OUTCOME ASSESSMENT IN COMPLEX REGIONAL PAIN SYNDROME

OUTCOME ASSESSMENT

IN

COMPLEX REGIONAL PAIN SYNDROME

By

TARA PACKHAM, BHSc(OT)

A Thesis

Submitted to the School of Graduate Studies

In Partial Fulfilment of the Requirements

For the Degree

Master of Sciences

McMaster University

MASTER OF SCIENCES (2011)
(Rehabilitation Sciences)

McMaster University
Hamilton, Ontario

TITLE: Outcome measurement in Complex Regional Pain Syndrome

AUTHOR: Tara Packham

SUPERVISOR: Dr. Joy MacDermid

NUMBER OF PAGES: 136

Table of contents

OUTCOME ASSESSMENT IN COMPLEX REGIONAL PAIN SYNDROME

Abstract:

Complex regional pain syndrome is a neurological condition characterized by a constellation of variable and seemingly disparate signs and symptoms for which there is presently no definitive diagnostic test. The opportunity exists for development of a condition-specific outcome measure for complex regional pain syndrome affecting any limb(s) that could be used by therapists, physicians and researchers to evaluate their patients, make treatment decisions, and monitor the changes in both impairments and quality of life experienced by those affected individuals. This thesis addresses outcome assessment in complex regional pain syndrome, incorporating 2 papers. The first is a systematic review of the current scope and psychometric rigor of outcome assessments available to health professionals to guide their management of this condition. The second paper describes the preliminary development steps of a new measurement tool for complex regional pain syndrome, with a focus on a cognitive debriefing study of current assessment practices and preferences of a cross-section of health professionals used to inform the definitions and user manual for a multi-disciplinary assessment. Finally, the thesis addresses areas for future refinement and testing of the proposed outcome measure.

Introduction and Background

Complex regional pain syndrome (CRPS) is a form of neuropathic pain that may develop after trauma or surgery, although is sometimes described as occurring spontaneously (Bruehl, Harden, and Galer et al,1999; Schwartzman, Erwin and Alexander, 2009). It can be associated with a nerve injury, and the International Association for the Study of Pain (IASP) diagnostic classification emphasizes this distinction, with CRPSI being defined as occurring without injury to a major nerve, while CRPSII has an accompanying nerve trauma (Albazaz, Wong, and Homer-Vanniasinkam, 2008). While there are few population based epidemiological studies, De Mos et al (2007) reported the incidence in a Dutch cohort at 26.2 per 100,000 person years (95% CI: 23.0–29.7), and additionally found a >3:1 ratio of women to men affected. The sample also found a higher rate of involvement in the upper extremity (59.2% versus 39.1%, $p < 0.001$), and reported that fractures were the most common precipitating event, accounting for 44% of cases. Pain is considered to be the key feature of the syndrome, often described as disproportionate to the injury, and spreading regionally beyond the original insult. Other symptoms may include swelling and changes in blood flow; trophic features such as changes in hair, skin and nails; motor symptoms such as stiffness, dystonia and guarding; and sensory alterations such as hypersensitivity and cold intolerance (Marinus and Van Hilten, 2006).

In a follow-up study of patients seen in a pain centre with the diagnosis of complex regional pain syndrome (CRPS) within a one year period, all reported

severe pain as the first symptom, and all but one respondent (n=56) also reported weakness and swelling as early signs of the condition (Galer, Henderson, Perander and Jensen, 2000). All respondents also reported that their condition caused substantial interference in their general activity, sleep, work, mobility, recreation, social activities, mood and relationships (Galer et al, 2000). The majority of patients seen in a 2009 study (n=102; De Mos et al) reported impairments persisting for more than two years since onset, and 31% (95%CI 19-43) had been unable to return to work.

CRPS- the measurement challenges

Historically, there has been a strong emphasis on diagnostic criteria in the field of CRPS; the debate has continued even after IASP published a benchmark taxonomy in 1994 in an attempt to create a unifying nomenclature for research and clinical practice. This new taxonomy was intended to replace former terms such as reflex sympathetic dystrophy (RSD), algodystrophy and causalgia (Reinders, Geertzen, and Dijkstra, 2002). However, the ongoing lack of agreement, coupled with the variable clinical presentation, has resulted in many small studies that cannot be compared to each other and whose claims to validity are not evidence based (Harden et al, 2007). Further contributing to small studies are the IASP diagnostic criteria themselves, which divides CRPS into two sub-types (CRPS I and II), on the basis of whether or not there is an accompanying nerve injury. For homogeneity, many studies focus only on CRPS

I, often including just upper or just lower limbs, reducing not only the sample sizes but also the generalizability of the research to the entire population. The challenge and opportunity still exists for an outcome measure that could be applied for both CRPS I and II, and which is capable of being utilized for any patient, regardless of which limb(s) have been affected. Furthermore, even as the search continues for a pathognomonic test to diagnose CRPS, there remains a role for discriminative assessment tools which can assist in identifying those patients who will not only have the condition, but also those who will require early intervention (Harden et al, 2010).

The work presented in this thesis will describe the present state of outcome assessment in CRPS, and the early development of a new outcome measure intended to address these concerns.

Clinical measurement concepts and principles

Clinical assessment tools are used in rehabilitation to serve one of three purposes: 1) evaluation – measuring change over time, or 2) discrimination – screening or classifying individuals according to a particular trait or risk factor, or 3) prognosis – classifying individuals as to whether they will respond to treatment or not (Finch, Brooks, Stratford, and Mayo, 2002). As the focus of this thesis is outcome measurement, this section will concentrate only on the measurement concepts and principles related to evaluation.

Reliability can be described as a measure of the reproducibility, agreement, or degree of error inherent in any measurement or measurement tool. Finch et al (2002) represent two essential components: *relative reliability*, or the ability of a tool to differentiate between subjects, presented numerically as the intraclass correlation coefficient (ICC), and *absolute reliability*, or the measurement error intrinsic to the scale, expressed as the standard error of measurement (SEM). It encompasses several forms, including the agreement between raters (intra-rater reliability) and consistency of results with any given subject (test-retest reliability). In theory, the reliability of an assessment tool can be improved by reducing the amount of errors through standardization of administration methods, training of those who administer the assessment, and by using rating scales with sufficient precision for discrimination (Streiner and Norman, 2003; Preston and Colman, 2000).

Validity encompasses multi-faceted analyses of the veracity attributed to an assessment tool: commonly examined forms include content, criterion, and construct validity, which in turn look at how the assessment tool can be used to describe, predict, and evaluate the concepts of interest (Tickle-Degnan, 2002). The evaluation of validity can be described as an ongoing process of accumulating evidence on multiple aspects of the test (Sechrest, 2005), including content, relationships to other variables, the internal structure of the assessment (Goodwin, 2002), the response process, and consequences (Cook and Beckman, 2006).

Content validity is studied by testing the comprehensiveness and/or accuracy of an assessment tool in measuring or embodying the constructs it purports to represent. Face validity is an extension of content validity: it is tested by surveying whether the intended users and/or target populations subjectively judge the tool to be reasonably representative (Hulley, Martin and Cummings, 2007). Internal consistency is sometimes described as construct validity, other times as reliability.

Construct validity is the degree to which we can reasonably estimate that an assessment reflects the hypothetical attributes or behaviours of both the tool, and the constructs of interest; it should be considered an estimate, however, as these relationships are usually complex and non-linear (Sechrest, 2005). *Discriminative* validity is an extension of construct validity; it could be defined as the ability of a measure to discriminate between two known groups (i.e. those

who have acute symptoms and those who have chronic symptoms) – this is also referred to as known group validity (Finch et al, 2002). This should be distinguished from *discriminant validity*, or the relationship of the tool findings about the constructs measured in contrast to the findings of another tool intended to measure different constructs within the same sample (Streiner and Norman, 2003). Criterion validity is how the findings of a tool compare to the findings of a gold standard test or widely accepted assessment in concurrent measures of the population of interest (Finch et al, 2002).

Responsiveness is described as the ability to detect clinically meaningful changes in status; it can be represented by both the effect size (ES) and standardized response mean (SRM) (MacDermid, Richards, Donner, Bellamy and Roth, 2000).

Use of Cognitive Interviewing to inform outcome measures

With the increased focus on incorporating the perspective of the patient in health care measurement, there has been a proliferation of self-report measures, and a corresponding drive to make them more scientifically rigorous (Kayes and McPherson, 2010). Cognitive interviewing is one such qualitative method: it is used to examine how potential users interpret, retrieve relevant information to formulate a decision, and respond to survey questions (Housen et al, 2008). Traditionally, qualitative studies seeking to refine clinician based (CB) assessment tools have used a Delphi or other consensus method for validation

of the content coverage (Fink, Kosecoff, Chassin, and Brook, 1984; Brunner et al, 2008). However, consensus methods do not generally allow for an understanding of the sources of difference. The process of constructing judgements from observations and assessments within clinical decision-making has been studied both quantitatively and qualitatively within the context of diagnosis (Norman, Coblentz, Brooks and Babcock, 1992; Regher and Norman, 1996), however, a recent (albeit not exhaustive) search of the health sciences literature could not produce a descriptive or qualitative study which focused on the process of observer variation across a variety of disciplines when scoring clinical attributes or status using clinician based (CB) assessment tools. Cognitive interviewing provides an opportunity to address the process by which health professionals arrive at a judgement when making scalar decisions about clinical signs or patient characteristics, and to investigate differences between the professions. This insight can inform the assessment tool and scoring manual to ensure that multidisciplinary values and practices are best reflected to enhance consistent application of the scoring guidelines. This thesis will present the results of a cognitive debriefing study of the a portion of an outcome measure for CRPS currently in development, the Hamilton Inventory for complex regional pain syndrome.

Development of the Hamilton Inventory for Complex Regional Pain Syndrome

Conceptual development of an outcome measure for CRPS

In 2007, several students in the MScOT program at McMaster University undertook a literature review under the supervision of the author (TP) to identify the characteristics attributed to CRPS, in an attempt to establish the parameters for content coverage for a new multi-disciplinary assessment tool, as recommended by Hicks (2004). The goals of developing a new tool were: 1) to provide a common taxonomic and assessment framework for therapists and physicians to evaluate their patients, make treatment decisions, and monitor progress; 2) to foster research by developing an inclusive tool that would allow patients with CRPSI or CRPSII, involving any extremity to be compared to each other; 3) to employ a broad framework which included impairment, activity and participation perspectives but was condition-specific; 4) to develop a reliable and valid tool that would yield useful and durable information for both research and clinical practice; and 5) potentially to identify patterns of symptoms, and thus discriminate between patient groups not on the basis of chronicity or nerve injury, but on the basis of functional symptom clusters. Interviews were also conducted with local 'champions' in CRPS care, and with patient 'stakeholders' to informally gather their input. An initial draft was developed which contained 99 items divided into two sections: a 44 item assessment to be completed by a health professional (covering the content areas of sensory symptoms, autonomic

function, trophic changes and motor function) and a self-report section containing 55 items addressing quality of life in three domains: physical symptoms, participation in daily activities, and socio-emotional impact. The tool was named the Hamilton Inventory for Complex Regional Pain Syndrome (HI-CRPS).

Initial steps of development

The first major refinement undertaken by the author was the revision of the scoring system. Many items were initially scored categorically, limiting the discriminative ability of the tool (Streiner and Norman, 2003; p. 31). Thus all a seven-point scale (0-6) was developed for all items, with a higher score indicating greater symptom severity. The seven-point scale was chosen for its anticipated contribution to reliability, validity and ease of completion (Preston and Colman, 2000). A combination of Likert and end-anchored adjectival forms were used, as incorporating both formats allowed easy ranking of both performance and attitudinal attributes. Ambiguous and double-barrelled items were also rewritten, and some questions were redistributed to minimize their potential influence on adjacent items.

It also became apparent that the 17 pain descriptors included in the symptoms subscale added unintentional weighting to that scale relative to the other subscales (Steiner and Norman, 2003; p. 104). Consequently, a group of 19 patients with CRPS of either the upper or lower extremity were recruited to assist in reducing the number of pain descriptors included in the symptom subscale of Section B. From the initial list of 17 descriptors, patients were asked

to select the 7 words that they felt best described the pain they experienced with CRPS; the seven words selected most frequently were then retained for the questionnaire, and the remainder were deleted (refer to table 1 for summary of endorsement). The combined effects of these changes reduced the number of items in Section B to 45, with 14 items on the symptoms subscale, 19 items on the function scale, and 12 items on the socio-emotional scale.

Table 1. Summary of patient selections of pain descriptors

Pain descriptor	# of endorsements	Status
Sharp	12	Keep
Hot	8	
Dull	2	
Cold	3	
Sensitive	11	Keep
Unpleasant	8	
Itchy	0	
Numb	4	
Throbbing	13	Keep
Shooting	8	
Stabbing	15	Keep
Aching	11	Keep
Pricking	4	
Burning	11	Keep
Spontaneous	1	
Deep intense	16	Keep
Radiating	6	

At this point, a formal pilot-testing study was submitted to the Research Ethics Board of Hamilton Health Sciences and McMaster University, consisting of 2 phases. Phase one requested review of the entire HI-CRPS (Draft 14) from a selection of international experts who had published on CRPS in the past 5 years, as well as to a group of clinicians treating CRPS in Southern Ontario (n=25), inviting them to review the tool using a structured questionnaire, including

rating content coverage, and making suggestions for adding or deleting items. Only 3 responses were received: 2 from local clinicians and the third from an internationally recognized researcher. One of the respondents did not complete the numerical rankings of components, but did comment on content and taxonomy. Given the poor response rate, these results are not presented here as they cannot be considered generalizeable. Phase two involved pilot-tested of the entire tool for inter-rater and test-retest reliability with student evaluators; again, the recruitment rate was very low (n=5) due to the short time frame of the study.

Phase three was a cognitive debriefing study of the clinician-based portion of the tool, and marks the transition of this endeavour from a clinical research project to a scholarly program of study. The cognitive debriefing study is described in detail in chapter two of this thesis.

The overall purpose of this thesis is to critically examine outcome measurement in complex regional pain syndrome and to consider the multidisciplinary assessment values and taxonomy related to the evaluation of the signs and symptoms of CRPS. Firstly, we undertook a systematic review of the quality and extent of psychometric evaluations published for current outcome measures specific to CRPS: this is the basis of chapter one of this thesis. Secondly, we conducted a cognitive debriefing study of the concepts covered by the clinician-based portion of the HI-CRPS tool, gathering definitions, scale

anchors, and assessment practices and preferences from a sampling of the intended user groups; the methods and results comprise chapter two.

Preface

This dissertation follows a sandwich thesis format. I, Tara Packham, am the first author for the two journal papers presented within. The ideas and work, including the design, data collection, data analysis, and manuscript preparation, associated with each paper are primarily my own. Drs. Joy MacDermid, James Henry and James Bain were coauthors. In their role as my supervisors, they provided guidance, critical feedback, and suggestions in regards to study design, data analysis, and writing up the findings. Dr. MacDermid also assisted in completion of the systematic review, acting as the second reviewer of the studies therein.

CHAPTER ONE

A systematic review of outcome assessments for CRPS: describing the elephant

Abstract

Purpose: To conduct a systematic review of the quality and extent of psychometric examinations of disease-specific outcome measures for complex regional pain syndrome (CRPS).

Method: Health database searches yielded 23 papers covering 19 assessment instruments. Each article was scored for quality using a 12-item structured tool; data was also extracted for comparison of tool content.

Results: Article quality ratings ranged from 25% to 88%. Six of the tools were specific to the upper extremity; 5 for the lower extremities while the remaining 8 were general. Many 'general' tools focused on a single construct, such as pain, skin temperature or allodynia. Most psychometric data was based on small studies (mean n=33); only one study addressed all relevant issues of reliability, validity and responsiveness.

Conclusions: Despite the variety of outcome measurement tools reported for CRPS rehabilitation, large gaps in both comprehensiveness and supporting psychometric evidence remain. The existing state of evaluation for CRPS might be illustrated by the analogy of a blindfolded person examining parts of an

elephant and forming an inaccurate conclusion about the essence of an elephant. Comprehensive, relevant and psychometrically sound tools for monitoring treatment outcomes are needed to address the pain and functional limitations experienced by this population.

Introduction

Complex regional pain syndrome (CRPS) is a neurological disorder characterized by a variable collection of signs and symptoms affecting either the upper or lower extremities. These signs may include pain, sensory disturbances, trophic changes, altered vascular/thermal regulation, edema, joint stiffness, and other motor impairments [1, 2]. The condition typically develops as a seemingly disproportionate response to some form of trauma [3]. Often, the patient reports increasing burning pain, swelling and stiffness that appears days to weeks after the injury, characteristically extending beyond the site of injury. Initially, these features may be difficult to distinguish from post-traumatic findings, but ultimately, they exceed the expectations for known damage [1]. Symptoms may vary with activity, environment and stress [4, 5], and often leads to a decrease in the spontaneous movement and function of the affected limb, as well as a decrease in participation in daily activities [6, 7].

There is no single diagnostic test that can accurately diagnose CRPS; hence, the condition is identified primarily on the basis of clinical assessment [8,9]. The ongoing emergence of new etiological theories, along with an ever-evolving understanding of the syndrome has contributed to a proliferation of tools

to assist in diagnosis and monitoring of the signs and symptoms [10]. Tools that can accurately diagnose require discriminative properties; and the multidimensionality of CRPS and lack of a uniform specific biological finding limit the ability to develop accurate diagnostic tests.

Despite difficulties in accurately making the diagnosis, there remains a need to assess patient status over time. This requires a different type of tool that is able to assess changes in the sign/symptoms of CRPS. Since the condition is a composite of multiple signs and symptoms, then the analogy might be one of describing an elephant by looking at its characteristic components. A number of assessment tools have been proposed for measuring the outcomes of treatments for CRPS. However, there have been no syntheses of the available information; nor consensus on which tools should be used in clinical research studies or practice. Conversely, recent systematic reviews addressing both medical and rehabilitative treatment recommendations for CRPS [11,12,13] have highlighted the inconsistent use of outcome measures to evaluate this population.

The purpose of this study was to conduct a systematic review of the quality and extent of psychometric evidence for disease-specific outcome measures for complex regional pain syndrome, and to consider if there is sufficient weight of evidence to recommend utilization of any of these tools for clinical practice and /or research involving persons with CRPS.

METHODS:

In September 2010, a systematic search was performed using Pub Med, Embase, Ovid Healthstar and Medline; see table 1 for search terms. The search was limited to articles available in English. The search revealed

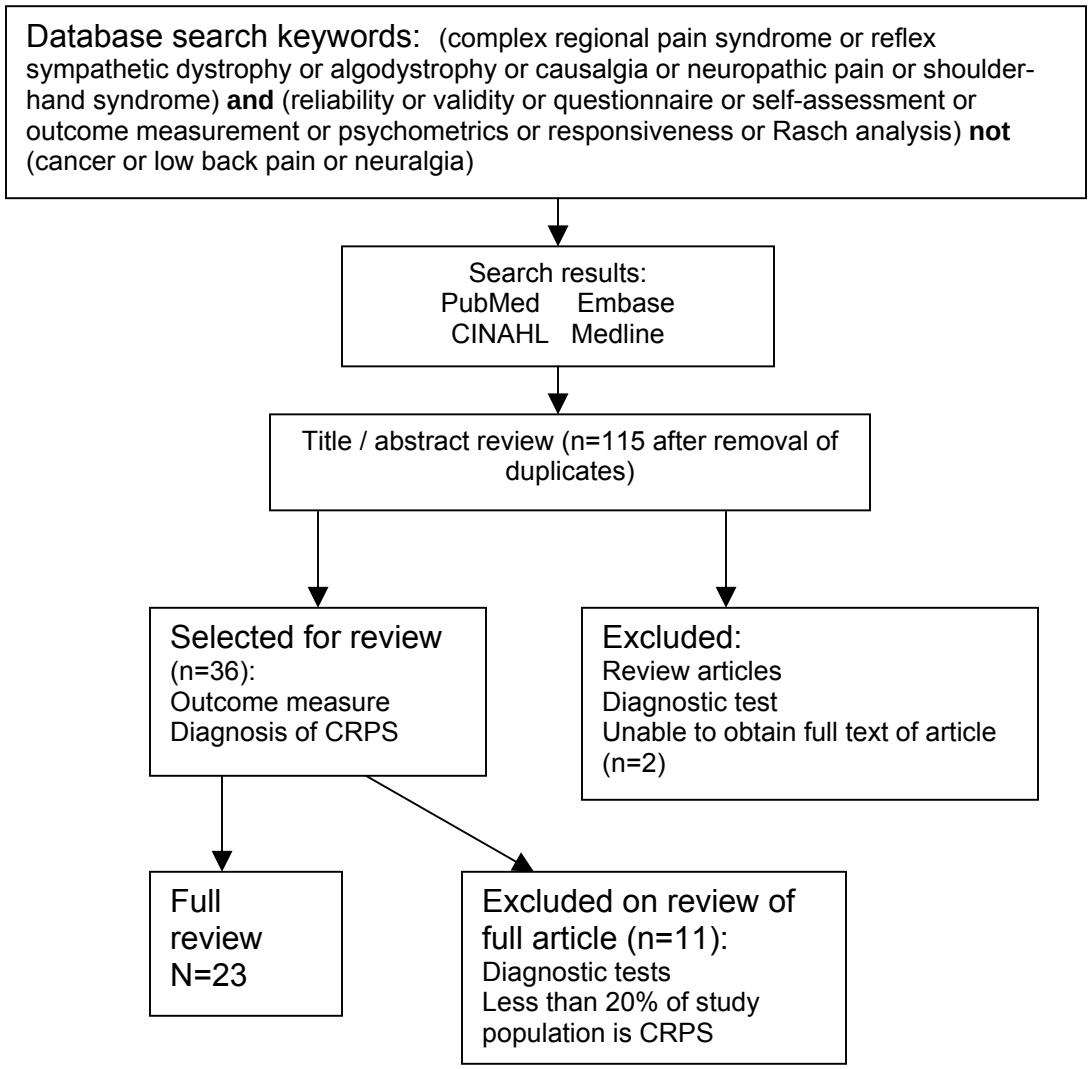
Table 1: Search terms for systematic review

Complex regional pain syndrome	Reliability
Reflex sympathetic dystrophy	Validity
Causalgia	Psychometrics
Algodystrophy	Questionnaire
Neuropathic pain	Responsiveness
Shoulder-hand syndrome	Rasch analysis
	Outcome measurement
	Self assessment
Terms within each column were combined with "or" and then the results of each of these searches were combined with "and"	

two distinct categories of evaluations: those intended for diagnosis, and those developed to measure outcomes. Since the focus of this review was tools to measure change in status over time, diagnostic tools were excluded. All abstracts were reviewed and diagnostic articles were removed from the set, leaving 36 articles for full review. We were unable to obtain two of the papers in print or electronic form. A further 11 articles were eliminated after examination of the full paper because a) they did not meet the inclusion criteria or b) less than 20% of the study population (or less than n=50) had a diagnosis of CRPSI or II. This yielded 23 papers to be critically appraised and scored for quality. Refer to figure 1 for a flow diagram of the systematic review process. These 23 studies addressed 19 different instruments; 9 were self-report and 8 were performance-based, and two combine components of both. Twelve of the instruments were

comprised of multiple subscales; 7 were unidimensional.

Figure 1: Flow diagram for systematic review



Critical appraisal and psychometrics summaries

All studies were appraised that a) addressed some aspect of psychometrics of an outcome assessment and b) at least 20% of the test population was comprised of persons with CRPS. Data on internal consistency, reliability, validity, and responsiveness was extracted from each paper using a structured instrument. This standardized 12-item evaluation instrument and accompanying guide was developed for appraisal of psychometric studies and has been used in previous systematic reviews of shoulder function, neck disability, and obesity-specific quality of life measures [14,15,16]. The 12 evaluation questions address the research question, study design, measurements, analysis and recommendations: each of the items is rated using a 3-point scale (0-2) to yield a total quality score out of 24, and the tool has a guide that provides descriptors to assist in ranking the items [17]; see appendix A for a copy of the tool. The reviewers (TP and JM) initially scored two articles jointly for a practice consensus round to ensure common interpretation of scoring guidelines, then appraised each of the 23 studies independently, and met to compare and discuss ratings until consensus was reached on the 12 quality score items for each article (see Table 2). All independent ratings were recorded and entered into SPSS 16.0 for analysis of agreement. Individual items had an average agreement rating of kappa = 0.87; the agreement for the total scores was ICC=0.98 (95%CI 0.96 - 0.99).

Table 2. Consensus quality ratings of paper (arranged highest to lowest)

Study Reference	Tool	Sample size	Quality Rating Items (see key below)												(%)Total
			1	2	3	4	5	6	7	8	9	10	11	12	
Geertzen et al [19]	Grip strength	29	2	1	2	2	0	2	2	2	2	2	2	2	**87.5**
Kemler and deVet [27]	Foot function	100/ 20	2	1	2	2	1	NA	2	2	2	2	1	1	**81.8**
Burnham et al [35]	Skin temp (IR thermom.)	17/ 17	2	1	1	2	0	NA	2	2	2	2	1	2	**77.3**
Foufanzar et al [34]	VAS pain	54	2	1	1	1	1	1	1	2	2	2	2	2	**75**
Perez et al [29]	Rising and Sitting Q; Walking Stairs Q	21	1	2	1	1	0	2	2	2	2	2	2	1	**75**
Roorda et al [31]	Walking Stairs Q	759	2	1	2	1	1	NA	2	1	2	2	1	1	**72.7**
VanEijs et al [33]	Brush Allodynia	36	1	2	1	1	1	0	2	1	2	2	2	2	**70.8**
Roorda et al [28]	Walking activity Q	981	2	1	2	1	1	NA	2	1	2	1	1	1	**68.2**
Geertzen et al [18]	ROM	29	1	2	1	1	0	NA	2	2	1	2	1	1	**63.6**
Heitz et al [30]	Rising and Sitting Q; Walking Stairs Q; Walking activity Q	52	2	2	1	1	1	NA	1	1	1	2	1	1	**63.6**
Roorda et al [32]	Rising and Sitting Q	759	1	1	1	1	1	NA	2	1	2	1	1	1	**62.5**
Krause and Backonja [38]	Neuropath. Pain Q	528/ 149	2	1	1	1	1	NA	2	1	1	1	0	2	**59.1**
Galer and Jensen [39]	NPS	288/ 78	2	1	1	2	1	1	1	1	1	1	1	1	**58.3**
Collins et al [40]	TReND symptom inventory	42/ 26	1	1	1	1	0	2	1	1	1	2	2	1	**58.3**
Oerlemans et al [21]	Impairment sum score (U/E)	45	1	1	1	1	0	0	2	2	1	2	1	1	**54.2**
Oerlemans et al [24]	Radboud skills Q (RSQ)	54	1	2	1	1	1	1	1	1	1	1	1	1	**54.2**
Dworkin et al [37]	SF McGill	882	1	2	1	2	1	1	1	1	1	1	0	1	**54.2**

Schasfoort et al [22]	Upper Limb Activity Monitor (ULAM) RSQ	30	1	1	2	1	0	NA	1	1	1	1	1	1	**50**
Brunner et al [25]	RSQ	57	2	1	0	0	1	0	0	1	1	2	2	1	**45.8**
Schasfoort et al [23]	Upper Limb Activity Monitor	4	1	2	1	1	0	2	1	1	1	0	0	0	**45.5**
Bianchi et al [20]	CRPS evaluation	31 (25)	0	1	0	0	0	1	2	1	1	1	0	2	**37.5**
Perez et al [26]	Impairment sum score (L/E)	43/ 58	1	1	1	1	0	0	1	2	0	1	0	0	**33.3**
Davidoff et al [36]	RSD assessment process	17	1	1	0	1	0	0	1	1	0	1	0	0	**25**

Key to Quality Rating Items			
1	Research Question (clarity and rationale)	7	Measurement procedures
2	Description of setting and participants	8	Standardization and reduction of bias
3	Hypotheses about reliability and validity	9	Analysis related to hypotheses
4	Scope of psychometrics under examination	10	Selection of appropriate statistical tests
5	Sample size	11	Use of benchmarks and confidence intervals
6	Recruitment and retention	12	Conclusions and recommendations

RESULTS

Summary of the outcome measures

Table 3 contains a brief descriptive summary of the 19 different assessment tools covered by the 23 papers, including the target populations, number and types of items included in the evaluation, administration time, and equipment required. None of the papers indicated that a manual was available for any of the assessment tools. Six upper extremity tools were revealed by this review: ROM of the upper limb [18]; grip strength dynamometry [19]; a composite CRPS

evaluation [20] combining pain visual analogue scale (VAS), swelling, AROM measures, and grip strength; the Impairment Sum Score (ISS) [21], including edema measured by volumeter, skin temperature measured by infrared thermometry, an AROM score, pain VAS, and pain descriptors from the short-form McGill; the Upper Limb Activity Monitor [22, 23], a portable sensor array that measures frequency and intensity of upper limb motion over a 24hr period in the clients' home; and the Radboud Skills Questionnaire (RSQ)[24, 25], a 45 item self-reported evaluation of personal care, domestic activities, and other activities (including work, social and leisure activities). Five lower extremity instruments also emerged: a lower extremity version of the ISS [26]; a four-part foot function evaluation [27] of 1) forward/backward shifting and 2) lateral shifting of a foot panel, 3) alternately touching two bells, and 4) depressing a pedal; and a trio of self-report questionnaires. The Walking Activity questionnaire [28, 29] covers walking indoors and outdoors, walking speed and the use of aids with 35 questions. The Walking Stairs questionnaire [29, 30, 31] includes 15 items, such as time, effort, avoidance, use of aids, and the need for assistance. The Rising and Sitting questionnaire [29, 30, 32] is a unidimensional scale consisting of 39 items, such as sitting/rising from high and low seats, getting on/off toilet, and getting in/out of the car.

Table 3. Description of Tools

Tool	Population	Items	Admin time	Cost/equipment
Impairment Sum Score [21,26]	CRPS 1 (U/E and L/E); also compared with normals for L/E group	Pain (VAS and partial McGill) Skin temperature Volume ROM	Not reported	Volumeter for edema measurement, goniometer for ROM, IR thermometer for temperature measures
Grip strength [19]	CRPS1	Dynamometry: Full fist grip 3 point pinch grip Lateral pinch grip	Not reported	Calibrated dynamometers
Foot function [27]	Normals and CRPS	Forward/backward shifting and lateral shifting of a foot panel Alternately touching 2 bells Depressing a pedal	Not reported	Wooden footboard with pedal and foot panel (generic instructions given for construction)
SF McGill [37]	Individuals with chronic pain, including neuropathic pain	22 items: pain descriptions rated on a 10 pt scale	Not reported	Pen and paper; licensing fee when used in research by for-profit organizations
Rising and Sitting [29, 30, 32]	CRPS1	39 items, such as: Sitting/rising from high and low seats Getting in/out of car Getting in/out of bed Getting on/off toilet	About 10 minutes	Pencil and paper
Walking Stairs [29, 30, 31]	CRPS1, other disorders affecting l/e function i.e. CVA, OA	15 items, such as: Time Effort Use of aids Need for assistance Avoidance	About 10 minutes	Pencil and paper
Walking Activity [28, 30]	CRPS1	35 items, such as: Walking indoors Walking outdoors Walking speed Use of aids	Not reported	Pencil and paper
RSD assessment process [36]	RSD (CRPS1)	Pain from joint palpation Edema Skin temperature AROM McGill VAS for pain	45 min	Volumeter for edema measurement, goniometer for ROM, IR thermometer for temperature measures

NPS [39]	Neuropathic pain, including CRPSI and CRPSII	10 questions: pain sharpness, heat/cold, dullness, intensity, surface / deep pain, and overall unpleasantness	15 min	Pen and paper
Radboud skills questionnaire [24, 25]	CRPS 1 (u/e)	45 items: Personal care Domestic activities Other activities (social, work, recreation)	10-15 min	Pen and paper
TreND symptom inventory [40]	CRPS1 and fibromyalgia	164 items in 10 subscales, i.e.: sensory, trophic, autonomic, motor, visceral symptoms	18-22 min	Pen and paper
ROM [18]	CRPS1 u/e	AROM of shoulder, elbow and wrist	Not reported	Inclinometer and goniometer
VAS pain [34]	CRPS1	Average pain intensity	Not reported	100mm VAS scales and rulers
Neuropathic Pain Questionnaire [38]	Neuropathic pain, including CRPS1	12 items:	Not reported	Pencil and paper
ULAM [22, 23]	CRPS1	Activity limitations based on recorded patterns of limb use, including speed and frequency of movements	Recordings based on a 24 hour monitoring period	Acceleration sensors, portable recorder, and software for analysis
Skin temperature (IR thermometry) [35]	Normals and CRPS (not specified if type I or II)	Skin temperature side differences: can be used on upper or lower extremities	Not reported	IR thermometers range from under $100 to $1500
CRPS evaluation [20]	CRPS (Not specified if type I or II)	Pain, swelling, ROM, grip strength	Not reported	Dynamometer and goniometer
Brush Allodynia [33]	CRPS1	Brush-evoked allodynia	Not reported	Soft brush

The remaining 8 tools addressed more global concepts including brush evoked allodynia [33]; reported average pain intensity with a VAS [34]; and skin temperature [35] with 3 different measurement instruments (thermistor, tympanic

and skin thermometers). The RSD assessment process [36] was a composite score including a numeric rating scale (NRS) for pain from joint palpation, edema measures, skin temperature, AROM, McGill, and VAS for pain, again reflecting a mix of observed and self-reported components. The most recent version of the short-form McGill pain questionnaire (SF-MPQ-2) [37] has been expanded to include 22 items rated on 10-point metric in 4 subscales (continuous pain, intermittent pain, predominantly neuropathic pain, and affective descriptors). The Neuropathic Pain Questionnaire (NPQ) [38] asks respondents to rank 12 items from a single scale with a NRS 0-100, including burning pain, increased pain due to weather changes, and questions such as “How overwhelming is your usual pain”, while the Neuropathic Pain Scale (NPS) [39] has 10 questions rated 0-10 such as pain sharpness, heat/cold, dullness, pain intensity, surface / deep pain, and overall unpleasantness. Finally, the Trauma Related Neuronal Dysfunction (TReND) questionnaire [40] is a self-report including 164 items in 10 subscales incorporating sensory, trophic, autonomic, motor, and visceral domains.

Quality of research

The consensus quality ratings of individual items for each paper can be found in table 2. Of the 23 studies reviewed covering 19 different outcome assessments, only 8 papers achieved a quality score of greater than 70% (or 18/24). These papers included examinations of 8 different tools, with the Dutch Walking Stairs questionnaire being the only tool represented by more than one higher quality

study. The highest quality score was 21/24 (88%); this was assigned to Geertzen et al [19] for their study of grip strength measurements in persons with CPRS; the next highest (82%) was Kemler and deVet [27] for their novel assessment of foot function. The lowest score was assigned to the earliest paper of the series by Davidoff et al [36] (RSD assessment process); however, this work was expanded by Oerlemans et al [21] to form the foundation for the Impairment Sum Score assessment tool. Overall, the mean quality score was 60%(± 16).

Study populations and sample size

One major limitation of all of the papers reviewed was the failure to address sample size and statistical power. When the study population was comprised exclusively of persons with CRPS, sample sizes ranged from 4 to 58, with an average of 33 participants (Table 2). Developers of new tools may be anxious to put them out for publication before large sample sizes have been acquired; however, most authors failed to acknowledge this is a limitation. Precise estimates of population scores and standard error depend on adequate sampling [41]. In the interest of homogeneity, most studies also limited their focus to persons with CRPSI and excluded those with CRPSII. Assessment tools were also frequently designed to focus on one area of the body, with 6 tools either designed or tested for the upper extremity; 5 for the lower extremities, and 8 generic tools (see table 4).

Depth of exploration of tools

None of the outcome tools have been studied for the full spectrum of psychometric properties; refer to table 4 for a summary of the examinations. The paper by Burnham et al [35] looking at skin temperature measurements reflects the greatest breadth of study: the paper covered test-retest reliability, construct validity, criterion validity, and responsiveness in a small sample of CRPS patients (n=17); internal consistency was not addressed as it did not apply (skin temperature was the singular measurement/construct). While many authors identified the need for reliable and valid outcome measures for use with the CRPS population, their concluding remarks often failed to recognize the need for further research to address gaps in the available psychometric data before considering the evidence produced by any given tool as sufficiently robust for use in research and clinical practice.

Summary of findings by psychometric properties

Internal consistency

Internal consistency reflects the degree to which the items on the tools, or a subscale, demonstrate commonality or inter-correlation. It can be measured by the correlation of individual items scores to the scores of the scale (or relevant subscale) as a whole [item total correlations, or ITC] and is often expressed using Cronbach's alpha [41]. It was frequently labeled as reliability in the papers we reviewed. However, since internal consistency reflects a consistent

Table 4. Outcome Measure Summaries

Tool	Constructs examined	Quality score (%)	Languages	Self report	Properties Studied			
UPPER EXTREMITY TOOLS								
Impairment sum score [21]	Composite score of ROM, edema, temp diff, and pain	54.2	Dutch version of McGill	One component	#		◊ *	%
Grip strength [19]	Grip strength	87.5	N/a	N	_	$ @		
ULAM [22,23]	Frequency and velocity of u/e movements	50 45.5	N/a	N			*	
Radboud skills questionnaire [24, 25]	Activity and participation limitations	54.2 50 48.5	Dutch Dutch German	Y	#	$ @	◊ * *	
CRPS evaluation [20]	CRPS symptom inventory	11/24	Italian	N				%
ROM [18]	Joint ROM	13/22	N/a	N	_	$ @		
LOWER EXTREMITY TOOLS								
Impairment sum score [26]	Composite score of ROM, edema, temp diff, and pain	33.3	Dutch McGill	One component			*	%
Walking Stairs [29, 30, 31]	Stair climbing in community dwelling adults	81.8 75 63.6	Dutch German	Y	 # 	$ $ 	♪ ♪ *	
Rising and Sitting [29,30,32]	Sit to stand	75 63.6 62.5	Dutch German	Y	 # 	$ $ 	 * ◊ *	
Walking activity [28, 30]	Walking in community dwellers	77.3 63.6	Dutch German	Y	#		♪ *	
Foot function [27]	Foot function	87.5	N/A	N		$ @	*	

Tool	Constructs examined	Quality score	Languages	Self report	Properties studied			
GENERALIZED ASSESSMENT TOOLS								
TReND symptom inventory [29]	Symptoms of CRPS, FM or RSI	58.3	Dutch	Y		$	*	
VAS pain [34]	VAS pain	75	N/A	Y	_	$	*	
RSD assessment process [36]	Composite assessment incl. McGill, skin temperature differences, edema	25	N/A	N	#			
NPS [39]	Neuropathic pain	58.3	English	Y			♪	%
SF McGill [37]	Pain descriptions and severity	50	English	Y	#		* ◊ ▲	
Neuropathic Pain Questionnaire [38]	Neuropathic pain	59.1	English	Y	#		♪ ▲	
Skin temperature (IR thermometry) [35]	Skin temperature side differences	77.3	N/a	N	_	$	■ *	%
Brush Allodynia [33]	Brush-evoked allodynia	75	N/a	N	_		♪	

KEY:
# internal consistency
$ test-retest reliability
@ inter-rater reliability
* construct validity
◊ content validity
♪ discriminant validty
% responsiveness
■ criterion validity
▲ factorial validity

response to different items of the scale at a single point in time, it provides little direct evidence on the important aspects of reliability for clinical measurement [42]. Analysis of internal consistency can be performed on both cross-sectional and longitudinal data; however, in our review, this information was only reported in 8 of 19 studies where it would have been possible and appropriate to collect and analyze this data (see table 4). Using longitudinal data, Oerlemans et al [24] reported high Spearman's coefficients of >0.80 for tester A and "mostly above 0.60" (p. 238) for tester B using the RSQ with persons experiencing CRPSI in a single upper extremity. Krause and Backonja [38] found high internal consistency (Cronbach's alpha= 0.95) amongst the initial 32 items of the NPQ. Since overly high internal consistency indicates redundancy of items it was appropriate that they proceeded to develop a shorter version of the scale; however internal consistency was not reported for the final 12-item scale.

Reliability

Reliability is a reflection of the consistency or true variability reflected in the scores of an assessment tool; higher levels of reliability can give the user confidence that any differences in measurements between individuals are due to true variation rather than random or systematic error [41]. Most of the studies that assessed reliability (46% of the total papers) used reliability coefficients expressed as the intra-class correlation coefficient (ICC). These values indicate relative reliability as they compare the variability between subjects to the total variability as a ratio. Inter-rater reliability does not apply to self-report measures;

and was reported in 4/9 of the clinician-based measures. Kemler and DeVet [27] reported inter-rater reliability ranging from ICC=0.85-0.99 while simultaneously scoring patients using their foot function test board, however, these ratings were based only on the normative sample of healthy adults. Oerlemans et al [24] used a different method of comparison to evaluate the oral administration of the RSQ. Using 2 different raters, they compared the coefficient of variation (CV) between them, reporting a median CV of 2.3% (range 0%± 5 .1%) for tester A, while the median for tester B was 4.5% (range 0%± 18 .9%).

Test-retest reliability, which assesses stability of measures over periods of time where the patient should is not changing, can apply to both self-report and clinician-based measures and was assessed in 9/19 of the measures (refer to table 4). The Dutch Walking Stairs Questionnaire (WSQ) and Rising and Sitting Down Questionnaire (RSDQ) were examined by Perez et al [26] in persons with CRPSI of a single lower extremity: they reported test–retest reliability as measured by ICCs >0.78 (WSQ) and >0.84 (RSDQ). Excellent test-retest reliability was reported for the TReND questionnaire with ICCs > 0.90 in both the CRPS sample and overall group of persons with CRPS and fibromyalgia [40]. Burnham et al [23] compared three different devices for measuring skin temperature, and found high ICCs (>0.96) for all three methods: thermistor, tympanic thermometer, and IR skin thermometer.

Reliability coefficients can also be determined by using generalizability theory to simultaneously consider different sources of error variance; however,

this approach is uncommon, being only reported by a single paper [18] in this review. That study examined the reliability of ROM measures considering multiple sources of measurement error and found that all variables (ROM measures for individual joints) were influenced by the interaction of the patient, session and observer: this resulted in ICCs ranging from 0.55 to 0.97 for different joints. The smallest detectable difference, an extension of the standard error of measurement, ranged from 7 degrees for elbow flexion to over 25 degrees for external rotation of the shoulder.

Validity

When considering the measurement properties of a tool, it is possible to examine multiple forms of validity – all intended to test different facets of the truthfulness of the data gathered by the assessment tool. See table 4 for a compiled account of the types of validity investigated by the papers included in this review.

While multiple papers attributed *content validity* to their tool (see table 4), only two [21, 24] described in detail a process for formal testing. Content validity methods included a Delphi consensus round with experts to verify that their outcome measures reflected the spectrum of symptoms experienced by persons with CRPS, and were not missing any key signs or symptoms. Other authors invoked content validation based on literature reviews or when test items appeared to cover areas identified as key domains by others [29,30]. While Dworkin et al [37] did not overtly refer to content validity, they did in fact conduct

focus groups to assist in the generation of meaningful descriptors of pain qualities, and then had over 800 persons living with pain review the items developed for the newest iteration of the SF-MPQ-2 as part of an on-line survey.

Factor analysis was employed by Dworkin et al [37] to scrutinize their revised scale. They used data from their web survey to cluster the new items into subscales, and then performed a confirmatory analysis of the old and new sub-scales (goodness of fit index >0.90 for all 4 sub-scales) once the overall tool revisions had been finalized – this could be considered *factorial validity*. Krause and Backonja [38] also used factor analysis to assist in reducing items for the NPQ from the initial group of 32 down to 12 items, ensuring they retained items that loaded strongly (>0.60) onto each of the factors that related to their core theoretical concepts.

Five papers examined *discriminant validity*, which includes both the ability of the tool to identify known groups and the aspect of contrasting against other tools measuring different constructs. Galer and Jensen [39] demonstrated the ability of four items from the NPS [sharp, cold, sensitive and itchy) to discriminate between persons with post-herpetic neuralgia and those with other types of neuropathic pain – including those with CRPS- using repeated analysis of variance (ANOVA) and $p<0.01$. Kemler and deVet [27] validated their assessment of foot function by comparing the bilateral scores of healthy participants with bilateral scores from a group of persons with CRPS of one lower extremity; they ratified their hypothesis that persons with CRPS of a single lower

extremity would perform poorer than healthy controls (matched for age, sex and dominance) on both their affected and unaffected side ($p<0.05$). Van Eijs et al [33] validated the ability of brush-evoked allodynia testing to predict which persons with CRPSI would respond positively to spinal cord stimulation; they found a mean allodynia score of below 2.5/10 predicted those who would benefit from the procedure (sensitivity of 0.75 and specificity of 0.81).

Construct validity, or how closely the tool behaves to how the developers predict it will (based on their theoretical foundation) was widely considered by fourteen of twenty-one papers; often this took the form of comparing the results of the measure against the results of other tools. Burnham et al [35] found strong correlations between several methods for measuring skin temperature (thermistor, tympanic or skin) and intramuscular temperature readings at the same body sites ($r >0.90$). Kemler and deVet [27] verified their postulation that the assessment of foot function would have low correlations with existing tools such as the '3 minute walk' ($r=0.30$, $p>0.05$) and 'timed up and go'($r= -0.01$, $p>0.05$) because the individual tasks (pressing a pedal, ringing a bell) measured different components of foot impairment as compared to the mobility-based assessments. Dworkin et al [37] extensively evaluated the construct validity of the SF-MPQ-2 by making multiple comparisons to other recognized measures of pain intensity, activity and participation limitations related to pain, finding (among others) positive correlations with the Brief Pain Inventory (BPI) average pain score ($r=0.60$, $p<0.001$) and Multidimensional Pain Inventory interference

scale(r=0.54, p<0.001), and total numbers of days spent in bed because of pain (r=0.31, p<0.001). In a cross-sectional study, Heitz et al [30] contrasted three lower extremity functional questionnaires [Walking Stairs, Rising and Sitting, and Walking Activity] with concurrent VAS scores for both pain (R=0.25, p<0.001) and restrictions in activities of daily living (R=0.37, p<0.001).

The least frequently examined aspect of validity was that of *criterion validity*, or how well the results of an assessment correlated with the findings of a gold standard: only one paper weighed their evaluations against a recognized measure. Burnham et al [35] compared readings taken by two types of infrared thermometers against the skin temperature readings taken by a thermistor. The infrared device measurements were significantly cooler than those recorded by the thermistor, with the mean infrared skin thermometer measurement being 0.2°C cooler and the mean infrared tympanic thermometer measurement 0.5°C cooler (p<0.001).

Responsiveness

Responsiveness (the ability of the assessment tool to detect change over repeated evaluations when true changes exist) was addressed for 5 of the 19 tools (refer to Table 4). This is prerequisite if the measure is to be used to assess clinical changes over time [43].

The CRPS evaluation scale proposed by Bianchi et al [20] reported responsiveness in a subset of upper extremity subjects; they reported the mean score dropped from 77% at baseline to 9% at one year after treatment with

corticosteroids and physiotherapy, but did not calculate effect size. Both the upper extremity [21] and lower extremity [26] versions of the Impairment Sum Score were examined for responsiveness. They reported adequate responsiveness in both studies, but each measured the concept differently. Oerlemans et al [21] demonstrated a large treatment effect for treatment with radical scavengers, reporting a 44% reduction in ISS over a 12-month period, with an effect size of 2.29. Perez et al [26] compared percentage changes in ISS with patient-reported global perceived change scores (using a 4 point scale of amelioration, no change, deterioration, and unknown/undecided) and physicians' perceptions of CRPS severity for the individual subjects (using a 5 point scale none/light/moderate/severe/extreme). They described ISS score changes as demonstrating a statistically significant difference between those groups reporting improvement (mean ISS dropped by 6.8 points, $p=0.01$) vs. no change (ISS decreased by 0.9 points, $p=0.07$) vs. deterioration (mean ISS increased by 3.8, $p=0.02$), and acknowledged the need for more rigorous testing of responsiveness [26].

Burnham et al [35] evaluated the responsiveness of 3 different skin temperature measurement instruments by comparing temperatures before and after IV regional sympathetic blockade for treatment of persons with CRPS: the sample included 5 upper and 12 lower limbs. They reported responsiveness as an index calculated as described by Guyatt et al [44]: the values for the three instruments ranged from 3.6 to 4.2, and were all considered to be good.

Galer and Jensen [39] looked at the mean changes in scores of individual items on the NPS in a heterogeneous chronic pain population examining pre and post treatment for two different interventions, and used ANOVA to identify which items differed significantly over time: this is not a recommended method of evaluation [41]. The “unpleasant” and ”deep” pain descriptors were the only items to reach statistical significance ($p<0.01$) for the interaction of pre/post differences and method of treatment. They did not use standard responsiveness indices, nor did they determine responsiveness for the NPS as a whole. Dworkin et al [37] also investigated the responsiveness of the SF-MPQ-2, however, their study population was a subset of the overall study comprised exclusively of persons with diabetic neuropathy. Thus the responsiveness of the short-form McGill still needs to be examined with a CRPS population. While Geertzen et al [18] did present data on the smallest detectable difference in their cross-sectional study of ROM, no studies in this review included longitudinal data on minimum detectable change or clinically important differences.

Item –response or Rasch analysis

The majority of psychometric evidence found was based on classical test theory (CTT). However, the use of more recent clinical measurement (item response) theory was evident in two of the more recent papers. Roorda et al [31] found that the Walking Stairs scale was suitable for measuring patients, as demonstrated by a strong fit with the monotone homogeneity scale ($H=0.50$) and was suitably hierarchical ($H^1=0.58$) based on a non-parametric item response

Mokken scale analysis in a heterogeneous sample of patients with CRPS1 of the lower limb. The authors also examined concepts that would be comparable to classical reliability (intratest reliability coefficient = 0.90, acceptable for individual patient decisions), and divergent validity (differential item functioning [DIF]). No differential item functioning was found for age or sex, allowing comparisons between age groups and males/females; however some DIF was evident for diagnostic groups, limiting the confidence to compare across diagnostic groups [31]. In a subsequent paper, Roorda et al [32] also used IRT to model unidimensionality of the Rising and Sitting Down questionnaire in persons with mobility disorders (including CRPS1 of the lower limb). This essentially tests the assumption that the items only test a singular trait (comparable to the CTT concept of factorial validity). Thirty-nine of 42 items loaded on 1 component, explaining 59% of the variance; these were maintained for the final version of the scale. Good intratest reliability was demonstrated, with Cronbach's alpha = 0.96 for the scale [32].

DISCUSSION

This systematic review indicated that there is a wide variety of assessment tools that have been reported for potential use in complex regional pain syndrome. However, significant gaps in both comprehensiveness and supporting psychometric evidence still exist. The current state of evaluation for CRPS might be illustrated by the analogy of a blindfolded person examining parts of an

elephant and then conveying to others with great inaccuracy their perspective on the essence of an elephant. Scales that focus on limited elements of CRPS, or that have not been validated against “the truth” are likely to provide unstable and potentially false representations of the nature and severity of this complex syndrome.

We conducted a systematic review as a way of synthesizing the content and quality of existing knowledge about disease-specific measures for complex regional pain syndrome. We found 19 tools that had been described by 23 articles available in English: this included 3 tools developed in English, 8 measurement tools studied only in Dutch, German, and/or Italian; and 8 tools where scoring was not language-dependent (see table 4). Nine of the tests were self-reported questionnaires available in the public domain at no cost, and able to be administered with a pencil and paper; 6 more utilized standard equipment such as goniometers and hand dynamometers typically available to clinicians in their practice settings (refer to table 3). Consensus rating of quality ranged from 25% to 88% (mean 60%) suggesting that the overall quality of studies in this area is relatively low. This indicates the need for additional research that would either develop new psychometrically solid evaluation tools or contribute stronger efforts on existing tools. While some promising objective assessments exist for key syndrome components like skin temperature differences, foot function and grip strength, this review did not identify a comprehensive assessment tool with sufficient psychometric support to provide a

reliable outcome tool incorporating the unique signs and symptoms experienced by these patients. To return to our analogy, focusing on isolated parts like the trunk or the tail may limit one's ability to recognize the elephant.

The concepts of reliability, validity and responsiveness as we have described them in this paper are based on classical test theory, a theoretical framework that focuses on the test as a whole, rather than individual items. When this framework is utilized, it constrains comparison of the results (like the calculations of reliability) to only like populations [41]. For this systematic review, only papers including persons with CRPS above a minimum threshold of 20% or n=50 in their study samples were included so as not to violate this principle.

In this review, 11 of 23 papers looked at some form of reliability. Not included in this number are two papers [21, 26] based on the ISS tool: while both papers reported that they considered reliability, they did not actually calculate the overall reliability estimates of the assessment. Instead, they simply reported previously published values for reliability of each of the components; they justified this approach because they felt the ISS was simply a mathematical calculation "… *independent from the observer. Therefore, computing test-retest and interobserver reliability was considered to be redundant*" (p. 986, Oerlemans et al, [21]).

The difficulty in establishing "gold standard" comparators for issues of pain and disability is reflected in the paucity of examinations of criterion validity found in this review. However, given the diverse spectrum of established tools

used for construct validity comparisons when evaluating newer tools, it is clear that a singular standard does not presently exist in this area. Additionally, this reinforces that there is a wide variety of clinically important criteria for the assessment of CRPS signs and symptoms: the same challenge that has bedeviled those seeking consensus for diagnostic standards [8,45,46]. Although a new severity scale has just been proposed as a comprehensive diagnostic test that is sensitive and specific to the complexities of this syndrome [10], no single evaluative standard has yet been recognized for outcome measurement. A broad range of concepts is dissected in the existing assessments, including pain, swelling, ROM, strength, skin temperature, mobility, participation and independence: these are appraised from multiple frameworks, including both impairments measured by health professionals and patient self-reports of disability. Nonetheless, a comprehensive and rigorous assessment for CRPS remains an unattained goal.

Although component assessments like skin temperature [35] and foot function [27] were based on high quality studies, no existing outcome measure has yet demonstrated an accurate and inclusive approach to outcome measurement in CRPS. As clinicians and scientists continue in pursuit of such a goal, it may be worthwhile to consider the possibilities of developing and testing tools that do not further divide this population into the subcategories of CRPSI and CRPSII and upper and lower extremity problems. This would facilitate larger samples in all forms of research and encourage wider use of these tools across

health care, enabling clinicians from many domains to utilize common assessment techniques and taxonomies. Additionally, given that almost half of the published tools are not currently available in English, translation into English and other major languages (and appropriate cultural validation) would also facilitate international collaborations and comparisons. Further research is also required to solidify the psychometric foundations of most of the tools, including:

1. filling in methodological gaps. For example, 74% of the tools have not been assessed for responsiveness, or the research has not yet been published. None of the papers reported information on minimum detectable differences or clinically important differences, further inhibiting the incorporation of these measures into clinical practice.
2. repeating evaluations with larger sample sizes of adequate power to confirm the preliminary results found with smaller samples, and
3. including expanded descriptions of study populations in publications to elucidate demographic comparators such as duration of symptoms, work status, and third-party funding issues that help clinicians to identify whether the population in the studies are similar to the population that they see in their practice.

The second purpose of this systematic review was to consider if there is sufficient weight of evidence to recommend utilization of any particular tool or tools for clinical practice and /or research involving persons with CRPS. While several tools have a well-designed and described evaluation of their specific construct (i.e. grip strength [19]; skin temperature [35]; foot function [27]), their generalizability is limited by small sample sizes. Information on minimum important differences and clinically important differences was not presented. When all of these considerations are taken into account, it is challenging to make any unequivocal recommendations in favour of any particular tool presented herein. Evidence-informed health professionals will need to weigh the information for each unique tool to consider which assessments can provide clinically relevant information to make individual patient decisions for the types of CRPS patients seen in their practice, and the outcomes that they and their clients consider important [43]. If our shared goal is to find successful treatments to address the pain and functional limitations experienced by this population, then we must not only seek to unlock the mechanisms by which it develops, and provide timely and accurate diagnosis, but we must also be able to effectively monitor the recovery process. Only then may we truly be able to describe the elephant.

Chapter One References

[1] Vacariu G. Complex regional pain syndrome. Disabil Rehabil 2002; 24(8):435-42.

[2] van Rijn M, Marinsu J, Putter H, and van Hilten JJ. Onset and progression of dystonia in complex regional pain syndrome. Pain 2007;130(3):287-93.

[3] Taha R, Blaise G. Editorial: Is complex regional pain syndrome an inflammatory process? Theories and therapeutic implications. Can J Anesth 2007; 54(4):249-53.

[4] Li Z, Paterson Smith B, Smith TL, Koman LA. Diagnosis and Management of Complex Regional Pain Syndrome Complicating Upper Extremity Recovery. J Hand Ther 2005 18(2):270-8.

[5] Geertzen JHB, Van Wilgen JB, Schrier E, Dijkstra PU. Chronic pain in rehabilitation medicine. Disabil Rehabil 2006; 28(6):363-67.

[6] Schasfoort F, Bussmann JB, and Stam H. Impairments and activity limitations in subjects with chronic upper limb CRPS I. Arch Phys Med Rehabil 2004;85(4):557-66.

[7] Marinus J, Van Hilten JJ. Clinical expression profiles of Complex Regional Pain Syndrome, Fibromyalgia, and a-specific Repetitive Strain Injuries: more common denominators than pain? Disabil Rehabil 2006; 28(6):351-62.

[8] Harden RN, Bruehl S, Perez RSGM, Birklein F, Marinus J, Maihofner C et al. Validation of proposed diagnostic criteria (the ‘‘Budapest Criteria”) for Complex Regional Pain Syndrome. Pain 2010; 150:268–74.

[9] de Mos M, Huygen FJPM, van der Hoeven-Borgman M, Dieleman JP, Stricker BHC, Sturkenboom MCJM. Outcome of the Complex Regional Pain Syndrome. Clin J Pain 2009; 25:590–7.

[10] Harden RN, Bruehl S, Perez RSGM, Birklein F, Marinus J, Maihofner C, Lubenow T, Buvanendran A, Mackey S, Graciosa J, Mogilevski M, Ramsden C, Schlereth T, Chont M, Vatine JJ. Development of a Severity Score for CRPS. Pain 2010;151:870-6.

[11] Simpson EL, Duenas A, Holmes MW, Papaioannou D, Chilcott J. Spinal cord stimulation for chronic pain of neuropathic or ischaemic origin: systematic review and economic evaluation. Health Technol Assess 2009;13(17):1-154.

[12] Daly AE, Bialocerkowski AE. Does evidence support physiotherapy management of adult Complex Regional Pain Syndrome Type One? A systematic review. Eur J Pain 2009 Apr;13(4):339-53.

[13] Ezendam D, Bongers RM, and Jannink MJ. Systematic review of the effectiveness of mirror therapy in upper extremity function. Disabil Rehabil 2009;31(26):2135-49.

[14] Roy SB, MacDermid JC, Woodhouse LJ. Measuring shoulder function: a systematic review of four questionnaires. Arthritis Rheum 2009; 61(5):623–32.

[15] MacDermid JC, Walton DM, Avery S, Blanchard A, Etruw E, McAlpine C, and Goldsmith CH. Measurement properties of the neck disability index: a systematic review. J Orthop Sports Phys Ther 2009; 39(5):400-17.

[16] Forhan M, Vrkljan B and MacDermid JC. A systematic review of the quality of psychometric evidence supporting the use of an obesity-specific quality of life measure for use with persons who have class III obesity. Obesity reviews. 2010; 11:222–228.

[17] MacDermid JC. Critical appraisal of study design for psychometric articles evaluation form and interpretation guide. In: Law MC, MacDermid J (eds). Evidence Based Rehabilitation: A Guide to Practice. Slack: Thorofare, NJ, 2008, p. 387–92.

[18] Geertzen JHB, Dijkstra PU, Stewart RE, Groothoff JW, Jan Ten Duis H, Eisma WH. Variation in measurements of range of motion: a study in reflex sympathetic dystrophy patients. Clin Rehabil 1998; 12:254–64.

[19] Geertzen JHB, Dijstra PU, Stewart RE, Groothoff JW, tenDuis HJ, and Eisma WH. Variation in measurements of grip strength: a study in RSD patients. Acta Orthop Scandan 1998 Suppl 279; 69:4-11.

[20] Bianchi C, Rossi S, Turi S, Brambilla A, Felisari G, Mascheri D. Long-term functional outcome measures in corticosteroid-treated complex regional pain syndrome. Eura Medicophys 2006; 42(2):103-11.

[21] Oerlemans HM, Goris RJA, Oostendorp RAB. Impairment level sumscore in reflex sympathetic dystrophy of one upper extremity. Arch Phys Med Rehabil 1998;79:979-90.

[22] Schasfoort FC, Bussmann JBJ, Krijnen HJ, Stam HJ. Upper limb activity over time in complex regional pain syndrome type 1 as objectively measured with an upper limb-activity monitor: An explorative multiple case study. Eur J Pain 2006;10(1):31-9.

[23] Schasfoort FC, Bussmann JB, Stam HJ. Correlation between a novel upper limb activity monitor and four other instruments to determine functioning in upper limb complex regional pain syndrome type I. J Rehabil Med 2005; 37:108–14.

[24] Oerlemans HM, Cup EH, DeBoo T, Goris RJ, Oostendorp RA.
The Radboud skills questionnaire: construction and reliability in patients with reflex sympathetic dystrophy of one upper extremity. Disabil Rehabil 2000; 22(5): 233-45.

[25] Brunner F, Heitz C, Kissling R, Kessels AGH, Perez RSGM, Marinus J, ter Rien G, Bachmann LM. German translation and external validation of the Radboud Skills Questionnaire in patients suffering from Complex Regional Pain Syndrome 1 BMC Musculoskeletal Disorders 2010; 11:107.

[26] Perez R, Oerlemans HM, Zuurmond W, De Lange J. Impairment level Sum Score for lower extremity Complex Regional Pain Syndrome type I. Disabil Rehabil 2003; 25(17):984-91.

[27] Kemler MA, De Vet HCW. An objective and standardized test of foot function: normative values and validation in patients with reflex sympathetic dystrophy. Arch Phys Med Rehabil 2000; 81:1401-7.

[28] Roorda LD, Roebroeck ME, van Tilburg T, Molenaar IW, Lankhorst GJ, Bouter LM, and the Measuring Mobility Study Group. Measuring activity limitations in walking: development of a hierarchical scale for patients with lower-extremity disorders who live at home. Arch Phys Med Rehabil 2005; 86:2277-83.

[29] Perez RS, Roorda LD, Zuurmond WW, Bannink II, Vranken JH, de Lange JJ. Measuring perceived activity limitations in lower extremity Complex Regional Pain Syndrome type 1 (CRPS I): test-retest reliability of two questionnaires. Clin Rehabil 2002;16(4):454-60.

[30] Heitz C, Bachmann LM, Leibfried A, Kissling R, Kessels AGH, Perez RSGM, Marinus J, Brunner F. Translating the Dutch walking stairs, walking ability and rising and sitting questionnaires into German and assessing their concurrent validity with VAS measures of pain and activities in daily living. BMC Musculoskeletal Disorders 2010; 11:108.

[31] Roorda LD, Roebroeck ME, van Tilburg T, Lankhorst GJ, Bouter LM, Measuring Mobility Study Group. Measuring activity limitations in climbing stairs: development of a hierarchical scale for patients with lower-extremity disorders living at home. Arch Phys Med Rehabil 2004; 85:967-71.

[32] Roorda LD, Molenaar IW, Lankhorst GJ, Bouter LM, and the Measuring Mobility Study Group. Improvement of a questionnaire measuring activity

limitations in rising and sitting down in patients with lower-extremity disorders living at home. Arch Phys Med Rehabil 2005;86:2204-10.

[33] van Eijs F, Smits H, Geurts JW, Kessels AGH, Kemler MA, van Kleef M, Joosten EAJ, Faber CG. Brush-evoked allodynia predicts outcome of spinal cord stimulation in complex regional pain syndrome type 1. Eur J Pain 2010; 14(2):164-9.

[34] Forouzanfar T, Kemler M, Kessels AG, Köke AJ, van Kleef M, Weber WE. Comparison of multiple against single pain intensity measurements in complex regional pain syndrome type I: analysis of 54 patients. Clin J Pain 2002; 18(4): 234-7.

[35] Burnham RS, McKinley RS, Vincent DD. Three types of skin-surface thermometers: A comparison of reliability, validity, and responsiveness. Am J Phys Med Rehabil 2006; 85:553–8.

[36] Davidoff G, Morey K, Amann M, Stamps J. Pain measurement in reflex sympathetic dystrophy syndrome. Pain 1988; 32(1):27-34.

[37] Dworkin RH, Turk DC, Revicki DA, Harding G, Coyne KS, Peirce-Sandner S, Bhagwat D, Everton D, Burke LB, Cowan P, Farrar JT, Hertz S, Max MB, Rappaport BA, Melzack R. Development and initial validation of an expanded and revised version of the Short-form McGill Pain Questionnaire (SF-MPQ-2). Pain. 2009;144(1-2):35-42.

[38] Krause SJ and Backonja MM. Development of a Neuropathic Pain Questionnaire. Clin J Pain 2003; 19(5):306-14.

[39] Galer BS, Jensen MP. Development and preliminary validation of a pain measure specific to neuropathic pain: the Neuropathic Pain Scale. Neurology 1987; 48(2):332-8.

[40] Collins S, van Hilten JJ, Marinus J, Zuurmond WW, de Lange JJ, Perez RS. Development of a symptoms questionnaire for complex regional pain syndrome and potentially related illnesses: the Trauma Related Neuronal Dysfunction Symptoms Inventory. Arch Phys Med Rehabil 2008; 89:1114-20.

[41] Streiner DL and Norman GR. Health Measurement Scales: a practical guide to their development and use (3rd ed). New York: Oxford University Press; 2003.

[42] Henson RK. Understanding internal consistency reliability estimates: a conceptual primer on coefficient alpha. Meas Eval Couns Devel 2001; 34:177-89.

[43] MacDermid JC, Grewal R, Macintyre NJ. Using an evidence-based approach to measure outcomes in clinical practice. Hand Clin 2009; 25:97–111.

[44] Guyatt G, Walter S, Norman G. Measuring change over time: Assessing the usefulness of evaluative instruments. J Chron Dis 1987; 40:171–8.

[45] Bruehl S, Harden RN, Galer BS, Saltz S, Bertram M, Backonja M, Gayles R, Rudin N, Bhugra MK, Stanton-Hicks M. External validation of IASP diagnostic criteria for Complex Regional Pain Syndrome and proposed research diagnostic criteria. Pain 1999; 81(1-2):147–54.

[46] Brunner F, Lienhardt SB, Kissling RO, Bachmann LB, Weber U. Diagnostic criteria and follow-up parameters in complex regional pain syndrome type I - a Dephi survey. Eur J Pain 2008;12(1):48-52.

APPENDIX A Scoring sheet and scoring guidelines for critical appraisals

Critical Appraisal Of Study Quality For Psychometric Articles

Reference:

		Descriptors
Study question		
	Score	
1		Research question
Study design		
2		Setting and Participants
3		Hypotheses and types of reliability and validity
4		Scope of psychometric properties
5		Sample size
6		Recruitment and retention
	Measurements	
7		Measurement procedures
8		Standardization
	Analyses	
9		Relation to hypotheses

10		Appropriateness of statistical tests
11		Benchmarks and CIs
Recommendations		
12		Conclusions and clinical recommendations

Critical Appraisal Of Study Quality For Psychometric Articles Interpretation Guide

To decide which score to provide for each item on your quality checklist, read the following descriptors. Pick the descriptor that sounds most like the study you were evaluating with respect to a given item.

		Descriptors
Study question		
Score		
1	2	The authors: - performed a thorough literature review indicating what is currently known about the psychometric properties of the instruments or tests under study - presented a critical, and unbiased view of the current state of knowledge - indicated how the current research question evolves from a current knowledge base - Established a research question based on the above.
	1	All of these above criteria were not fulfilled, but a clear rationale was provided for the research question
	0	A foundation for the current research question was not clear or was not founded on previous literature
Study design		
2	2	Specific inclusion/exclusion criteria for the study were defined, the practice setting was described and appropriate demographic information was presented yielding a study group generalizable to a clinical situation.
	1	Some information on person and place is provided (NOT ALL). For example, age/sex/diagnosis and the name of the practice (clinic name) without additional information. Information on the type of patients is briefly defined, but it is insufficient to allow the reader to generalize the study to a specific population
	0	No information on type of clinical settings or study participants is provided.
3	2	Authors identified specific hypotheses which included the specific type of reliability (intra/inter-rater or test-retest) or validity (construct/ criterion/ content; longitudinal/concurrent; convergent/divergent) being tested. For validity, expected relationships or constructs were defined.
	1	Types of reliability and validity being tested were stated, but not clearly defined in terms of specific hypotheses.
	0	Specific types of reliability or validity under evaluation were not clearly defined nor were specific hypotheses on reliability and validity stated. ("*The purpose of this study was to investigate the reliability and validity of*…" can be rated it is zero if no further detail on the types of reliability and validity or the nature of specific hypotheses is stated)

4	2	An appropriate scope of psychometric properties would be indicated by 1. A detailed focus on reliability that included multiple forms of reliability (at least two of – intra-rater, inter-rater, test-retest) where both relative and absolute reliability were addressed. (e.g. ICCs and SEM/MID) 2. A detailed focus on validity that included multiple forms of validity (content- judgmental; structured e.g. expert review/survey or qualitative interviews) or statistical (e.g. factor analyses), construct (known group differences; convergent/divergent associations), criterion (concurrent/predictive), responsiveness; predictive, evaluative or discriminative properties were established 3. Some aspects of both reliability and validity were examined concurrently using multiple approaches/analyses.
	1	Two psychometric properties were evaluated, however, the scope of both was superficial or narrow (e.g. point estimates used for one type of reliability and only a single unidimensional validity hypotheses tested)
	0	The scope of psychometric properties was very narrow as indicated by only one form of reliability or validity hypothesis estimated/tested.
5	2	Authors performed a sample size calculation and obtained their recruitment targets. Post-doc power analyses and/or confidence intervals confirm that the sample size was sufficient to define relatively precise estimates of reliability or validity.
	1	The authors provide a rationale for the number of subjects included in the study, but did not present specific sample size calculations or post-doc power analyses.
	0	Size of the sample was not rationalized or is clearly underpowered.
6	2	90% or more of the patients enrolled for study were re-evaluated.
	1	More than 70% of the eligible patients were re-evaluated.
	0	Less than 70% of the patients eligible for study were re-evaluated
Measurements		
7	2	The authors provided or referenced a published manual/article that outlines specific procedures for administration, scoring (including scoring algorithms handling of missing data) and interpretation that included any necessary information about positioning/active participation of the client, any special equipment required, calibration of equipment if necessary, training required, cost, examiner procedures/actions. Text describes key details of procedures.
	1	Procedures are referenced without any details or a limited description of procedures is included within text.
	0	Minimal description of procedures without appropriate references

8	2	All of the measurement techniques, including administration and scoring of the measurements were performed in a standardized way. This would include calibration of any equipment; use of consistent measurement tools and scoring, a priori exclusion of any participants likely to give invalid results/unable to complete testing (no exclusion of after enrollment participants); use of standardized procedures
	1	No obvious sources of bias, but minimal attention or description to ascertain the extent to which the above standards were maintained.
	0	No description of the extent to which the above standards were maintained or an obvious source of bias in data collection methods
Analyses		
9	2	Authors clearly defined which specific analyses were conducted for the stated specific hypotheses of the study. This may be accomplished through organization of the results under specific subheadings or by demarcating which analyses addressed specific psychometric properties. Data was presented for each hypothesis.
	1	Data was presented for each hypothesis, but authors did not clearly link analyses to hypotheses.
	0	Data was not presented for each hypothesis or psychometric property outlined in the purposes or methods
10	2	Appropriate statistical tests were conducted: 1. Reliability (e.g. Relative=ICCs for quantitative, Kappa for nominal data); absolute (SEM)) 2. Clinical relevance – e.g. minimal detectable change, minimally important difference, number needed to treat 3. Validity a. Validity associations- e.g. Pearson correlations for normally distributed data, Spearman rank correlations for ordinal data; or other correlations if appropriate b. Validity tests of significant difference- e.g. an appropriate global test like analysis of variance was used where indicated, with post-hoc tests that adjusted for multiple testing 4. Responsiveness- e.g. standardized response means or effect sizes or other recognized responsiveness indices were used.
	1	Appropriate statistical tests were used in some instances but suboptimal choices were made in other analyses.
	0	Inappropriate use of statistical tests

11	2	For key indicators like reliability coefficients indices at least 2 of the following were presented 1.appropriate confidence intervals, 2. Comparison to appropriate benchmarks or standards or 3. SEM. Correlation matrices for validity analysis may not require that each individual correlation be presented with its associated confidence intervals; however, however confidence intervals and benchmarks should be used according to standards for that type of analysis.
	1	Either confidence intervals or appropriate benchmarks were used-not both
	0	Inappropriate use of benchmarks or confidence intervals or neither included
Recommendations		
12	2	Authors made specific conclusions and clinical recommendations that were clearly related to specific hypotheses stated at the beginning of the study and supported by the data presented.
	1	Authors made conclusions and clinical recommendations that were general but basically supported by the study data; OR authors made conclusions and clinical recommendations for only some of the study hypotheses
	0	Authors made vague conclusions without any clinical recommendations; conclusions or recommendations were in contradiction to the actual data presented

CHAPTER TWO

Do you see what I see? A cognitive debriefing study of the CB-HI-CRPS.

Abstract

Objectives: The Hamilton Inventory for Complex Regional Pain Syndrome (HI-CRPS) is a multidisciplinary condition-specific assessment tool under development, including both patient-report (PR) and clinician-based (CB) assessment components. This study was intended to inform the content and structure of the CB-HI-CRPS, as well as subsequent training and user manuals for target users.**Methods**: Semi-structured cognitive interviews were conducted with a purposive sample of 20 health care professionals (HCP) from a spectrum of different disciplines working with CRPS. Assessment practices and scaling preferences for 16 concepts related to CRPS assessment were collected; these concepts relate directly to individual items on the CB-HI-CRPS. Interview recordings were transcribed and coded with emergent themes and preference and practice data was compiled.

Results: Four overarching themes emerged from the interviews: assessment beliefs and values, beliefs about CRPS, professional roles and multidisciplinary functions, and knowledge translation. Participants reported using the concepts represented by the CB-HI-CRPS 85.2% of the time. Physicians and nurses generally preferred present/absent judgements, while therapists used none/mild/moderate/severe scaling.

Discussion: In order for us to see our patients in the same way we must start to come to some consensus on what we look for (assess) and how we interpret those findings. The lack of uniformity in terminology and assessment behaviours underscores the need for the user manual of the CB-HI-CRPS to have a clear scoring framework and standardized assessment instructions to improve reliability across a multidisciplinary user group.

Keywords: complex regional pain syndrome, outcome assessment, cognitive debriefing

INTRODUCTION

Complex regional pain syndrome (CRPS) is a perplexing neurological condition that commonly arises following a traumatic injury, and can be associated with a peripheral nerve injury.[1,2,3] De Mos et al[4] calculated the incidence of CRPS at 26.2 per 100,000 person years in a retrospective cohort in the Netherlands; CRPS is thought to affect up to 30-40% of patients following upper extremity injuries or surgeries, and becomes a chronic condition in just under 2% of these patients.[5,6]

The presentation of the syndrome is variable; usually affecting a single limb, but occasionally seen to spread bilaterally or hemispherically. [1] The symptoms are variable: while most patients report some form of burning pain,

they may also have swelling, circulatory changes, skin changes, sensory complaints, stiffness and altered movement patterns. [7, 8] CRPS is often divided into two subtypes (CRPSI and CRPSII), distinguished by the presence of a known nerve injury in CRPSII.[2]

Although consensus-based diagnostic criteria[1] and assessment recommendations exist,[9] there is as yet no gold standard for diagnosis.[2,3,10] Despite this, there have been attempts to quantify some of the symptoms associated with CRPS.[11,12] For the most part these have focused on specific symptoms and have limited validation; no comprehensive CRPS scale has been accepted into practice or research. Preliminary work has been undertaken to develop a condition specific outcome measure, the Hamilton Inventory for Complex Regional Pain Syndrome (HI-CRPS).

Previous Development of the HI-CRPS

Development of the HI-CRPS was undertaken with the goal of providing clinicians and patients with a condition-specific tool for outcome assessment of both CRPSI and CRPSII affecting any limb. A foundational literature review found 96 different descriptors or evaluation points: these concepts were formulated into signs for clinician measurement (20 items) and symptoms for patient self-report (45 questions) and reviewed by patients and experts for clarity, face and content validity. This was combined with item-total correlation data from a small pilot study to remove redundant items (T. Packham, unpublished data, 2007), resulting in the current prototype [16 multidisciplinary items for

clinician based (CB) assessment, and a 35-item patient self-report (PR)]. This study focuses on the clinician-based (CB) portion of the assessment tool (the CB-HI-CRPS).

Cognitive Interviewing

Cognitive interviewing is a qualitative method that can be used to examine how participants interpret and respond to survey questions.[13,14] The debriefing data can help survey developers to not only discover errors made by respondents from the target population, but also to identify where those errors arise in the response process, thus facilitating item revision and the development of new items for self-report (SR) assessments.[15,16] Traditionally, qualitative studies seeking to refine clinician based (CB) assessment tools have used a Delphi or other consensus method for validation of the content coverage,[17,9] but these do not generally allow for an understanding of the sources of difference. This mixed methods study employed a novel application of cognitive debriefing, where those interviewed were the potential future users of a clinician-based assessment tool (the HI-CRPS), rather than the patients completing a self-report questionnaire.

Improving observational and direct measures: do you see what I see?

Reliability can be described as a measure of the reproducibility, agreement, or degree of error inherent in any measurement or measurement tool.[18] Inherent in classical measurement theory is the principle that every clinical measure is composed of the true score and a component of measurement error. Error can arise during clinician-based assessment because of a lack of clarity or consistency in how a measure is conceptualized/interpreted or applied[19]: others may not see what we see simply because they are looking elsewhere. For this reason, understanding the cognitive underpinnings of how clinicians interpret measures is a potential basis for enhancing reliability. Others have focused on how reliability of an assessment tool can be improved by reducing the amount of errors through standardization, training of assessors, and using precise rating scales.[18, 20] Few studies have used qualitative approaches to understanding measurement error.

Validity focuses on the extent that a tool can measure what is purported to measure and thus encompasses multi-faceted analyses[21], including *face validity* (the overt ways that the tool reflects concepts that the reader or user feels are important) and *content validity* (the degree to which the instrument includes the spectrum of assessment concepts). Measurement tools can be used to describe, predict, and evaluate the concepts of interest.[22] It is important for instruments to clearly define which of these purposes are being undertaken when

a tool is developed; and to test whether it can perform these different measurement functions.

This study examined the assessment practices, beliefs and preferences of the intended users of the CB-HI-CRPS with the intent of improving the reliability, face validity and content validity for future testing. The specific areas of exploration were intended to define:

1) user-generated definitions and descriptions for assessment techniques and scale anchors,
2) similarities and differences in how health care professionals from different disciplines assess and formulate judgements on the clinical signs of complex regional pain syndrome, and
3) implications for altering the measurement properties of a CRPS prototype tool with the intent of reducing possible sources of user error, and ultimately potentially improving reliability, content and face validity.

MATERIALS AND METHODS

Present Study Overview

This study consisted of a series of semi-structured cognitive debriefing interviews conducted by a team of researchers from McMaster University. Interview content was intended to reflect the key assessment concepts included in the clinician portion of the Hamilton Inventory for Complex Regional Pain Syndrome (CB-HI-CRPS); see Table 1 for a listing of the concepts. Participants

were asked to provide definitions for the concepts covered by the assessment (for example, allodynia and guarding) as well as to provide detailed description for the end anchors of the scales, or for individual scale points. Opportunity was also given for participants to state their preferences for different formats of scales that could be used for ranking the individual constructs. Although a working draft of the CB-HI-CRPS tool was provided to participants for reference, the focus was content rather than the format. While cognitive debriefing typically is used for the refinement of patient self-report measures,[15,16,14] this study undertook identification of potential sources of consensus and difference in clinical

Table 1: Key assessment concepts of the CB-HI-CRPS

Allodynia	*Hyperpathia – pinprick*
Cold intolerance	*Guarding*
Skin temperature	*Mottling*
Hyperhydrosis	*Edema*
Hair growth	*Skin quality*
Nail quality	*Muscle Tone*
Incoordination	*Movement given time since initial injury*
Movement given severity of initial injury	

`

assessment strategies across the professional groups, and to utilize this information for tool refinements that should strengthen statistical reliability and correspondingly, validity.

As part of the cognitive debriefing interview, participants were asked to describe their current assessment practices. The descriptive methodology of contemporary content analysis[23] was then used to garner insight into how individual clinicians have addressed the clinical challenge of CRPS assessment.

Participants and Research Team

Professionals with an interest in or experience with the treatment of complex regional pain syndrome were purposively selected to reflect the anticipated future users of the CB-HI-CRPS from six different professions: anaesthesia, occupational therapy, orthopaedic surgery, plastic surgery, physiatry, and physiotherapy. The goal of including multiple professions was to gain perspective on the scope of professional variations[24] and ensure that the language of the tool was clear and inclusive for all disciplines.

The research team for this project consisted of an occupational therapist (TP: author of the HI-CRPS) and two occupational therapy students from McMaster University. One or two members of the team conducted each interview after a joint initial practice interview. The sessions were audio-recorded, then transcribed verbatim for analysis. Reflecting the mixed methods design, the analyses had two forms: 1) quantitative content analysis of demographics, assessment practices and scaling preferences using an item by item format, and 2) descriptive thematic analysis and coding of the transcripts. The research team met for coding and consensus sessions four times over a one-year period.

This study received approval from the joint Research Ethics Board of McMaster University and Hamilton Health Sciences in Hamilton, Ontario.

Procedures

Sampling

The intended sampling strategy was to include at least 3 representatives of each of the target professional groups. Invitations to participate in the study were sent to members of a multidisciplinary clinical and academic network (MacHANd: McMaster Hand, Arm and Nerve), and therapist interest groups (Hamilton Hand Interest Group and London Hand Interest Group). Interviews were conducted at a place of the interviewee's choosing: ranging from offices to clinics to cafes. In the one-year study period, interviews were conducted with 20 clinicians; however, 2 interviews were lost to technology failures (degradation of audio-cassettes), leaving 18 transcripts for review. Unfortunately, despite multiple requests only a single orthopaedic surgeon responded and provided consent to participate; subsequently this interview consent was withdrawn because the surgeon felt they could not contribute the amount of time required for the interview (45 min-1 hr). Using a snowballing technique, those who participated were given the opportunity to recommend other professionals whom they knew to be interested in the area of CRPS; these individuals were then also invited to participate in an interview. While nursing had not originally been included as a target group, other professionals recommended nursing colleagues as part of the "snowballing", and these RNs were accordingly incorporated into

the sampling frame. No volunteers who gave consent to participate in the full interview process were excluded.

We were unable to meet initial multidisciplinary recruitment targets because of the lost data and imbalanced recruitment: only 2 transcripts were available from each of the plastics, physiatry and physiotherapy professions (see Table 2 for participant demographics). However, six occupational therapists (Ots) volunteered for the study and were interviewed. Five of the six OT participants each represented a distinct geographic or practice area, and were included by the research team because of the breadth they contributed to the overall study. The 2 Ots included from the same facility represented very different levels of experience and training: again, the team felt this reflected the spectrum of future users of the CB-HI-CRPS.

Table 2. Description of participants

	Mean	SD	Range
Years experience	14.9	9.9	3-35
Percentage of practice in CRPS	10.6	7.0	2-20
Self-rated expertise in CRPS (0-10)	7.5	1.9	5-10
Self-rated expertise in assessment skills (0-10)	8.1	1.7	5-10
Profession	Anaesthesia = 3, OT = 6, PMR = 3, PT = 2, Plastics = 3, RN = 3		
Practice area	Paediatrics = 2 , Pain = 4 , Hands = 8 , General = 6		

Several potential biases are acknowledged: 1) a high number of participants specialized in upper extremity practice, and 2) the interviews were limited geographically within the area of Southern Ontario for ease of conducting face-to-face interviews. It is also important to note that the study was closed after a one-year period when leads on new participants had been seemingly exhausted, rather than because data saturation or informational redundancy[9] had been achieved. New themes continued to emerge throughout the interview and review process, and this study cannot be interpreted as an exhaustive review of the rich scope of ideas generated from this examination.

Data collection

After giving informed consent, participants underwent a 'verbal probing' format of cognitive interview. Using a combination of established questions and responsive probing, subjects were asked to describe their assessment practices, define concepts, and indicate scale preferences related to each item (refer to Table 3 for sample questions). The face-to –face interviews, typically about an hour long, were audio-recorded then transcribed for analysis. Order bias can arise when respondents modify their responses based on 1) how they have previously answered other questions, 2) fatigue or 3) boredom [25,26], so the fifteen items for discussion were presented in a random order to each participant. After the first three interviews were completed and transcribed, the team reviewed the transcripts individually and jointly to generate themes and to modify

Table 3. Sample questions from cognitive interview structure

Guarding behaviour
What do you think this question [from CB-HI-CRPS] is asking you to assess?
Which scale would you prefer to categorize guarding behaviour in your patients? Why?
What cues would you use to pick a response category?
How confident are you in making a judgement about this characteristic?
Hypersensitivity – allodynia
How do you define allodynia?
How do you assess allodynia in your patients?
Which scale would you prefer to assess allodynia in your patients? Why?
How would you characterize the anchors of the scale [*mild allodynia, severe allodynia*, *etc*]?
Can you describe a patient to me that you think would fit this category?
How hard is it to judge allodynia?

the interview format to reflect emergent concerns. More themes were added as the interviews continued, with subsequent recoding of previously completed interview transcripts. Each of the joint sessions included time for reflection and discussion of how each individual member of the research team guided the interviews with their questions, and responded to the opinions and emotions of the participants.

Several issues that emerged from the reflection sections were team perceptions around experience and relationships. Student members of the research team felt their limited clinical experience constrained the interviews, and they tended to follow the question framework closely. However, as the study progressed and they became more familiar with the subject matter, they became more confident in pursuing alternate lines of questioning when participants'

responses indicated an opportunity to explore an alternate area. However, the interviews conducted by the students were generally richer in descriptive detail as participants seemed to assume a teaching role. Conversely, the clinical researcher (TP) was sometimes seen as an expert, and participants would ask questions and seek opinions on the basis of the perceived expertise. These interviews tended to be more broad-ranging in scope, and raised global issues such as: 1) timing and patterns of referrals within the health care system, 2) shifting theoretical frameworks for treatment, and 3) issues related to research, including development and use of outcome measures.

Analysis Plan

The coding system was developed after each reviewer listened to the first 3 interviews to generate some overall categories; more categories were added as they emerged. Working independently to reduce bias[27] each member of the research team identified categories within each anonymous transcript using a colour coding system; this coding was then reviewed by the team and discussed until consensus was reached. Using a cross-case analytical approach,[28] data from the interviews was then transcribed into Microsoft Excel in an item-by-item format, with respondents identified only by discipline. This second coding focused more on content-specific analysis that relates to the individual items and measurement concepts, while the initial thematic analysis highlighted the similarities and differences among the disciplines related to the process of making clinical judgements. Contemplation on each individual category

generated the sub-categories as the team attempted to reflect the scope of each meaning within each grouping. Subsequent reflections and discussions by the research team then contracted the categories into the four overarching themes. This level of analysis is more in-depth than the tradition of cognitive debriefing[15, 16] yet was in keeping with the overall study plan of not only refining the CB-HI-CRPS tool, but understanding the collective wisdom of health care professionals relating to outcome measurement for this condition.

RESULTS

DESCRIPTIVE CONTENT ANALYSIS FINDINGS

Although we do not presume to have exhausted the potential of themes arising from cognitive debriefing on this topic, four overarching themes emerged from these interviews:

1. Assessment beliefs and values
2. Beliefs about CRPS
3. Professional roles and multidisciplinary functions
4. Knowledge translation

The four themes arose from 10 content categories; these are presented with related qualifiers, and quotes reflecting their essence are listed in Table 4. The remainder of this section will define each theme and expand on each facet of the categories.

Table 4. Qualitative themes identified as arising from interviews

Themes		
Categories	***Qualifiers***	***Illustrative Quotes***
Assessment beliefs and practices		
The role of experience in assessment	*Necessary for competence*	"I have seen beads of sweat on their hand and they are doing nothing – they are sitting! … It's rare- but it does happen. But I've been in practice for 27 years, and I don't know how many would have seen that."
	Key factor for confidence in assessment skills	"I work with CRPS patients but not that many. I think this [confidence] would come with more experience. I think this [guarding behaviour] is sometimes hard to discriminate this between the increased muscle tone'"
	Assessment approaches evolve with skills and experience	*"*Again, it (Incoordination) is something that I am looking at more closely now, but I would probably rank myself quite low…. *[Interviewer asked: So you feel like as you do it more you will develop proficiency*?] Yes*."*
	Contributed to precise definitions for scale anchors	"The people who I would put in the severe category… is when you start noticing, on the dorsum of the hand below the nail, that they start to get a little bit sweaty there, that is very unusual, you won't see that is people without ncoordinatio… ."
The role of comparing to unaffected limb	*Integral to the assessment process*	"I'll note it that there is something different, that it is a true difference within that individual compared to the other side" *[in reference to nail quality]* "I look for it *[ncoordinatio]* and then I feel. So if they're like 'this palm is always sweaty', then I feel the two palms."
The need for observation, patient report and direct measurement	*Varies at different points in the assessment process*	"…usually when you first start with somebody, it's just going to be through questioning 'do you have pain', 'where is it and how much is it', 'does it hurt when you touch things', before I even do any hands on assessment" *.*
	Correlating patient reports and direct measurements	*"W*hen I am doing ROM you can almost feel the hand get cold and clammy a lot of the time or change in terms of temperature… then I will ask them as well, how does it feel? Does it feel cold, does it hot, to make sure that we correlate."
	Finding a balance between subjective and objective measures	"…less subjective from the patient and more objective from us would be a huge benefit" *"…*I think it is ok to use the descriptive factors, but I need some kind of objective measurement to back up my observations."
	Time pressures influence assessment behaviours	*"…*that's a good one because there are lots of ways to measure it [edema], but how often to we take the time to actually use them – that's another matter."
	The perceived value of direct measurements	"If you are going to do ROM, then you have got to measure it. If you are not going to measure it, then you have to have some standardized functional task…."

Professional roles and functions		
Diagnostic evaluation vs. outcome evaluation	*Both perspectives are useful*	"Unfortunately with CRPS, and sometimes it's *[hair growth]* very subtle, it doesn't mean that it's mild or mild disease. It just means there are mild changes …usually by the time they're having changes that are more severe, they're also having a huge decrease in function. Whereas my approach to CRPS is more if I think it's present, I treat it aggressively. And all I'm looking for is evidence of CRPS." "The present or absent is simpler, and maybe better for diagnosing, but in terms of looking at change over time…it depends what you want the tool for."
	Certain features are more useful for diagnosis than measuring outcome	*"I think, that at that point, present and absent would be sufficient information in the clinical needs to really come to any conclusions about CRPS.* [in regards to mottling]" *"I tend to use the allodynia piece as information to help me figure out if they do have CRPS…it's one of those real trigger things. But do I use it as a tool to see whether or not they're getting better…? I don't know, I don't think I would re-evaluate that in any way."*
	Utility for ruling in and ruling out	But if it is present, it is a diagnostic sign…. But if it isn't [present], it does not rule out CRPS.(in reference to skin temp)
	Focus reflected in scaling preferences	*"… you either have it* [allodynia] *or you don't. If you have it – mild, moderate, severe, it doesn't matter – it is the presence of it that matters*." "I would use the 4 pt scale and I'll tell you why. Because when that patient comes back or if the CRPS is resolving, I want to be able to look at my descriptors and say it was moderate, and now is maybe mild or none…." *[in reference to ▯ incoordination]*
Health care professional (HCP) roles and scope of practice	*Specialization supports skill for CRPS assessment*	`` *I think that question* [movement given severity of injury] *would be answered slightly differently, based on what their frame of reference was. A generalist may not know enough about what to expect for a particular injury."* *""I think in a pain clinic – for sure; but I think your average therapist in general practice…I think they might see the pathological ▯ incoordination in CRPS, but in terms of [judging] how extreme it can go, I don't know."*
	Assessment partnerships within the multidisciplinary team	*"We [RNs] don't do that. Our physiotherapists do it as part of their assessment and treatment plan…"* *"I would rely on the notes from the hand therapist to give me that information* [related to functional assessment]." "That *[allodynia]* would make me think, oh I really have to talk to the doctor about this person."

Professional roles and functions (con't)		
	Practice environment influences roles and scope of practice	"[Interviewer] How hard is it to judge allodynia? It depends on the age of the patient. So in paediatrics, it can be pretty tough. We have some kids with cognitive impairments, and other kids are very young and just not able to verbalize… some of them are just scared that they are even there, that someone is going to touch them, so it may be anxiety related rather than true hyperalgesia.`` "…I am dealing with patients in their more acute phase, and passing them along to pain management ncoord … so most of the patients I see will not be in the severe stage…."
	Roles influence scope of practice, and assessment practices and opportunities	*"… we* [RNs] *don't clinically do that* [assess Incoordination]*. To be quite honest, I don't know that I would have good cues to know how to assess that."* "But it is difficult to tease out the separate components when you are treating someone as a therapist… you are doing it all at once." "[Interviewer] How confident are you in making a judgement about this characteristic [allodynia]? *OK, that's within our scope."*
Beliefs about CRPS		
Physical vs. psychological symptoms	*Reflective of attitudes towards signs vs. symptoms*	*"Well, you are always relying on the patients' perceptions, so if the patient wants to fool you, they can fool you. It's pain, so it is subjective…".* "…you think they may be malingering, but if you step back and think "wait a minute, they haven't actually used this arm normally for a while, and it's *[coordination]* not just going to come back like that *[snapped fingers]."*
	Relationship to impaired perceptions and/or altered body image	*"Kids come in with slight swelling, and they come in saying "Look at how bad my swelling is" and you are looking at it saying "Riiight!" because you can barely see it…* [Interviewer asked: Do you attribute that to the child just being really focused or perseverative on it, or do you attribute it to changes in the sensory map in the brain, and body perception in CRPS?] …I think it is both: a perceptual component, and I believe there is a huge psychological component to chronic pain. Huge. There has got to be. I mean they are hypervigilant, it's almost like an anorexic with their weight…. *"… they'll have the sensation of swelling and I try and distinguish that and actual swelling…I'll say 'You might have the sensation of swelling, but is that swelling actually present? Look at the other hand and see if you can see any changes'"*

Beliefs about CRPS (con't)		
The influence of pain on categorical judgements	*Avoiding assessments that might be painful*	*"I would use the FROM to the distal palmar crease, because I am not going to measure each joint one by one, it's ridiculous with someone in pain."* *"I assess it* [edema] *by godette manoeuvre. You apply pressure, and if the fingernail stays printed in the skin of the patient, that is a positive sign for edema. But in the majority of these patients doesn't* [sic] *allow you to apply this kind of pressure in the region. So this is kind of difficult…"* "*I find it too traumatizing for the patients to pull out a pin (to assess hyperpathia) when they're having allodynia."*
	Rapport valued over complete assessment	*"I don't see the point of subjecting them to a stimulus that I know is noxious, and I know they are hypersensitive… I don't think it benefits the theraputic relationship that I have with them….It's the same thing as if they have told you their bone is broken – why are you doing passive ROM when you know it is counter-productive?"*
Severity vs. chronicity	*Some changes only present after a long time*	"[*Interviewer]* Do you see decreased hair growth? *Yes – later stage."* *"[Interviewer]* Do you look at nail quality or changes in the nail growth? *Yah- again, to us, those are late signs*." "Shiny waxy hands in more advanced CRPS, but shiny waxy feet more – you see, it's more that dependent circulation problem too. Do you see the changes in the feet earlier than the hand? *Yeah, maybe the vascular symptoms in the lower extremity are more exaggerated because of the dependence."*
	Signs found in every stage of syndrome, but may change with severity	"…Well, I have seen incoordinatio in every stage of CRPS" "… it depends on how long you've had it, because there would be more guarding behaviours and more contractures the longer you've had it."
	Different assessment practices for different phases	*"[Interviewer]* Do you see altered hair growth? *Rarely. We try to get them in quick – we try to see potential new CRPS patients as soon as possible, so it is pretty rare that we would see changed hair growth."* *"I think probably 0-3 would be the best I could do in terms of assessment. You have got to remember those nail changes happen over a long period of time, and I am dealing with patients in their more acute phase…*. " "… if they are in the later stages, and you are not really going to do anything for them, then what is the point of assessing them, other than to document that they can't go beyond a certain range, like a joint contracture?"

Beliefs about CRPS (con't)		
	Factoring in deconditioning to judgements about movement in the later phases	"I don't expect to see it *[Incoordination]* in CRPS, unless it is in the limbs, in the lower limbs. Because the weakness in one leg is going to make you have a wobbly type of walk and it's not going to be that coordinated, because you have a deconditioned muscle." "The difference is, how long it has been since the injury, is in the conditioning…. So I don't think that much on how much they can do – I think whether they can or cannot do it…if they cannot walk a certain distance, because the leg gets tired, it could just be part of de-conditioning…."
Symptom variability as an assessment challenge	*Hard to make judgements based on a single time point*	"…they will say that 'When I wake up in the morning it's mottled, or on and off it's mottled all day', but of course the hour that you see them, nothing is present…." *"And just assess it throughout treatment, see what they present with then they get there, see what they present with partially through, monitor it throughout treatment and see what it looks like at the end… and let me know what happens after they leave therapy."*
	Underscores need for listening to the patient	"I've had patients who say 'normally the hand is really hot, but it's really cold in your office so it's not too bad today, it actually feels normal'. I would say that in my dictation, they're having a day where it's not as bad, normally it feels much warmer than the other limb."
Knowledge Translation		
	Practice evolves as new research is incorporated	"In the beginning I did used to do it *[measure skin temperature]* because I thought that it had to be cold, but the more I read about it, it really doesn't matter, because the cold can be presented late, late in the problem – in the beginning it can be warm."
	Clinicians unsure if their practices are up-to-date	*"I don't know why we don't* [assess pinprick hyperpathia]*, but we don't. I wonder if it's paediatrics, I wonder if they are already so anxious and so fearful, that if we started putting painful stimuli….* *[Interviewer] In actual fact, we are likely going to drop this item, because most clinicians told us they would never do this. When we sent the tool out to experts, they thought it was important, but most of them were researchers, and doing it in that context. But clinicians told us they would not test this. Thank God, I was thinking, shoot, are we supposed to be doing that?"*
	Clinicians expressing lack of certainty around knowledge and practices	"…lack of hair….so why is that? This is for my own education. So people can have lack of hair with the sympathetic response as well?" "[Interviewer] Do you see this [increased hair growth] in children? *No. Do you?* *[Interviewer] We see it in adults frequently.* ***Really? Nope, not in kids."***

Knowledge Translation (con't)		
	KT barriers and supports	*"…maybe I should do the figure-8 method* [to measure edema] *… but it takes time to change your practice and so when you're already in the routine of doing something a certain way, it's hard to change."* **"I don't have a huge strong preference but what I chart like is more typically like B [a 4 point scale]. I know that the likert scale is the 7 point scale, and probably has the best research to support it."**
	Using interviews to reflect on evidence-based practice	*"This could be done differently now that it occurred to me."* *"And after I was reading it over again, I was thinking, oh maybe I should be assessing that...."* "…when I'm testing for sensation I'm not looking necessarily for allodynia, I'm looking for their ability to sense what I'm trying to...Maybe that's wrong, I don't know." "Now you are making me think about it! In my mind if you use the term ncoordinatio, it is to the point that the person needs to seek treatment for it, to do something to the nerves or glands or whatever…. Yes I am comfortable *[making a judgement about ncoordinatio],* but I hesitate now just thinking about how I have used the term a bit differently."

Assessment beliefs and values

The role of experience in assessment

Many participants articulated that clinical experience was a general pre-requisite for competent assessment of this population. Others saw experience as a critical factor to their confidence in their own assessment skills. The evolution of skill mirroring the evolution of practice with experience was also reflected on. The research team noted that experience also appeared to contribute to advanced description for scale end anchors.

The importance of comparison to the unaffected limb

This topic emerged so frequently throughout the interviews, it appears to be a foundational assessment value. In fact, it was articulated by every participant and referred to in relationship to every concept on the CB-HI-CRPS.

The need for observation, patient report, and direct measurement

Clinicians often raised the idea of the importance of using different sources of assessment information, including objective measurements, clinical observations, and listening to the subjective reports of the patient being assessed. Value was attributed to all of these forms of information; however, participants also expressed a need to find a balance between the different sources of information, suggesting that all forms were not valued equally. This was reflected in the process of assessment by seeking to correlate subjective data with objective findings. The taxonomy used by the participants also appeared to reflect a hierarchical value structure, as "subjective" was used far more frequently than "patient-reported".

The influence of time pressures on assessment choices was also noted, as it appeared that some participants made assessment choices based on time constraints rather than best practices. Others expressed strong opinions about the value of making actual measurements of range of movement (ROM) as opposed to the judgements about movement relative to anticipated recovery proposed by the current draft of the CB-HI-CRPS.

Professional roles and functions

HCP roles and scope of practice

Participants were from a purposive multidisciplinary sample of health care professionals (HCPs), and encompassed physicians, nurses and therapists from a variety of practice roles, including pain management, paediatrics, and hand therapy or hand surgery (refer to Table 2). However, most clinicians in the sample considered themselves specialists having a moderate to high level of expertise in CRPS, and this was reflected in their comments. The practice environment was also cited as a key influence on the scope of practice and assessment considerations.

Interviewees shared how their professional roles sometimes delineated what and how they would assess, and their comfort level with the assessment components. They also talked about the collaborative interactions between members of the multidisciplinary team as it pertained to the assessment of CRPS, and expressed that they valued the team approach.

Diagnostic evaluation vs. outcome evaluation

Assessment of CRPS can have a diagnostic focus or can be used for measuring progress over time, and/or final outcomes. Participants reflected the value of all of these perspectives. However, certain features of CRPS were thought to be more useful as diagnostic evidence than for monitoring progress – refer to Table 5.

Scaling preferences also reflected the purpose of assessment, with dichotomous scaling (present/absent) being preferred for features that were perceived as

Table 5: Numbers of participants who labelled CB-HI-CRPS items as diagnostic or outcome

	DIAGNOSTIC ITEM	OUTCOME ITEM
Hyperhydrosis	*3*	*1*
Hair growth	*0*	*5*
Skin temperature	*4*	*1*
Allodynia	2	*2*
Nail quality	4	0
Skin quality	2	*0*
Mottling	*4*	0
Incoordination	*0*	2
Cold intolerance	*1*	*0*
Muscle Tone	*1*	*1*
Hyperpathia – pinprick	*4*	*0*
Guarding	*1*	*3*
Edema	*2*	*7*
Movement given time since initial injury	*2*	*0*
Movement given severity of initial injury	*1*	*2*

diagnostic, and more descriptive scales (4 or 7 point) being preferred for monitoring change over time. Practice roles appeared to drive preferences: both occupational therapists and physiotherapists preferred to look at progress and outcome evaluation, while physicians' preferences reflected a diagnostic focus.

Beliefs about CRPS

Physical vs. psychological symptoms

Clinicians' descriptions of how patients recounted their symptoms sometimes reflected differing attitudes towards items seen as measurable clinical signs as compared to symptoms subject to the perceptions of the patient. Patient descriptions of pain intensity, guarding and neglect, incoordination, and fluctuating swelling were thought to be difficult to correlate or reconcile with behavioural observations and measurements such as ROM and volumetry. Several participants also related psychological symptoms to impaired perception or altered body schema, reflecting the evolving understanding in the literature[29-31] of the relationships between sensation, perception and altered cortical representations in CRPS.

The influence of pain on categorical judgements

Another subject that emerged was how the presence of pain might influence both the assessment process and the ranking of other symptoms. Clinicians sometimes avoided testing situations that they thought would induce or increase pain in persons with CRPS. In fact, the majority of our participants said they would not test pinprick hyperpathia (see Table 6), citing concerns for rapport and the therapeutic alliance.

Severity vs. chronicity

Interviewees discussed the challenge of separating the severity of this neuropathic pain syndrome from the duration of symptoms experienced by the person. They felt that some changes were only present in patients who had been affected by CRPS for a greater time interval. Other signs were seen in

every phase of the syndrome, although the severity of the findings might increase over time. Differences in assessment findings pertaining to timing was also influenced by practice area, with some paediatric practitioners indicating unique patterns of signs in children, and other clinicians citing their acute practice focus as precluding seeing what they perceived as later signs of CRPS. The influence of deconditioning with prolonged CRPS was also raised as a consideration when making comparative judgments on the severity of impairments e.g. coordination or range of movement; given that these could be affected by the duration of time since the original injury.

Symptom variability as an assessment challenge

The inherent variability of the signs and symptoms of CRPS was shared by many participants as a challenge and a barrier to relying solely on a single time-point of assessment. Health professionals from all disciplines stressed that this variability underscored the need to listen to the report of the patient in addition to undertaking objective physical examination.**Knowledge translation**

Knowledge translation can be defined as group of methods to close the gap between new research findings and clinical practice.[32] In the health care arena, it includes aspects of awareness of evidence, willingness to change, and environmental / system supports in order for health professionals to alter assessment and treatment practices.[33] During our interviews, the clinicians

Table 6. Assessment practices and scaling preferences

Assessment Concepts	Scaling preferences				Do you currently assess? (% yes)
	Present / Absent	4 point (0 – 3)	7 point (0-6 or –3 to 3 Likert)	Written Description/ Other Scale	
Allodynia	20%	45%	15%	- / 5%	85%
Hyperpathia – pinprick	15%	5%	5%	- / 10%	35%
Cold intolerance	20%	15%	5%	10% / 5%	55%
Guarding	30%	50%	10%	_	90%
Skin temperature	20%	15%	_	5% / 45%	85%
Mottling	50%	40%	_	_	90%
Hyperhydrosis	45%	30%	10%	-	85%
Edema	10%	75%	_	_	85%
Hair growth	30%	25%	5%	10% / 15%	85%
Skin quality	35%	15%	_	35% / _	85%
Nail quality	40%	25%	5%	5% / 10%	85%
Muscle Tone	15%	15%	15%	_ / 15%	60%
Movement control	15%	20%	10%	20% / _	65%
Movement given severity of initial injury	10%	10%	40%	5% / 10%	75%
Movement given time since initial injury	15%	15%	35%	5% / 10%	80%

made reference to how their practice had evolved in step with new research and ideas. Some also expressed concern that they were not keeping up with contemporary evidence, occasionally asking the interviewer if what they were doing was correct. Many participants also used the cognitive debriefing interviews as an opportunity to muse about their own practices, sometimes questioning what they were doing and why they were doing it. They also reflected on the barriers to knowledge translation like lack of time, and established practice patterns and identified strategies that could support change such as working towards change as a team.

Additional considerations

Several themes are perhaps noteworthy by their absence. No mention was ever made of gender, and all participants used gender-neutral terms like person, client or patient when referring to this population. Additionally, the sole references which could be attributed to culture or ethnicity pertained to a) the difficulty in assessing mottling in persons with darker skin pigment, and b) the need to carefully compare to the unaffected side, factoring in skin and hair colour when assessing changes in hair growth. The role of culture in the assessment of pain, and the influence of culture on the expression of pain was not overtly addressed.

The findings of these interviews must be viewed in the context of the research team. One of the researchers (TP) is the developer of the HI-CRPS, and was involved in the analysis of input on her tool from her colleagues and

peers. Her interviews often took on a different flavour, as participants sometimes asked questions of the interviewer, seeking additional information or verification for some of their beliefs and practices. Issues raised in these interviews sometimes reflected a broader spectrum of issues, including referral patterns and health system concerns. The impact of this dynamic bears consideration when reviewing these findings.

The other members of the research team were students, and they sometimes perceived a power imbalance when asking questions of experienced clinicians. Conversely, the participants often gave richer answers to the students, appearing to undertake an educational role with the students by providing detailed descriptions and definitions. By contrast, the interviewer with clinical experience sometimes received briefer answers, as if participants assumed she already knew the information.

Responding to perceived time pressures was a challenge experienced by all of the interviewers. The research team would edit the interview structure to fit the available time period, and omit questions if repeated probes were failing to elicit the depth of information sought. This experience seemed to justify the *a priori* decision to randomize the order of the concepts covered by each interview, but it is unclear to what extent the veracity of the interviews was compromised by these omissions. Finally, the large proportion of participants from the occupational therapy profession must also be acknowledged. The potential exists that this may bias the evaluation towards the taxonomy and theoretical

frameworks of client-centred practice and occupational performance that underpin the profession.[34]

In summary, the cognitive debriefing interview format created an opportunity for our participants to reflect not only on the content of the CB-HI-CRPS, but also on a variety of practice challenges and concerns, beliefs and attitudes related to the assessment of complex regional pain syndrome. Categories like diagnostic evaluation vs. outcome evaluation, and the need for observation, patient report, and direct measurement reflected our own struggles with developing a clinical assessment tool with a trans-disciplinary focus (T. Packham, Canadian Society of Hand Therapists, 2009: Toronto, Ontario). Issues surrounding knowledge translation, although not an anticipated theme, also resonated with the research team. The themes, categories and content we have shared here are not an exhaustive exploration of these issues, but are intended to give a broader context and depth of meaning when entwined with the quantitative analysis that follows.

QUANTITATIVE ANALYSIS

This section is intended to share the numerically descriptive data collected from the interview process. It is important to note that while 20 health care professionals participated in this study, not all participants provided answers for every question, so the number of respondents varies from question to question. Additionally, two audio recordings were lost due to technical malfunctions prior to

transcription, but their demographic and preference data (recorded on paper) are included in the overall summaries.

Scaling Preferences

Participants were asked to identify which scale they would prefer to use for scoring each individual concept on the CB-HI-CRPS (see Table 1 for the list of concepts). Grouped responses are summarized in Table 6; the data illustrates how preferences changed according to concept. For example, the dichotomous present / absent scale was preferred by 50% of respondents for rating mottling, but 75% preferred a 4 point scale (none, mild, moderate and severe) to rate edema, and 35% indicated that they would rather just write a description of skin quality instead of making a scalar judgement. Seven point scales were rarely selected (receiving no endorsements for four of the concepts) with two exceptions: the movement scales scored with a 7 point likert agreement scale (from strongly disagree to strongly agree) were preferred by 35% (duration) and 40% (severity) of participants. However, when participants were grouped according to their self-reported roles with a) physicians and RNs [*diagnostic group*] and b) therapists (both occupational and physiotherapists) [*outcome group*], a general trend did emerge. The diagnostic group tended to endorse the present / absent scales, in keeping with the type of judgments required to utilize the International Association for the Study of Pain (IASP) diagnostic criteria for CRPS, and with their consultative practice patterns. Conversely, therapists in the outcome grouping chose 4- point scales that reflected their current behaviours for

assessment and monitoring change over time, and were in keeping with seeing persons with CRPS for repeated visits over a longer period of time.

Assessment Practices

One of the purposes of the study was to examine health care professionals' current assessment practices for CRPS, and to compare and contrast those with the assessment template provided by the CB-HI-CRPS. The most striking departure from CB-HI-CRPS related to the concept of pinprick hyperpathia. Only 35% of participants currently assess this routinely: furthermore, amongst those who did assess it, it was used selectively on a case-by-case basis rather than as a standard part of their initial test battery. Hyperpathia for cold (used by 55%) and muscle tone (used by 60%) were also less consistently assessed concepts. However, there was some confusion with respect to the terminology: muscle tone was chosen as a neutral term encompassing both elevated or lowered levels, but some participants associated the phrase "muscle tone" with spasticity (12.5%) while others interpreted it as muscle strength (12.5%) as opposed to the intended classic definition of the amount of background contraction of a muscle at rest (43.8%). This is reflected in the 15% of participants who endorsed "other scales" as the way they would assess this concept: standardized grading systems for manual muscle testing and for the assessment of spasticity were among the additional scales cited by the participants as their current practices.

Overall, the participants reported using the concepts represented by the CB-HI-CRPS 85.2% of the time (SD= 11.1). When analyzed by professional groups, physicians followed the concepts a little more closely (mean 89.5%), therapists were consistent with the overall scores at 85%, and RNs tended to leave off a few more items (they used the concepts only 75.6% of the time). This represented a statistically significant correlation between professional group and use of the CB-HI-CRPS concepts (p=0.002).

User manual definitions and recommendations for standardization

Within the interviews, participants were asked to define or describe each concept covered by the CB-HI-CRPS, first as a check to ensure they understood what the question was asking them to assess, and secondly as a way of generating definitions for the user manual that reflected a multi-disciplinary lexicon. Many participants were quick to define allodynia using formal terms that reflected the IASP definition, which clearly states the pain is generated from "*a stimulus that does not normally provoke pain*" (p. 18),[35] but several participants substituted the word hypersensitivity and made reference to an exaggerated reaction to a noxious stimulus. The interview questions did not overtly ask participants to define hyperpathia, but instead asked them how they would assess pinprick hyperpathia. 25% of participants did not answer the question as they did not include it in their assessment; 10% directly referred to testing protective sensation as sharp/dull, while another 15% did refer to their testing procedure as testing hyperpathia, but whose methods in doing so did not reflect

the concept accurately [i.e. using Semmes-Weinstein monofilaments, which are intended to measure threshold.[36] Only 30% described an appropriate testing procedure (it is perhaps noteworthy that all of the anaesthetists fell into this group). Clinicians were much more comfortable with the concept of cold hyperpathia: they equated it with cold intolerance. However, those who included it in their assessment were evenly split on whether to objectively assess or simply ask the patient to describe their perceptions.

The differing interpretations of the term muscle tone have already been alluded to: a precise definition in the user manual may help to clarify the issue, but that presumes that all users of the instrument will familiarize themselves with the operational definitions before utilizing it in clinical practice. Brunner et al[9] proposed the term 'motor changes' to encompass dystonia, weakness, bradykinesia, and tremor. However, such broad umbrella terms present a challenge to reliability as different raters may introduce variability into the scoring by perceiving all of those symptoms need to be present for severity vs. any single attribute being scored for severity.[18]

Skin temperature was another category demonstrating great variation within the assessment process. 77% of those who currently measured temperature were doing so simply by touching the patient and comparing to the other side, while the remaining 23% are using some form of thermometer; however, an additional 23% stated they would like to measure temperature formally, but did not have the equipment available. No clinicians relied on patient

self-report for this information. The lack of uniformity in terminology and assessment behaviours underscores the need for a user manual with a clear scoring framework and standardized assessment instructions to improve reliability across a multidisciplinary user group.

DISCUSSION

Reflections

This mixed methods study used cognitive debriefing interviews to investigate the presence and patterns of how health care professionals from different disciplines define, assess and formulate judgements on the clinical signs of complex regional pain syndrome. The study also explored the potential to use the content of the interviews to modify the scoring system of a condition-specific outcome measure currently under development (the CB-HI-CRPS) to eliminate potential sources of difference between user groups, with the goal of improving reliability, content and face validity in future psychometric testing. Using semi-structured interviews, we asked a small purposive sample of health professionals from six different disciplines to define the concepts of the CB-HI-CRPS and describe their current assessment practices and potential scaling preferences. The interview transcripts were reviewed using a descriptive content analysis paradigm to examine the scope, themes and relationships of the information[37] and quantitative data reflecting practices and preferences was also collected and analyzed. Four interrelated themes emerged: 1) assessment beliefs and values, 2) beliefs about CRPS, 3) professional roles and functions, and 4) knowledge

translation. The health care professionals in our sample raised fundamental issues related to the assessment of signs and symptoms including balancing objective measurements with clinical observations and patient report, and the essential differences between assessment for the purpose of diagnosis and assessment for the purpose of measuring changes over time.

We also found that assessment practices appeared to differ in specific ways across professional groups; however it is difficult to make generalizations given our small sample. While there was good overall agreement for all of the occupational groups, the current practices of physicians reflected most closely the literature-based concepts of the CB-HI-CRPS, followed by therapists, then nurses. Preferences for dichotomous scales were higher among the professionals who saw their primary role as diagnosis or screening (physicians and nurses), while therapists generally preferred 4 point scales, seeing them as more responsive and thus reflective of their need to evaluate the same patient over time.

Limitations

One of the challenges of this study was the level of self-reflection or reflexivity required of the research team. Interviewing and reviewing data from colleagues was particularly challenging for the experienced clinician (TP) who was aware of the ethical tensions with both recruitment and participation and of the inherent bias of the direct involvement of the tool developer in validation of an

assessment tool. While a strong rapport often existed which may have fostered authenticity in the interviews, the semi-structured format of the interviews created by the research team nevertheless represented an intrinsic bias towards specific content areas. This could also be considered a weakness of the cognitive debriefing interview format as a vehicle for qualitative exploration, as the interviews are necessarily built around the questions from tool under examination.[38]

This study also has several limitations related to sampling. It was conducted over a one-year period, but failed to meet the initial recruitment targets for the sample. All of the interviews were conducted within the Southern Ontario region (ranging from London to Toronto to St. Catherines); however, several of the clinicians interviewed had international training and clinical practice experience in diverse areas such as India, Saudi Arabia, Mexico, Brazil, and the UK. While the range of time in clinical practice spanned from 3 to 35 years, the average experience of participants was 14.9 years, representing a relatively experienced group. It would have been valuable to include clinicians with less experience to provide a greater breadth; however, the snowballing sampling method may have unwittingly eliminated this population, as clinicians tended to nominate other more experienced colleagues, and those with less experience may have been hesitant to volunteer if they did not perceive themselves to have a requisite level of expertise. The large component of occupational therapists included in the sample may also add a bias towards the foundational theories

and values of that profession, including occupational performance and client-centred practice.[39]

Implications

This mixed methods study is intended to supplement and complement the traditional quantitative approach initially used to develop the CB-HI-CRPS. The information gathered from the clinicians will help to inform the scoring system and user manual, adding reliability and validity by strengthening the clinical utility for target users from a variety of disciplines. Moving forward, the intent is to test a revised version of the CB-HI-CRPS that omits the pinprick hyperpathia item; other categories such as muscle tone may be better served with clearer definitions or different terminology. The opportunity exists to test alternate forms of the assessment, using either 4-point or 7-point scales with expanded descriptors/anchors on the individual assessment items in a larger study of reliability and validity, and comparing the results to recommend the most reliable scaling format.

The second contribution of this study is to add insight into how individual clinicians have grappled with the challenges of assessing persons with CRPS. The rich descriptions provided by the participants highlight their struggles to balance observations with patient report and direct measurement: this illustrates the challenge of "objectifying" the variety of signs and symptoms found within this syndrome.[40] A holistic perspective appears to underpin their collective

awareness of the multi-factorial influences on impairments; and they harnessed their experience to identify symptom patterns within the inherent variability of CRPS. Although our sample was very experienced with a mean 14.9 years spent in clinical practice, this ability to be adaptive and address variability could be considered true expertise. [41]

This study underscores the realities of knowledge translation: health care professionals need clear reasons to inform behaviour change, and support to overcome barriers to practice change like time constraints in the clinical environment.[31,32, 42] In our study, the assessment practices of health professionals tended to mirror their current conduct; in essence, they saw their own behaviours as the standard. This may be explained in part by their self-rankings of expertise, with an average of 7.5/10. However, some practitioners did acknowledge the need to evaluate their practices against ever-evolving theories and research, and were actively seeking input to do so: this reflects behaviour that supports the development and maintenance of expertise in health care professionals. [41, 43] Additionally, the insight into the process differences in making clinical judgements between different health professions may provide the foundation for a future grounded theory study on models of clinical judgement formation, and correlation studies on the differences in assessment skills between professional groups, as well as between experienced and novice clinicians, both in the arena of CRPS assessment and beyond.

Do you see what I see?

In order for us to see our patients in the same way we must start to come to some consensus on what we look for (assess) and how we interpret those findings. Standardized scales can serve a role in helping to create more uniform assessments and provide structure to how they are interpreted. Expert consensus about the components of assessment and their interpretation could foster comprehensive assessment of CRPS leading to informed treatment decisions that will continue to advance evidence-based care of persons with this challenging syndrome.

Chapter Two References

1. Stanton-Hicks M, Burton A, Bruehl S, et al. An updated interdisciplinary clinical pathway for CRPS: report of an expert panel. Pain Practice 2002; 2(1): 1-16.

2. Bruehl S, Harden RN, Galer BS, et al. External validation of IASP diagnostic criteria for Complex Regional Pain Syndrome and proposed research diagnostic criteria. Pain 1999; 81(1-2):147–154.

3. Shipton EA. Complex Regional Pain Syndrome – Mechanisms, diagnosis and management. Curr Anaes Crit Care 2009; 20 (5-6): 209-214.

4. de Mos M, de Bruijn AG, Huygen FJ, et al. The incidence of complex regional pain syndrome: a population-based study. Pain 2007;129:12–20.

5. Albazaz, R., Wong, Y.T., Homer-Vanniasinkam, S. Complex regional pain syndrome: a review. Annals of Vascular Surgery 2008; 22(2):297-306.

6. DeMos M, Huygen FJPM, van der Hoeven-Borgman M, et al. Outcomes of the Complex Regional Pain Sydrome. Clin J Pain 2009;25(7):590–597.

7. van Rijn, M., Marinsu, J., Putter, H., et al. Onset and progression of dystonia in complex regional pain syndrome. Pain 2007;130(3): 287-293.

8. Schwartzman RJ, Erwin KL, Guillermo MA. The Natural History of Complex Regional Pain Syndrome. Clin J Pain 2009; 25(4): 275-280.

9. Brunner F, Lienhardt SB, Kissling RO, et al. Diagnostic criteria and follow-up parameters in complex regional pain syndrome type I - a Dephi survey. Eur J Pain 2008;12(1): 48-52.

10. Harden RN, Bruehl S, Perez RSGM, et al. Development of a Severity Score for CRPS. Pain 2010;151(3): 870-876.

11. Oerlemans HM, Goris RJ, Oostendorp RA. Impairment Level Sum Score in Reflex Sympathetic Dystrophy of One Upper Extremity. Arch Phys Med Rehab 1998; 79(Aug): 979-89.

12. Collins S, van Hilten JJ, Marinus J, et al Development of a symptoms questionnaire for complex regional pain syndrome and potentially related illnesses: the Trauma Related Neuronal Dysfunction Symptoms Inventory. Arch Phys Med Rehabil 2008; 89(6):1114-1120.

13. Napoles-Springer A, Santoyo-Olsson J, O'Brien H, et al. Using cognitive interviews to develop surveys in diverse populations. Medical Care, 2006; 44(11): Suppl 3, S21-S30.

14. Housen P, Shannon GR, Simon B, et al. What the resident meant to say: use of cognitive interviewing techniques to develop questionnaires for nursing home residents. The gerontologist. 2008; 48(2):158-169.

15. Willis G, Caspar R, Lessler J. Cognitive Interviewing: a how-to guide. 1999. http://appliedresearch.cancer.gov/areas/cognitive/interview.pdf Retrieved October 7, 2008.

16. Knafl K, Deatrick J, Gallo A, et al. The Analysis and Interpretation of Cognitive Interviews for Instrument Development. Res Nurs Health 2007; 30: 224–234.

17. Fink A, Kosecoff J, Chassin M, et al. Consensus methods: characteristics and guidelines for use. Am J Public Health, 1984;74(9): 979-983.

18. Streiner DL, Norman GR. Health Measurement Scales: a practical guide to their development and use (3rd ed). New York: Oxford University Press; 2003.

19. Kayes NM, McPherson, KM. Measuring what matters: does 'objectivity' mean good science? Disabil and Rehabil 2010; 32(12): 1011–1019.

20. Preston C, Colman AM. Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. Acta Psychol 2000;104(1): 1-15.

21. MacDermid JC, Stratford P. Applying evidence on outcome measures to hand therapy practice. J Hand Ther 2004; 17: 165-173.

22. Tickle-Degnen L. Communicating evidence to clients, managers and funders in Law, M. (ed). Evidence Based Rehabilitation: a guide to practice. Thorofare, NJ: Slack Inc.; 2002; 221-254.

23. Hsieh H, Shannon SE. Three Approaches to Qualitative Content Analysis. Qual Health Res 2005;15(9):1277-1288.

24. Goldsmith L. Sample size determination in qualitative research: when do you know that you have "enough"? Thesis dissertation. Hamilton, Ontario: McMaster University; 1997.

25. Schwarz N. Self Reports: how the questions shape the answers. Am Psychol 1999; 54(2): 93-105.

26. Meyer T, Deck R, Raspe H. Problems completing questionnaires on health status in medical rehabilitation patients. J Rehabil Med 2007; 39(8): 633-639.

27. Pope C, Ziebland S, Mays N. Qualitative research in health care: Analyzing qualitative data. BMJ 2000; 320 (7227): 114-116.

28. Patton MQ. Qualitative Evaluation and Research Methods (2nd ed.). Thousand Oaks, California: Sage Publications; 1990.

29. Lewis JS, Kersten P, McCabe CS, et al. Body perception disturbance: a contribution to pain in Complex Regional Pain Syndrome. Pain 2007;133:111–9.

30. Lewis JS, Kersten P, McPherson KM, Taylor GJ, Harris N, McCabe CS, Blake DR. Wherever is my arm? Impaired upper limb position accuracy in Complex Regional Pain Syndrome. Pain 2010;149: 463–469.

31. Swart CMA, Stins JF, Beeks, PJ. Cortical changes in complex regional pain syndrome (CRPS). Eur J Pain 2009;13:902–907.

32. Straus S E, Tetroe J, Graham I. Defining knowledge translation. Can Med Assoc J 2009;181(3-4):165-168.

33. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. The Lancet, 2003; 362 (9391): 1225-1230.

34. Law M, Baum C. Measurement in Occupational Therapy. In Law M, Baum C, Dunn W. (eds) Measuring Occupational Performance: supporting best practice in occupational therapy.

35. Turk, D.C., and Okifuji, A. Pain terms and taxonomies of pain. In Loeser, J.D. (ed) Bonica's Management of Pain (3rd ed). Philadelphia: Lippincott Williams and Wilkins; 2001.

36. Spicher CJ. Handbook for Somatosensory Rehabilitation (English edition). Paris, France: Sauramps Medical; 2006.

37. Rice PL, Ezzy D. Qualitative research methods: a health focus. Melbourne: Oxford University Press; 1999.

38. Ojanen V, Gogates G. A briefing on cognitive debriefing. Good Clin Pract J, 2006;13(12): 25-29.

39. Sumsion T, Law M. A review of evidence on the conceptual elements informing client-centred practice. CJOT 2006; 73:153-162.

40. Harden RN. Objectification of the diagnostic criteria for CRPS. Pain Med 2010; 11:1212-1215.

41. Alderson D. Developing expertise in surgery. Medical Teacher 2010; 32: 830-836.

42. MacDermid JC, Grewal R, Macintyre NJ. Using an evidence-based approach to measure outcomes in clinical practice. Hand Clinics 2009; 25(1): 97–111.

43. Black LL, Jensen GM, Mostrom E et al. The first year of practice: an investigation of the professional learning and development of promising novice physical therapists. Phys Ther 2010; 90: 1758-1773.

Appendices for Chapter Two

Appendix B: CB-HI-CRPS

Appendix C: User manual for CB-HI-CRPS

Appendix D: SR-HI-CRPS (attached at end of thesis document because of formatting difficulties)

Appendix E: Questions for cognitive debriefing interviews

Clinician Based HI-CRPS

Demographics

DATE: _____ (yy / mm / dd)

SEX: FEMALE □ MALE □ **DATE OF BIRTH**: / / (yy / mm / dd)

SMOKER : YES □ NO □

UPPER EXTREMITY □ **LOWER EXTREMITY** □ **BOTH** □

PRECIPITATING EVENT: YES □ NO □
SPECIFY:_____
_____

INTERVAL BETWEEN PRECIPITATING EVENT AND ONSET OF SYMPTOMS:
□ < 2 weeks □ 2-4 weeks □ 4-6 weeks □ 6-8 weeks □ > 8weeks

TIME SINCE ONSET OF SYMPTOMS:
□ < 3 months □ 3-6 months □ 6-9 months □ 9-12 months □ > 12 months

EMPLOYMENT: Working prior to onset □ YES □ NO
Current status: □unable to work □modified duties □no change

SENSORY COMPONENTS:

Please rate the patients symptoms of:

1. **Hypersensitivity (allodynia):**

0	1	2	3
None	Mild	Moderate	Severe

2. **Hyperpathia: cold**

0	1	2	3
None	Mild	Moderate	Severe

3. **Guarding behavior**

0	1	2	3
None	Mild	Moderate	Severe

Subtotal: /9

AUTONOMIC COMPONENTS:

Check the boxes that describe the patient's current autonomic changes when compared to the opposite side.

4. Skin temperature differences:

0	1	2	3
None	Mild	Moderate	Severe

Affected side is: ☐ **Colder** ☐ **Hotter**

5. Vascular function: mottling

0	1	2	3
None	Mild	Moderate	Severe

6. Sweating (hyperhydrosis)

0	1	2	3
None	Mild	Moderate	Severe

Anhydrosis Y N

7. Edema

0	1	2	3
None	Mild	Moderate	Severe

Subtotal: /12

TROPHIC COMPONENTS:

Check the boxes that describe the patient, comparing to the unaffected limb.

8. Changes in hair growth patterns

0	1	2	3
None	Mild	Moderate	Severe

9. Changes in nails

0	1	2	3
None	Mild	Moderate	Severe

10. Changes in skin quality

0	1	2	3
None	Mild	Moderate	Severe

Subtotal: /9

Please check the box that best describes the patient's current motor changes.

11. The movement is less than would be expected for the patient's initial degree of injury

☐ Strongly Agree ☐ Agree ☐ Slightly Agree ☐ Disagree

12. The movement is less than would be expected for the patient's stage of healing/duration of time since injury.

☐ Strongly Agree ☐ Agree ☐ Slightly Agree ☐ Disagree

13. Abnormal Muscle Tone ☐ Hypotonic ☐ Hypertonic

0 None 1 Mild 2 Moderate 3 Severe

14. Incoordination

0 None 1 Mild 2 Moderate 3 Severe

Subtotal:_____/12

Scoring:

Sensory: /9 Trophic: /9

Autonomic: /12 Motor: /12

TOTAL SCORE: /42 = %

USER MANUAL

CLINICIAN-BASED

HAMILTON INVENTORY FOR

COMPLEX REGIONAL PAIN SYNDROME

(CB-HI-CRPS)

SENSORY COMPONENTS

1) **Allodynia:**

Definition: Pain or unpleasant sensation is experienced when an ordinarily painless stimulus is applied.

Testing: Lightly stroke the affected limb 3 times with a cotton swab; rate response (use both verbal and behavioural feedback)
Instructions: I am going to touch you lightly with this swab; tell me how it feels. (Allow patient to respond then ask) Does it hurt?

0 = None, no complaints of pain
1 = Mild, patient reports discomfort when asked but no physical behaviours evident
2 = Moderate, patient reports pain, may show a behavioural response such as flinching, grimacing, or vocalizing discomfort
3 = Severe, patient reports pain and has a clear behavioural response; may decline to be tested

2) **Cold Hyperalgesia:**

Definition: an exaggerated painful sensation evoked by low-temperature stimulation (Verdugo et al., 2004, p. 369).

Testing: Touch test tube of cold water to skin for 3 seconds. Repeat over 3 different zones within affected area. Rate response as below.
Instructions: I am going to touch you with this test tube of cold water; tell me how it feels to you. (Allow patient to respond then ask) Does it hurt?

0 = None, no complaints of pain; may report that tube feels cold.
1 = Mild, patient reports discomfort with cold but no physical behaviours evident
2 = Moderate, patient reports pain, may show a behavioural response such as flinching, grimacing, or vocalizing discomfort
3 = Severe, patient reports pain and has a clear behavioural response; may decline to be tested

3) Guarding:

Definition: behaviours in which one engages in as a means to protect their affected body part from becoming injured or re-injured; may include postures, use of assistive devices (slings, splints, braces) and vigilance.

Testing: Observe and rate evidence of bracing, holding affected body part close to their body, or avoidance of activity i.e. avoiding social situations for fear of someone hitting affected body part) and reduced interaction with the immediate environment (i.e. does not weight-bear or use limb to hold or manipulate objects. May keep body part covered with clothing or avoid clothing. May use braces or splints as external protection.

0 = None, no guarding behaviour observed or reported.
1 = Mild, patient reports needing to protect limb, may demonstrate some protective behaviour such as altered posture but is able to interact with the environment
2 = Moderate, patient demonstrates clear postural alteration and protective behaviours such as covering limb but is able to interact with the environment to some degree
3 = Severe, patient adopts non-functional postures, demonstrates vigilance and/or restricts environmental interactions; may refuse aspects of assessment

AUTONOMIC COMPONENTS

4) **Skin temperature asymmetry:**

Definition: temperature of the skin that is noticeably warmer or colder than that of the unaffected or contralateral limb.

Testing: Wipe skin areas to remove moisture. Measure surface skin temperature using infra-red thermometer over a distal area of skin (preferably without hair); compare to opposite limb. Note if temperature is increased or decreased compared to unaffected side.

0 = None, temperature side differences of less than 0.5 degrees C measured.
1 = Mild, temperature side differences of 0.5 degrees C to 1.0 degrees measured
2 = Moderate, temperature differences of 1.0 degrees C to 1.5 degrees C measured
3 = Severe, temperature side differences of greater than 1.5 degrees C

5) **Vascular function - mottling:**

Definition: skin colouring appears uneven, patchy, or mottled. May be seen together with redness or cyanosis.

Testing: Patches of red and/or white spots throughout the affected body part greater than observed in contra-lateral unaffected limb. **Instructions:** (If not observed) Does this hand/foot ever get blotchy or patchy looking? How often does that happen?

0 = None, no differences observed.
1 = Mild, slight differences evident and/or patient reports it happens infrequently
2 = Moderate, clear differences evident and/or patient reports it occurs regularly
3 = Severe, clear differences compared to other side and intense coloration of patches evident (i.e. very purple or very red blotches); patient reports limb is always discoloured.

6) **Sweating: Hyperhydrosis**

Definition: excessive sweating or moisture of the skin as compared to the unaffected side. Sweating may occurs in areas not normally prone to perspiration (i.e. the dorsum of the hand).

Testing: Skin may be moist to very wet to touch; look for sheen and/or beads of moisture; moisture quickly reappears after drying with cloth, etc. This differs in frequency and severity compared to opposite limb. Also note moisture in areas not normally prone to sweating, such as around the nails.
Instructions: (if not observed) Does this hand/foot ever get moist or sweaty? How often does that happen?
If patient denies any sweating, then ask:
Has your hand/foot stopped sweating since your injury? Does it stay dry even if you are working up a sweat on the other side? ** score as anydrosis

0 = None, no differences observed.
1 = Mild, slight differences evident and/or patient reports it happens infrequently
2 = Moderate, clear differences evident and/or patient reports it occurs regularly
3 = Severe, clear differences compared to other side; visible beading of sweat. Patient reports limb is always moist or sweaty.

7) **Edema**

Definition: characterized by an accumulation of fluid resulting in an increase in tissue volume.

Testing: Volumetry or figure of 8 measurements recommended for objective measurements; may also consider whether edema is localized or generalized in the limb.

0 = None, no swelling observed.
1 = Mild, slight differences evident between limb size and/or patient reports it comes and goes intermittantly
2 = Moderate, clear differences evident between limbs; may impact on ROM
3 = Severe differences compared to other side; significant impact on ROM; tissues may have pitting or boggy end feel

TROPHIC COMPONENTS

8) **Changes in hair growth patterns**

Definition: characterized by a change in hair, including: a) colour, b) texture, c) distribution, or d) density of follicle growth; can be increased or decreased.

0 = None, no differences observed.
1 = Mild, slight differences evident in one or two characteristics
2 = Moderate, differences evident
3 = Severe, clear differences compared to other side

9) **Changes in nails**

Definition: characterized by changes in nail a) colour, b) ridging, c) thickness.

0 = None, no differences observed.
1 = Mild, slight differences evident in comparison to contralateral limb
2 = Moderate, clear differences evident
3 = Severe, clear differences compared to other side.

10) **Changes in skin quality**

Definition: characterized by thickening or thinning of the epidermis; may be shiny or dull in appearance.

0 = None, no differences observed.
1 = Mild, slight differences evident in comparison to contralateral limb
2 = Moderate differences evident
3 = Severe, clear differences compared to other side; patient may report or show evidence of poor wound healing or ulcerations.

MOTOR DYSFUNCTION COMPONENTS

11) **The movement is less than would be expected for the patient's initial degree of injury**

Definition: Degree of movement is less than would be anticipated in relation to initial injury: based on clinical judgement related to a clear understanding of initial injury vs. the movement loss appears to be the direct result of CRPS.

12) **The movement is less than would be expected for the patient's stage of healing/duration of time since injury**

Definition: Degree of movement is less than would be anticipated in relation to a) time elapsed since initial injury and b) treatment protocols utilized to guide rehabilitation

13) **Abnormal muscle tone**

Definition: characterized by decrease in muscle tone (hypotonic) or increase in muscle tone (hypertonic). Muscle tone is defined as the amount of contraction in a muscle at rest.

0 = None, no differences observed.
1 = Mild, slight differences evident in comparison to contralateral limb
2 = Moderate, clear differences evident.
3 = Severe, clear differences compared to other side.

Testing: Palpation of muscle belly at rest; observations of muscle wasting. Objective assessment can include manual muscle testing and/or evaluations of resistance to passive movement such as the Ashworth scale.

14) **Incoordination**

Definition: the lack of skilful and balanced movement of the affected extremity; may include d*ysdiochokinesia* [the ability to make rapid, smooth, and alternating movements] and/or *dystonia* [uncoordinated muscle movements caused by prolonged contractions].

Testing: rapid alternating movements of hands or feet (ie pronation/supination or inversion/eversion); finger to nose test; heel/toe walking

0 = None, no differences observed.
1 = Mild, slight differences evident in comparison to contralateral limb
2 = Moderate differences in tone evident
3 = Severe, clear differences compared to other side.

Appendix E
Probing Questions for Cognitive Interviews

*** uses likert 7 point scale**

Hypersensitivity – allodynia

How do you define allodynia?

How do you assess allodynia in your patients?

Which scale would you prefer to assess allodynia in your patients? Why?

How would you characterize the anchors of the scale [*mild allodynia, severe allodynia*, *etc*]? Can you describe a patient to me that you think would fit this category?

How hard is it to judge allodynia?

Hyperpathia – pinprick *

What types of words would you use to talk to a colleague about hyperpathia?

How would you assess protective sensation (or sharp/dull sensation) in your patients?

Which scale would you prefer to categorize pinprick hyperpathia in your patients? Why?

How would you characterize the anchors of the scale [*hypalgesia, hyperalgesia*, *etc*]? Can you describe a patient to me that you think would fit this category?

What are you thinking about when you look at the response categories?

Hyperpathia – cold *

How would you assess cold sensation in your patients?

Which scale would you prefer to categorize cold intolerance in your patients?

Why?

How would you characterize the anchors of the scale [*hypalgesia, hyperalgesia*, *etc*]? Can you describe a patient to me that you think would fit this category?

What are you thinking about when you look at the response categories?

Guarding behavior

What do you think this question is asking you to assess?

Which scale would you prefer to categorize guarding behavior in your patients?

Why?

What cues would you use to pick a response category?

How confident are you in making a judgement about this characteristic?

Skin temperature

Do you assess skin temperature in your patients? How?

Which scale would you prefer to categorize skin temperature in your patients?

Why?

What cues would you use to pick a response category? How hard is this to judge?

Mottling

What types of words would you use to describe mottling to a colleague?

How would you assess mottling in your patients?

Which scale would you prefer to categorize mottling in your patients? Why?

What are you thinking about when you look at the response categories?

How confident are you in your abilities to judge mottling accurately?

Hyperhydrosis *

What do you think this question is asking you to assess?

Which scale would you prefer to categorize hyperhydrosis in your patients?

Why?

What cues would you use to pick a response category?

How confident are you in making a judgement about this characteristic?

Edema

Do you assess edema in your patients? How?

Which scale would you prefer to categorize edema in your patients? Why?

What cues would you use to pick a response category?

How confident are you in making a judgement about this characteristic?

Hair growth *

What descriptors would you use as anchors for the ends of this scale?

Which scale would you prefer to categorize hair growth in your patients? Why?

What cues would you use to pick a response category?

How confident are you in making a judgement about this characteristic?

Nail quality

Which scale would you prefer to categorize nail changes in your patients? Why?

What descriptors would you use as anchors for the ends of this scale?

How confident are you in making a judgement about this characteristic?

Skin quality *

Do you assess skin quality in your patients? How? Would that differ between the upper to lower extremity [*if applicable*]?

Which scale would you prefer to categorize changes in skin quality in your patients? Why?

What descriptors would you use as anchors for the ends of this scale?

Movement – expected for degree of injury

What do you think this question is asking you to assess?

What information would you need to make a decision?

Which scale would you prefer to categorize movement in your patients? Why?

How confident are you in making a judgement about this characteristic?

Movement expected for duration since injury

What do you think this question is asking you to assess?

What information would you need to make a decision?

Which scale would you prefer to categorize movement in your patients? Why?

How confident are you in making a judgement about this characteristic?

Muscle tone *

Do you assess muscle tone in your patients? How?

Which scale would you prefer to categorize muscle tone in your patients? Why?

What are you thinking about when you look at the response categories?

What descriptors would you use as scale anchors?

Incoordination

Do you assess incoordination in your patients? How?

Which scale would you prefer to categorize incoordination in your patients?

Why?

What cues would you use to pick a response category?

How confident are you in your abilities to judge incoordination accurately?

Scale 1

Present **Absent**

Scale 2

0	1	2	3
none	mild	moderate	severe

Scale 3a

0	1	2	3	4	5	6
none						marked

Scale 3b (likert)

-3	-2	-1	0	1	2	3
hypo			none			hyper

Thesis Conclusion:

The variability of symptoms, both in scope and intensity, complicate the diagnosis of complex regional pain syndrome. This difficulty in classifying the composite of sensory, autonomic, trophic, and motor features continues in the evaluation of those persons in whom this condition has already been identified. Although the literature contains many tools developed to assist in the challenge of diagnosis, the focus of this dissertation is outcome measurement in complex regional pain syndrome.

The systematic review contained in Chapter One examined condition-specific outcome measures for CRPS, and identified 19 potential assessment tools for use in clinical practice and research. However, many of the tools were narrow in scope (i.e. focused on a single construct such as grip strength or foot function) and the supporting psychometric literature (in English) varied in quality, with an average consensus score of 60% on a standardized evaluation of merit. Many of the assessments were validated only with a homogeneous population of CRPS I, and did not address application to CRPSII. Over half of the assessments were specific to a particular region of the body (usually upper or lower limb). Lack of a gold standard for criterion measurement of pain and disability was also evident. Nonetheless, this study serves to contrast and compare the scope and quality of the current resources available and underscores the need for a comprehensive and rigorous assessment that

incorporates the spectrum of signs and symptoms at the theoretical levels of body structures and functions, as well as garnering the daily impact of CRPS on the activities and participation of the person in their environment. The Hamilton Inventory for CRPS (HI-CRPS), an outcome measure currently in development by the author, was initially conceived in response to these concerns, and this systematic review reinforces the many areas of psychometric evaluation that will need to be addressed as the development of this tool progresses.

Chapter two describes a cognitive debriefing study of the clinician-based portion of the HI-CRPS. While cognitive interviews have primarily been used for the refinement of self-reported assessments, this study employed the cognitive debriefing format to delineate the process used by health care professionals in assessing and formulating scalar judgements on the condition-specific concepts covered by the CB-HI-CRPS.

The descriptive examination of the constructs covered by the Hamilton Inventory in this evaluation has 1) iteratively yielded a common taxonomy that has been utilized to refine the user manual of the tool (see Appendix C); 2) recommended refining the scaling from 7- point to 4- point scales to reflect user preferences yet still maintain potential for discrimination and reliability; and 3) suggested modification of content coverage to reflect the multidisciplinary assessment values. The improved detail and standardization of the user manual should serve to improve reliability (Streiner and Norman, 2003) while moving to 4-point scales potentially may decrease reliability but still maintains the number

of choices within the recommended range of between 4 and 10 (Preston and Colman, 2000).

Future directions for HI-CRPS

Much work remains to realize the vision of a well standardized, reliable, valid and responsive assessment tool for complex regional pain syndrome that will assist clinicians across disciplines to identify patients needing treatment, set appropriate treatment goals and measure treatment outcomes. Going forward, the battery of psychometric testing for the HI-CRPS should include:

1. A measurement study which a) tests inter-rater (CB-HI-CRPS) and intra-rater (SR-HI-CRPS), as well as test-retest reliability on both sections; b) examines content validity as reflected by item-total correlations on the subscales and individual sections; and c) assesses validity of the assignment of items to the subscales through factor analysis
2. Initial steps of divergent validation based on the hypothesis that the clinician based and self-report sections will have a weak correlation, reflecting the findings of previous studies in this and other populations which have examined the relationship between signs and symptoms (or impairments and disabilities); foundational construct (convergent) validation comparing the Impairment Sum Score (Oerlemans, Goris, and Oostendorp, 1998) to the CB-HI-CRPS and the Radboud Skills questionnaire (Oerlemans, Cup, DeBoo, Goris, and Oostendorp, 2000) and Patient Rated Wrist and Hand Evaluation (MacDermid and

Tottenham, 2004) to the SR-HI-CRPS in an upper-extremity population and the Lower Extremity Functional Scale (Stratford et al, 2005) to the SR-HI-CRPS in a lower-extremity population.

3. Assessment of responsiveness with related estimates of minimum detectable differences, and clinically important differences in a heterogeneous CRPS population.
4. Continuing to build content validation by contrasting International Classification of Function (ICF) coding of SR-HI-CRPS with the current symptom, function and socio-emotional subscale designations.

As medical science continues unravel the complex etiology of complex regional pain syndrome, new treatment strategies will evolve and existing ones will be refined. Testing the efficacy of these approaches will require outcome measures with rigorous estimates of reliability, validity and responsiveness for evaluating the continuum of symptoms seen in persons with complex regional pain syndrome, and ultimately mapping the path to recovery. The opportunity exists for development of a condition-specific outcome measure for CRPS affecting any limb(s) that could be used regardless of the diagnostic classification (type I or II) by therapists and physicians to evaluate their patients, make treatment decisions, and monitor progress. This has the potential to support inclusive research with larger, more diverse samples involving the entire spectrum of the CRPS population.

The challenge is also to maintain a holistic framework while evaluating the condition-specific signs and symptoms of CRPS, and reflecting the multidisciplinary paradigm of activity, participation and the context of physical and social environments. Development of the Hamilton Inventory for CRPS was initiated with the goal of addressing these challenges. The work presented in this thesis will serve as a foundation upon which to build future psychometric evaluations of the Hamilton Inventory for CRPS as the process of tool development continues.

References:

Albazaz, R., Wong, Y.T., and Homer-Vanniasinkam, S. (2008). Complex regional pain syndrome: a review. *Annals of Vascular Surgery*, 22(2), 297-306.

Bruehl, S., Harden, R.N., Galer, B.S., et al. (1999). External validation of IASP diagnostic criteria for Complex Regional Pain Syndrome and proposed research diagnostic criteria. *Pain.* 81(1-2),147–154.

Brunner, F., Lienhardt, S.B., Kissling, R.O., Bachmann, L.B., and Weber U. (2008). Diagnostic criteria and follow-up parameters in complex regional pain syndrome type I - a Dephi survey*. European Journal of Pain.* 12(1), 48-52.

Cook, D.A. and Beckman, T.J. (2006). Current concepts in validity and reliability for psychometric instruments: theory and application. *American Journal of Medicine*. 119, 166.e7-166.e16.

Finch, E., Brooks, D., Stratford, P. and Mayo, N. (eds). (2002). Why Measurement Properties are Important. In *Physical Rehabilitation Measures: a guide to enhanced clinical decision making (2nd ed.).* (pp. 27-42). Hamilton, ON, Canada: BC Decker.

Fink, A., Kosecoff, J., Chassin, M., and Brook, R. (1984). Consensus methods: characteristics and guidelines for use. *American Journal of Public Health* 74(9), 979-983.

Galer, B.S., Henderson, J., Perander, J., Jensen, M.P. (2000). Course of symptoms and quality of life measurement in Complex Regional Pain

Syndrome: a pilot survey. *Journal of Pain and Symptom Management*. 20(4), 286-292.

Goodwin, L.D. (2002). Changing conceptions of measurement validity: an update on the new *Standards. Journal of Nursing Education*. 41(3), 100-106.

Harden, R.N., Bruehl, S., Stanton-Hicks, M., Wilson, P.R. (2007). Proposed new diagnostic criteria for complex regional pain syndrome. *Pain Medicine,* 8(4), 326–31.

Harden, R.N., Bruehl, S., Perez, R.S.G.M., Birklein, F., Marinus, J., Maihofner, C., Lubenow, T., Buvanendran, A., Mackey, S., Graciosa, J., Mogilevski, M., Ramsden, C., Schlereth, T., Chont, M., and Vatine, J.J. (2010). Development of a Severity Score for CRPS. *Pain* 151, 870-876.

Henson, R.K. (2001). Understanding internal consistency reliability estimates: a conceptual primer on coefficient alpha. *Measurement and Evaluation in Counselling and Development,* 34, 177-189.

Hicks, C. (2004). *Research Methods for Clinical Therapists: applied project design and analysis (4th ed.).* London: Churchill Livingston.

Housen, P., Shannon, G., Simon, B., Orlando Edelen, M., Cadogan, M., Sohn, L., Jones, M., Buchanan, J., and Saliba, D. (2008). What the resident meant to say: use of cognitive interviewing techniques to develop

questionnaires for nursing home residents. *The gerontologist.* 48(2). 158-169.

Hulley, S.B., Martin, J.N. and Cummings, S.R. (2007). Planning the measurements: precision and accuracy. (p.37-49). In Hulley SB, Cummings SR, Browner WS, Grady DG, and Newman TB (eds) *Designing Clinical Research (3rd ed.)*. Philadelphia: Lippincott Williams and Wilkins.

Kayes, N.M. and McPherson, K.M. (2010). Measuring what matters: does 'objectivity' mean good science? *Disability and Rehabilitation*, 32(12), 1011–1019.

MacDermid, J.C., Richards, R.S., Donner, A., Bellamy, N., and Roth, J.H. (2000). Responsiveness of the Short Form- 36, Disability of the Arm, Shoulder, and Hand Questionnaire, Patient- Rated Wrist Evaluation, and Physical Impairment Measurements in Evaluating Recovery After a Distal Radius Fracture. *Journal of Hand Surgery,* 25A(2), 330-340.

Marinus, J. and Van Hilten, J.J. (2006). Clinical expression profiles of Complex Regional Pain Syndrome, Fibromyalgia and a-specific Repetitive Strain Injury: More common denominators than pain? *Disability and Rehabilitation,* 28(6), 351–362.

Norman, G.R., Brooks, L.R., Coblentz, C.L., and Babcock, C.J. (1992). The correlation of feature identification and category judgements in diagnostic radiology. *Memory and Cognition*, 20(4), 344-355.

Oerlemans HM, Goris RJ, and Oostendorp RA. (1998). Impairment Level Sum Score in Reflex Sympathetic Dystrophy of One Upper Extremity. *Archives of Physical Medicine and Rehabilitation,* 79, 979-89.

Oerlemans, H.M., Cup, E.H., DeBoo, T., Goris, R.J., and Oostendorp, R.A. (2000). The Radboud skills questionnaire: construction and reliability in patients with reflex sympathetic dystrophy of one upper extremity. *Disability and Rehabilitation,* 22(5), 233-245.

Perez, R., Oerlemans, H.M., Zuurmond, W., and De Lange, J. (2003). Impairment level Sum Score for lower extremity Complex Regional Pain Syndrome type I. *Disability and Rehabilitation,* 25(17), 984-991.

Preston, C. and Colman,A.M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica,* 104(1), 1-15.

Regher, G. and Norman, G.R. (1996). Issues in cognitive psychology: implications for professional education. *Academic Medicine.* 71(9), 988-1001.

Reinders, M.F., Geertzen, J.H., and Dijkstra, P.U. (2002). Complex regional pain syndrome type I: use of the International Association for the Study of Pain diagnostic criteria defined in 1994. *Clinical Journal of Pain.* 18(4), 207 –215.

Schwartzman, R.J., Erwin, K.L., and Guillermo, M.A. (2009). The Natural History of Complex Regional Pain Syndrome. *Clinical Journal of Pain*, 25(4), 275-280.

Sechrest L. (2005). Validity of measures is no simple matter. *Health Services Research,* 40(5), 1584-1604.

Stratford, P.W., Hart, D.L., Binkley, J.M., Kennedy, D.M., Alcock, G.K., and Hanna, S.E. (2005). Interpreting lower extremity functional status scores. *Physiotherapy Canada,* 57(2), 154-162.

Streiner, D.L. and Norman, G.R. (2003). *Health Measurement Scales: a practical guide to their development and use (3rd ed.).* New York: Oxford University Press.

Tickle-Degnen L. (2002). Communicating evidence to clients, managers and funders in Law, M. (ed). *Evidence Based Rehabilitation: a guide to practice* (pp. 221-254). Thorofare, NJ: Slack Inc.

Date: _____

Section B: Function and Quality of Life Components

(To be completed by patient)

Please circle the number in the box that best describes the type of pain you experience.

1. The type of pain I experience is:

	Always (6)	(5)	Often (4)	(3)	Sometime (2)	(1)	Never (0)
Sharp	6	5	4	3	2	1	
Sensitive	6	5	4	3	2	1	
Throbbing	6	5	4	3	2	1	
Stabbing	6	5	4	3	2	1	
Aching	6	5	4	3	2	1	
Burning	6	5	4	3	2	1	

Please mark the boxes that best describe you over the last week.

2. I have enough energy to do everything I want to do.

❒ Strongly agree	❒ Agree	❒ Slightly agree	❒ Neutral	❒ Slightly disagree	❒ Disagree	❒ Strongly Disagree

3. I become irritated easily.

❒ Always	❒	❒ Often	❒	❒ Sometimes	❒	❒ Never

4. I am confident that I can manage my signs/symptoms.

❒ Strongly agree	❒ Agree	❒ Slightly agree	❒ Neutral	❒ Slightly disagree	❒ Disagree	❒ Strongly Disagree

5. I feel anxious about my symptoms.

❒ Strongly agree	❒ Agree	❒ Slightly agree	❒ Neutral	❒ Slightly disagree	❒ Disagree	❒ Strongly Disagree

6. Fear of hurting my affected limb prevents me from participating in activities.

❒ Always	❒	❒ Often	❒	❒ Sometimes	❒	❒ Never

7. Please circle the number in the box that best describes your difficulty doing:

	Always (6)	(5)	Often (4)	(3)	Sometimes (2)	(1)	Never (0)	N/A
Getting dressed	6	5	4	3	2	1	0	
Taking a bath	6	5	4	3	2	1	0	
Walking around the home	6	5	4	3	2	1	0	
Household chores	6	5	4	3	2	1	0	
Work	6	5	4	3	2	1	0	
Shopping	6	5	4	3	2	1	0	
Driving	6	5	4	3	2	1	0	
Hobbies	6	5	4	3	2	1	0	

8. My pain stops me from sleeping.

❒ Always	❒	❒ Often	❒	❒ Sometimes	❒	❒ Never

9. I get frustrated easily.

❒ Strongly agree	❒ Agree	❒ Slightly agree	❒ Neutral	❒ Slightly disagree	❒ Disagree	❒ Strongly Disgree

10. I feel my symptoms have affected my relationships.

❒ Strongly agree	❒ Agree	❒ Slightly agree	❒ Neutral	❒ Slightly disagree	❒ Disagree	❒ Strongly Disagree

11. I get tired easily.

❒ Strongly agree	❒ Agree	❒ Slightly agree	❒ Neutral	❒ Slightly disagree	❒ Disagree	❒ Strongly Disagree

12. I feel my affected limb is not a part of my body.

❒ Always	❒	❒ Often	❒	❒ Sometimes	❒	❒ Never

13. My symptoms affect my comfort level with intimacy.

❒ Strongly agree	❒ Agree	❒ Slightly agree	❒ Neutral	❒ Slightly disagree	❒ Disagree	❒ Strongly Disagree

14. I need to concentrate in order to move my affected limbs.

❒ Strongly agree	❒ Agree	❒ Slightly agree	❒ Neutral	❒ Slightly disagree	❒ Disagree	❒ Strongly Disagree

15. Pain prevents me from participating in activities throughout my day.

❒ Always	❒	❒ Often	❒	❒ Sometimes	❒	❒ Never

16. I am confident that I can manage different tasks and activities throughout the day.

❒ Always	❒	❒ Often	❒	❒ Sometimes	❒	❒ Never

17. I experience swelling in my affected limbs.

❒ Strongly agree	❒ Agree	❒ Slightly agree	❒ Neutral	❒ Slightly disagree	❒ Disagree	❒ Strongly Disagree

18. My signs and symptoms embarrass me.

❒ Strongly agree	❒ Agree	❒ Slightly agree	❒ Neutral	❒ Slightly disagree	❒ Disagree	❒ Strongly Disagree

19. I experience muscle cramps or muscle spasms.

❒ Always	❒	❒ Often	❒	❒ Sometimes	❒	❒ Never

20. I worry that people will not believe my symptoms are real.

❒ Strongly agree	❒ Agree	❒ Slightly agree	❒ Neutral	❒ Slightly disagree	❒ Disagree	❒ Strongly Disagree

21. I experience joint stiffness on my affected limb.

❒ Always	❒	❒ Often	❒	❒ Sometimes	❒	❒ Never

22. My swelling come and goes.

❒ Strongly agree	❒ Agree	❒ Slightly agree	❒ Neutral	❒ Slightly disagree	❒ Disagree	❒ Strongly Disagree

23. The people around me are supportive.

❒ Strongly agree	❒ Agree	❒ Slightly agree	❒ Neutral	❒ Slightly disagree	❒ Disagree	❒ Strongly Disagree