SJ. Neuropathic Pain Questionnaire –Short Form. *Clin J Pain.* 2003;19:315–316.
88. Galer BS, Jensen MP. Development and preliminary validation of a pain measure specific to neuropathic pain: the Neuropathic Pain Scale. *Neurology.* 1997;48:332–338.
89. Jensen MP, Gammaitoni AR, Olaleye DO, Oleka N, Nalamachu SR, Galer BS. The Pain Quality Assessment Scale (PQAS): Assessment of pain quality in carpal tunnel syndrome. *J Pain.* 2006;7:823–832.
90. Galer BS, Jensen MP, Ma T, Davies PS, Rowbotham MC. The lidocaine patch 5% effectively treats all neuropathic pain qualities: results of a randomized, double-blind, vehicle-controlled, 3-week efficacy study with use of the Neuropathic Pain Scale. *Clin J Pain.* 2002;18:297–301.
91. Jensen MP, Dworkin RH, Gammaitoni AR, Olaleye DO, Oleka N, Galer BS. Assessment of pain quality in chronic neuropathic and nociceptive pain clinical trials with the Neuropathic Pain Scale. *J Pain.* 2005;6:98–106.
92. Tai Q, Kirshblum S, Chen B, Millis S, Johnston M, DeLisa JA. Gabapentin in the treatment of neuropathic pain after spinal cord injury: a prospective, randomized, double-blind, crossover trial. *J Spinal Cord Med.* 2002;25:100–105.
93. Moseley GL. Why do people with complex regional pain syndrome take longer to recognize their affected hand? *Neurology.* 2004;62:2182–2186.
94. Rog DJ, Nurmikko TJ, Friede T, Young CA. Randomized, controlled trial of cannabis-based medicine in central pain in multiple sclerosis. *Neurology.* 2005;65:812–819.
95. Negri E, Bettaglio R, Demartini L, et al. Validation of the Italian version of the "Neuropathic Pain Scale" and its clinical applications. *Minerva Anestesiol.* 2002;68:95–104.
96. Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. *Pain.* 1992;50:133–149.
97. Von Korff M, Jensen MP, Karoly P. Assessing global pain severity by self-report in clinical and health services research. *Spine.* 2000;25:3140–3151.
98. Turner JA, Cardenas DD. Chronic pain problems in individuals with spinal cord injuries. *Semin Clin Neuropsychiatry.* 1999;4:186–194.
99. Hanley MA, Masedo A, Jensen MP, Cardenas D, Turner JA. Pain interference in persons with spinal cord injury: classification of mild, moderate, and severe pain. *J Pain.* 2006;7:129–133.
100. Turner JA, Jensen MP, Warms CA, Cardenas DD. Catastrophizing is associated with pain intensity, psychological distress, and pain-related disability among individuals with chronic pain after spinal cord injury. *Pain.* 2002;98:127–134.
101. Raichle KA, Osborne TL, Jensen MP, Cardenas D. The reliability and validity of pain interference measures in persons with spinal cord injury. *J Pain.* 2006;7:179–186.
102. Cardenas DD, Bryce TN, Shem K, Richards JS, Elhefni H. Gender and minority differences in the pain experience of people with spinal cord injury. *Arch Phys Med Rehabil.* 2004;85:1774–1781.
103. Putzke JD, Richards JS, DeVivo MJ. Quality of life after spinal cord injury caused by gunshot. *Arch Phys Med Rehabil.* 2001;82:949–954.
104. Putzke JD, Richards JS, DeVivo MJ. Predictors of pain 1 year post-spinal cord injury. *J Spinal Cord Med.* 2001;24: 47–53.
105. Putzke JD, Richards JS, Hicken BL, DeVivo MJ. Interference due to pain following spinal cord injury: important

predictors and impact on quality of life. *Pain.* 2002;100: 231–242.
106. Putzke JD, Barrett JJ, Richards JS, DeVivo MJ. Age and spinal cord injury: an emphasis on outcomes among the elderly. *J Spinal Cord Med.* 2003;26:37–44.
107. Summers JD, Rapoff MA, Varghese G, Porter K, Palmer RE. Psychosocial factors in chronic spinal cord injury pain. *Pain.* 1991;47:183–189.
108. Widerstrom-Noga EG, Turk DC. Types and effectiveness of treatments used by people with chronic pain associated with spinal cord injuries: influence of pain and psychosocial characteristics. *Spinal Cord.* 2003;41:600–609.
109. Widerstrom-Noga EG, Turk DC. Exacerbation of chronic pain following spinal cord injury. *J Neurotrauma.* 2004;21: 1384–1395.
110. Widerstrom-Noga EG, Duncan R, Turk DC. Psychosocial profiles of people with pain associated with spinal cord injury: identification and comparison with other chronic pain syndromes. *Clin J Pain.* 2004;20:261–271.
111. Khenioui H, Cahagne V, Brissot R. Évaluation fonctionnelle de la douleur chronique chez le patient blessé médullaire (Assessment of chronic pain as a disability in patients with spinal cord injuries). *Ann Readapt Med Phys*. 2006;49:125–137.
112. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. *Ann Acad Med Singapore.* 1994;23: 129–138.
113. Stroud MW, Turner JA, Jensen MP, Cardenas DD. Partner responses to pain behaviors are associated with depression and activity interference among persons with chronic pain and spinal cord injury. *J Pain.* 2006;7:91–99.
114. Jensen MP, Hoffman AJ, Cardenas DD. Chronic pain in individuals with spinal cord injury: a survey and longitudinal study. *Spinal Cord.* 2005;43:704–712.
115. Sawatzky BJ, Slobogean GP, Reilly CW, Chambers CT, Hol AT. Prevalence of shoulder pain in adult- versus childhood-onset wheelchair users: a pilot study. *J Rehabil Res Dev.* 2005;42:1–8.
116. Sipski ML, Richards JS. Spinal cord injury rehabilitation: state of the science. *Am J Phys Med Rehabil.* 2006;85:310–342.
117. Kennedy P, Frankel H, Gardner B, Nuseibeh I. Factors associated with acute and chronic pain following traumatic spinal cord injuries. *Spinal Cord.* 1997;35:814–817.
118. Rosenstiel AK, Keefe FJ. The use of coping strategies in chronic low back pain patients: relationship to patient characteristics and current adjustment. *Pain.* 1983;17:33–44.
119. Harland NJ, Georgieff K. Development of the Coping Strategies Questionnaire 24, a clinically utilitarian version of the Coping Strategies Questionnaire. *Rehabil Psychol.* 2003;48:296–300.
120. Jensen MP, Keefe FJ, Lefebvre JC, Romano JM, Turner JA. One- and two-item measures of pain beliefs and coping strategies. *Pain.* 2003;104:453–469.
121. Main CJ, Waddell G. A comparison of cognitive measures in low back pain: statistical structure and clinical validity at initial assessment. *Pain.* 1991;46:287–298.
122. Gil KM, Abrams MR, Phillips G, Keefe FJ. Sickle cell disease pain: relation of coping strategies to adjustment. *J Consult Clin Psychol.* 1989;57:725–731.
123. Roth RS, Lowery JC, Hamill JB. Assessing persistent pain and its relation to affective distress, depressive symptoms, and pain catastrophizing in patients with chronic wounds: a pilot study. *Am J Phys Med Rehabil.* 2004;83: 827–834.
124. Sullivan MJ, Thorn B, Haythornthwaite JA, et al. Theoretical perspectives on the relation between catastrophizing and pain. *Clin J Pain.* 2001;17:52–64.

ERRATUM

The byline for the following abstract, which was published in the previous issue, should have included co-author Marinella Galea. The correct citation is:

Post J, Spungen A, Zindelhadid R, Emmons R, Bauman W, Korsten M, Galea M. Renal safety of oral phosphate solution in patients with SCI undergoing colonoscopy. Abstract #P2. *J Spinal Cord Med.* 2007:30(4):399.