Original Article

Test-retest reliability of the Donovan spinal cord injury pain classification scheme

JD Putzke¹, JS Richards^{*,1}, T Ness¹ and L Kezar¹

¹University of Alabama at Birmingham, Birmingham, USA

Study design: Videotape rating by independent viewers. **Objective:** To determine the test–retest reliability of the Donovan spinal cord injury (SCI) pain classification scheme.

Setting: Rehabilitation Centre, Alabama, USA.

Methods: A total of 28 individuals with SCI reported 60 pain sites. A structured interview and physical exam were used to illicit information to classify each pain site according to the Donovan criteria. All structured interviews and exams were videotaped. Three independent raters viewed the videotapes on two occasions, separated by a 3-month interval, and classified each pain site using the Donovan pain classification scheme.

Results: Considering all three raters together, 78% of the pain sites were consistently classified from one period to the next. Within each rater, consistent classification ranged from 67 to 83%. However, inter-rater agreement for the classification of each pain site into the various types of pain was low for both periods (about 50–60%).

Conclusions: Pain classification within each rater generally showed adequate test-retest reliability when using the Donovan SCI pain classification scheme. However, reliability estimates of agreement across raters highlight the ongoing need to exam and improve the psychometric characteristics of the various pain classification schemes. *Spinal Cord* (2003) **41**, 239–241. doi:10.1038/sj.sc.3101434

Keywords: pain; classification; reliability; test-retest reliability; spinal cord injury

Introduction

Numerous pain classification schemes have been proposed for individuals with spinal cord injury (SCI).^{1–5} Unfortunately, relatively little is known about the psychometric properties of these classification schemes. We previously examined the inter-rater reliability of a common SCI pain classification scheme (Donovan *et al*³) and found agreement across three independent raters to range from 50 to 70%.⁶ Moreover, per cent agreement did not change as additional classification criteria were provided to each rater. However, raters' confidence ratings in the accuracy in their pain classification systematically increased as each additional criterion was provided.

The primary aim of the current study was to extend our previous research by examining the test-retest reliability of the Donovan SCI pain classification scheme. More specifically, the same three independent raters classified the same pain sites as previously reported following a 3-month interval. Classification agreement across periods was assessed.

Methods

Participants

A total of 28 individuals with traumatic onset SCI were recruited for the study. All study participants were at least 1 year postinjury, 18 years or older, and reported chronic (ie, 6 months or more) pain in one or more sites. Participants were the same as those previously reported.⁶ Exclusion criteria included a fourth grade reading level or lower, a history of traumatic brain injury, a history of chronic pain prior to SCI onset, and other medical conditions or complications that may account for chronic pain. Participants were recruited from the SCI clinic at the University of Alabama at Birmingham and through local advertisements, and were paid \$25 for their participation. np

^{*}Correspondence: JS Richards, Spain Rehabilitation Center, 1717, 6th Avenue South, Room 529, Birmingham, Al 35233-7330, USA.

Procedures and Measures

A more detailed description of the procedures are provided in a previous report.⁶ In short, participants completed a questionnaire assessing demographic and injury-related characteristics after obtaining informed consent. A semistructured interview and physical exam were used to elicit information on 'all of the places you have pain.' Participants were allowed to report multiple pain sites but were also told 'you may have pain in several different places that to you is the same kind of pain. If this is so, we will ask you to group those pains together and answer questions about them as a group.' Consistent with the Donovan classification scheme (see description below), patients were interviewed about each pain site to obtain information about each site's location, character (ie, verbal descriptors of the pain), length of time postinjury when the pain first began, duration (average length of a pain episode), and aggravating and diminishing factors.

All semistructured interviews and physical exams were videotaped. Three experienced clinicians (TN, SR, LK) independently classified the 60 pain sites using the Donovan classification scheme. Each clinician used information from the semistructured interview and the videotaped physical exam to classify each pain site. The first classification formed the basis of our inter-rater reliability study.⁶ After about a 3-month delay, the information from the semistructured interview and videotape were re-reviewed, and each pain site was classified again according to the Donovan classification scheme. Thus, the same classification information and the same stimulus materials were used for both assessment periods.

The Donovan classification scheme³ includes five pain types (ie, segmental nerve/cauda equina, spinal psychogenic). cord. visceral, mechanical, and The various types of pain are determined based on distinctions within four areas that include (1) time of onset (eg, days to months after injury), (2) verbal descriptors of the pain (eg, burning, aching, tingling), (3) duration of the pain experience (eg, seconds, constant), and (4) aggravating and mitigating factors (eg, rest, activity). Donovan also provided several case examples of each pain type to help facilitate pain classification.

Results

Table 1 presents a summary of the participants demographic and medical characteristics. Participants tended to be middle-aged (M = 46 years), male (82%), Caucasians (82%), with paraplegia (75%), and a greater

 Table 1
 Demographic and medical characteristics of participants

	M(SD)
Age Years since injury Number of pain sites	45.5 (10.02) 10.3 (7.20) 2.1 (0.89)
WRAT reading grade level	10 (4.0)
Education 11th grade or less High school or G.E.D More than high school	n (%) 5 (18) 8 (28) 15 (54)
Gender Males Females	23 (82) 5 (18)
Race Caucasian African American	23 (82) 5 (18)
<i>Level of impairment</i> Paraplegia, incomplete Paraplegia, complete Tetraplegia, incomplete Tetraplegia, complete	10 (36) 11 (39) 5 (18) 2 (7)
Etiology of SCI Vehicular Sports Violence Other	16 (57) 2 (7) 4 (14) 6 (21)
ASIA impairment index Complete injury Incomplete injury B Incomplete injury C Incomplete injury D	15 (54) 7 (25) 4 (14) 2 (7)

SCI = Spinal cord injury.

 Table 2
 Test-retest reliability of the Donovan pain classification scheme

Initial classification	Second classification (3-month separation)			
	Segmental nerve/cauda equina n (%)	Spinal cord n (%)	Visceral n (%)	Mechanical n (%)
Segmental nerve/cauda equina	32 (67)	10 (21)	1 (2)	5 (10)
Spinal cord	7 (11)	52 (84)		3 (5)
Visceral		1 (25)	3 (75)	
Mechanical	6 (9)	7 (11)	× ,	53 (80)

Note: Per cent of initial classification.

than high school education (54%). Etiology of SCI was predominantly motor vehicle accident (57%).

Table 2 presents the results of the test-retest reliability analysis. As can be seen, pain sites were classified into four of the five Donovan pain types at both assessment periods. Consistent classification between periods ranged from about 67 to 84% across the four pain types, with 'segmental nerve/cauda equina' and 'spinal cord' pain showing the lowest and highest consistent classification, respectively. The overall agreement for all three raters was 78% (140/180). The rate of agreement within each rater ranged from 67 to 83%. Consistent with our previous study, inter-rater agreement between the three raters for the second classification was 50%.

Discussion

Empirical validation of the various types of reliability (eg, inter-rater, test-retest) is one of the several psychometric characteristics that have yet to be examined among the numerous SCI pain classification schemes that have been proposed. Recently, we have examined the inter-rater reliability of several commonly used SCI pain classification schemes. The primary purpose of the current study was to extend our previous research by examining the test-retest reliability of the Donovan classification scheme, one of the more common SCI pain classification schemes.

In general, the test-retest reliability estimates for the Donovan classification scheme were in the moderately acceptable range, with 78% of the pain sites consistently classified over a 3-month interval. Moreover, test-retest reliability estimates increased to 87% (157/180) if inconsistent classification across the two different types of neuropathic pain were ignored (ie, collapsed into one group). It should be noted, however, that none of the pain types on the Donovan scheme demonstrated perfect agreement across classification periods. That is, between 16 and 34% of the pain sites were inconsistently classified across pain types.

The percentage of inconsistent classification was arguably more noteworthy since the same information was used to classify pain for each assessment. That is, raters viewed the same videotape and re-reviewed the results of the same structured interview when making the second classification after a 3-month delay. It may be argued that variability in both patient (eg, verbal pain descriptors used, pain characteristic variation (intensity), emotional status) and treatment provider characteristics (eg, variability in eliciting pain information) from one period to the next would likely result in a lower test-retest reliability estimate for the Donovan classification scheme. Thus, the test-retest reliability reported here may be considered the 'best-case' scenario since the exact same information was used for each classification.

It is important to note that the overall reliability of pain classification schemes includes examination across

various types of reliability estimates. For instance, testretest reliability of the Donovan scheme within each rater showed a consistent classification for about 75% of the pain sites. In contrast, inter-rater reliability estimates showed agreement across raters for only about 50% of the pain sites. The 50% inter-rater reliability estimate is consistent with our previous report after the initial pain classification by all three raters⁶ for these same participants. As previously discussed, it will be important in future research to more specifically distinguish the factors that contribute to disagreement, both within and across raters. In addition to reliability studies, it will be important to assess other psychometric characteristics of the many SCI pain classification schemes, including validity studies. For instance, the validity of a classification scheme to identify two broad types of pain, neuropathic versus musculoskeletal, may be assessed by comparing the classification of pain type to treatment response using medications known to influence either neuropathic or musculoskeletal pain. Clearly, there is an ongoing need to assess the psychometric properties of the various SCI pain classification schemes. Increased reliability and validity of pain classification schemes should facilitate SCI pain research and help target intervention efforts.

Acknowledgements

Resources for the production of this manuscript were provided by the University of Alabama at Birmingham from a Rehabilitation Research Training Center grant from the National Institute on Disability and Rehabilitation Research, Office of Special Education and Rehabilitation Services, Department of Education, Washington, DC, #H133B980016-00 and #H133N50009.

References

- 1 Hicken BL, Putzke JD, Richards JS. Classification of pain following spinal cord injury: literature review and future directions. Pain following spinal cord injury. International Association for the Study of Pain, 2001.
- 2 Berlowitz DR, Du W, Kazis L, Lewis S. Health-related quality of life of nursing home residents: differences in patient and provider perceptions. J Am Geriatri Soc 1995; 43: 799–802.
- 3 Donovan WH, Dimitrijevic MR, Dahm L, Dimitrijevic M. Neurophysiological approaches to chronic pain following spinal cord injury. *Paraplegia* 1982; **20**: 135–146.
- 4 Siddall PJ, Taylor DA, Cousins MJ. Classification of pain following spinal cord injury. *Spinal Cord* 1997; **35:** 69–75.
- 5 Tunks E. Pain in spinal cord injured patients. In: Bloch RF, Basbaum M (eds). *Management of Spinal Cord Injuries*. Williams & Wilkins: Baltimore, MD, 1986. pp 180–211.
- 6 Richards JS *et al.* Inter-rater reliability of the Donovan spinal cord injury pain classification Scheme. *Arch Phys Med Rehabil* 2002; **83:** 1290–1294.