

Measuring Pain in Systemic Sclerosis: Comparison of the Short-form McGill Pain Questionnaire Versus a Single-item Measure of Pain

GHASSAN EL-BAALBAKI, JANIE LOBER, MARIE HUDSON, MURRAY BARON, and BRETT D. THOMBS;
Canadian Scleroderma Research Group

ABSTRACT. *Objective.* Studies of pain in systemic sclerosis (SSc) have used a variety of measures, including single-item measures and the 15-item short-form McGill Pain Questionnaire (MPQ-SF). The objective of our study was to compare the performance of the MPQ-SF to a single-item pain numerical rating scale (NRS) and determine whether the MPQ-SF effectively differentiates between sensory and affective components of pain in SSc.

Methods. A cross-sectional, multicenter study of 1091 patients from the Canadian Scleroderma Research Group Registry who completed the MPQ-SF and pain NRS. Correlations of MPQ-SF total scores and pain NRS scores with relevant outcome measures (disability, quality of life, depressive symptoms) were compared. To assess whether the MPQ-SF differentiated between sensory and affective factors, confirmatory factor analysis modeling was used, and correlations of sensory and affective factor scores with other outcome measures were compared.

Results. MPQ-SF total score and the pain NRS correlated similarly with other outcome measures, as did the sensory and affective scores. MPQ-SF sensory and affective factors were highly correlated (0.92), and a single-factor model fit as well as a 2-factor (sensory and affective) model.

Conclusion. The substantial overlap between sensory and affective subscales of the MPQ-SF and the similarity of the MPQ-SF and NRS pain measures compared to other patient-reported outcomes suggest that the 15-item MPQ-SF does not provide tangible advantages compared to the single-item pain NRS. These findings support recommendations to use a single-item NRS pain measure in SSc as it is less burdensome to patients than the MPQ-SF. (J Rheumatol First Release Oct 1 2011; doi:10.3899/jrheum.110592)

Key Indexing Terms:

SYSTEMIC SCLEROSIS
MCGILL PAIN QUESTIONNAIRE

SCLERODERMA

PAIN
SINGLE-ITEM MEASURE

Systemic sclerosis (SSc), or scleroderma, is a rare chronic autoimmune connective tissue disease that affects multiple organ systems. SSc is characterized by thickening and fibrosis of the skin and internal organs due to an increase in the production and deposits of collagen^{1,2}. Patients are typically classified as having either limited or diffuse cutaneous SSc. Diffuse SSc is characterized by more widespread organ system involvement, and thus patients with diffuse SSc general-

ly have a worse prognosis than those with limited SSc³. Disease onset most commonly occurs between the ages of 30 and 50 years, and women are at 4–5 times increased risk compared to men⁴. Median survival time from diagnosis is about 11 years, and patients are 3.7 times more likely to die within 10 years of diagnosis than are individuals without SSc, even after controlling for sex, age, and race⁴. Overall, SSc has a substantial negative effect on quality of life^{4,5,6,7,8,9,10,11,12,13},

From the Departments of Psychiatry, Epidemiology, Biostatistics, and Occupational Health; and Medicine (Division of Rheumatology), McGill University, Montréal; and Department of Psychiatry, Center for Clinical Epidemiology and Community Studies, and Division of Rheumatology, Lady Davis Institute, Jewish General Hospital, Montréal, Québec, Canada.

Supported by a research grant from the Fonds de la Recherche en Santé Québec (B.D. Thombs) and by an American College of Rheumatology Research and Education Rheumatology Investigator Award (B.D. Thombs). Dr. Baron is the director of the Canadian Scleroderma Research Group, which receives grant funding from the Canadian Institutes of Health Research, the Scleroderma Society of Canada and its provincial chapters, Scleroderma Society of Ontario, Sclérodémie Québec, and the Ontario Arthritis Society, and education grants from Actelion Pharmaceuticals and Pfizer Inc. Drs. Thombs and Hudson are supported by New Investigator Awards from the Canadian Institutes of Health Research and Établissement de

Jeunes Chercheurs awards from the Fonds de la Recherche en Santé Québec.

G. El-Baalbaki, PhD; J. Lober, BSc, Department of Psychiatry, McGill University; M. Hudson, MD; M. Baron, MD, Department of Medicine (Division of Rheumatology), McGill University; Center for Clinical Epidemiology and Community Studies, and Division of Rheumatology, Lady Davis Institute, Jewish General Hospital; B.D. Thombs, PhD, Departments of Psychiatry, Epidemiology, Biostatistics, and Occupational Health, and Medicine (Division of Rheumatology), McGill University; Department of Psychiatry, Center for Clinical Epidemiology and Community Studies, and Division of Rheumatology, Lady Davis Institute, Jewish General Hospital.

Address correspondence to B.D. Thombs, Jewish General Hospital, 4333 Cote Ste Catherine Road, Montreal, Quebec H3T 1E4, Canada. E-mail: brett.thombs@mcgill.ca

Accepted for publication July 6, 2011.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2011. All rights reserved.

and patients with SSc report numerous painful symptoms, including skin breakdown, digital ulcers, Raynaud's phenomenon, musculoskeletal pain, and gastrointestinal symptoms^{14,15}.

Existing studies of pain in SSc have used a variety of pain assessment tools, including single-item pain visual analog scales (VAS)^{14,16}, pain numerical rating scales (NRS)¹⁵, and the 15-item short-form McGill Pain Questionnaire (MPQ-SF)^{1,17}. The 2003 OMERACT workshop on SSc (OMERACT 6) recommended the single-item VAS to measure painful symptoms of Raynaud's phenomenon and digital ulcers in SSc¹⁶. The MPQ-SF has not been specifically validated in SSc. Research from other settings, however, has suggested that the MPQ-SF may provide a more informative assessment than single-item measures of pain. Longer measures are often more reliable than short measures, and thus have stronger validity characteristics, which would suggest a potential advantage of the MPQ-SF compared to single-item pain measures. Another possible advantage is that the MPQ-SF provides separate factor scores for sensory and affective pain dimensions^{18,19,20}. Indeed, psychophysical studies show that the affective and sensory dimensions of pain relate differently to nociceptive stimulus intensity and are separately influenced by various psychological factors²¹. However, questions have been raised about the degree to which the MPQ-SF factor scores substantively differentiate these 2 dimensions¹⁸. Studies that have used confirmatory factor analysis (CFA) to assess the factor structure of the MPQ-SF have not directly compared 2-factor models to a single-factor model, but they have reported very high correlations between the sensory and affective factors (0.77 to 0.92)^{22,23,24}, which would suggest that they may be too highly associated to be substantively differentiated.

No studies have compared the performance of single-item measures of pain to the MPQ-SF among patients with SSc. In terms of feasibility and patient burden, single-item measures are clearly advantageous compared to the 15-item MPQ-SF. On the other hand, using the MPQ-SF would be a potentially better option if it performed differently as a measure of general pain intensity than single-item measures or if there were evidence that it produces substantively different sensory and affective pain ratings.

Rare diseases, such as SSc, are increasingly investigated in large, multicenter cohort studies with many different measures administered to patients. In this context, it is important to reduce the burden on patients as much as possible without compromising measurement. Thus, our study had 2 main objectives. The first was to compare the MPQ-SF to a single-item measure of pain among patients with SSc. To determine whether the MPQ-SF may be a more robust measure of pain than single-item methods, we compared correlations of the MPQ-SF total score versus correlations of NRS pain ratings to related outcome measures of depressive symptoms, disability, and mental and physical functioning. The second

objective was to assess the degree to which the MPQ-SF provides substantively distinguishable measures of sensory and affective pain. To do this, we assessed whether a 2-factor CFA model of pain (sensory and affective) provided a substantively better fit to the data than a 1-factor model among patients with SSc. In addition, we compared the correlations of sensory and affective subscale scores to other related outcome measures to determine the degree that they produced similar versus substantively different results.

MATERIALS AND METHODS

Patient sample. The study sample consisted of patients enrolled in the Canadian Scleroderma Research Group (CSRG) Registry. Patients in the registry were recruited from 15 centers across Canada. To be eligible for the Registry, patients must have a diagnosis of SSc made by the referring rheumatologist, be 18 years of age or older, and be fluent in English or French. At each annual registry visit, patients undergo an extensive clinical history, physical evaluation, and laboratory investigations and complete a series of self-report questionnaires. Patients from all centers provided informed consent, and the research ethics board of each center approved the data collection protocol. Only data from patients' initial Registry visit were included in our study.

Measures. Pain was assessed with the MPQ-SF and a pain NRS. Outcome measures used for comparisons included the Center for Epidemiologic Studies Depression scale (CES-D), the Health Assessment Questionnaire-Disability Index (HAQ-DI), and the Mental Component Summary (MCS) and Physical Component Summary (PCS) scores of the Short-form 36 Health Survey Questionnaire. These measures were chosen because they are common outcomes of pain in patients with chronic diseases, including SSc. In addition, we used the Present Pain Intensity Scale (PPI) of the MPQ-SF to define a subsample of patients with substantial pain, for sensitivity analyses.

Demographic and disease variables. Demographic information was based on self-report. Patients' medical histories and disease characteristics were obtained through clinical histories and examinations by study physicians. Limited skin disease was defined as skin involvement distal to the elbows and knees with or without face involvement²⁵. SSc disease duration was determined as the time from onset of non-Raynaud's symptoms based on a clinical history obtained by study rheumatologists. Skin involvement was assessed using the modified Rodnan total body skin score, with scores ranging from 0 to 51²⁶.

The Short-form McGill Pain Questionnaire (MPQ-SF). The MPQ-SF²⁰ consists of a checklist of 15 adjectives, including 11 sensory and 4 affective descriptors of pain. Each descriptor is rated on a 4-point Likert scale for pain intensity that ranges from 0 (no pain) to 3 (severe pain), providing a total score ranging from 0 to 45, as well as subscale scores ranging from 0 to 33 for the sensory subscale and from 0 to 12 for the affective subscale. The MPQ-SF has been shown to have excellent psychometric properties across a wide range of patient groups^{20,22,23,24}, but has not been evaluated in SSc.

Present Pain Intensity scale (PPI). The PPI is a single-item measure of general pain intensity. Patients rate their present level of pain on a 5-point Likert scale that ranges from 0 (no pain) to 5 (excruciating). The PPI is administered as part of the MPQ-SF²⁰, although scored separately.

Pain numerical rating scale (NRS). The study used an 11-point pain NRS. Patients rated their level of current pain intensity according to the statement, "In the past week, how much pain have you had due to illness?". Scores on the 11-point pain NRS range from 0 (no pain) to 10 (very severe pain). The pain NRS has strong validity and reliability across patient groups and performs similarly to the pain VAS²⁷. An advantage of the pain NRS is its increased ease of scoring compared to a VAS²⁷.

The Center for Epidemiological Studies Depression Scale (CES-D). The CES-D is a 20-item scale that measures the frequency of depressive symptoms over the past week on a Likert scale that ranges from 0 (rarely or none

of the time) to 3 (most or all of the time). Total scores range from 0 to 60, with standard cutoffs of > 16 for “possible depression” and > 23 for “probable depression”²⁸. Among a sample of 470 patients with SSc from the CSRG, the CES-D had good reliability and convergent validity with other self-report measures²⁹.

The Health Assessment Questionnaire–Disability Index (HAQ-DI). The HAQ-DI is a 20-item self-report questionnaire designed to assess functional ability in patients with arthritis⁶. The HAQ is the most widely used instrument among patient-reported measures of functional status in SSc. It has been shown to have good face and construct validity and reliability, high sensitivity to change and to predict survival in SSc^{12,30,31,32}. A higher score on the HAQ-DI indicates a greater level of disability, with a total score range of 0 (no disability) to 3 (severe disability).

The Short-form 36 Health Survey Questionnaire (SF-36). The SF-36^{33,34} is the most widely used and evaluated health outcomes measure. It has been shown to be valid and reliable in multiple populations, including SSc^{9,10,30}. The SF-36 comprises 8 domains: physical functioning, social functioning, role limitations related to physical problems, role limitations related to emotional problems, mental health, vitality, bodily pain, and general health perceptions. Each domain can be scored separately, with scores that can range from 0 (worst health state) to 100 (best health state). Domain scores are summarized into a Physical Component Summary (PCS) score and a Mental Component Summary (MCS) score. The PCS and MCS are scored with norm-based scoring based on a general population sample in order to produce T scores for each individual patient (mean of 50 and SD of 10). Our study used version 2 of the SF-36.

Data analyses. To determine whether the MPQ-SF total score is more robustly associated than the single-item NRS is with the set of outcome measures, Pearson’s bivariate correlations with 95% CI were computed. In addition to 95% CI for the correlations, 95% CI for the difference between the MPQ-SF and NRS correlations (Δr) with each of the other outcome measures were also calculated.

To assess the degree to which a 2-factor model (sensory and affective) of the MPQ-SF more accurately describes the data than a single-factor model, and thus to assess the degree to which the sensory and affective factors measure distinct components of pain among patients with SSc, we conducted CFA using Mplus³⁵. CFA is well suited to estimate true interfactor correlations because item error variances are explicitly modeled. In exploratory factor analysis, for example, the use of raw item scores would be expected to reduce interfactor correlations due to measurement error in the items. Because MPQ-SF item responses are measured on an ordinal Likert scale, the weighted least-squares estimator with a diagonal weight matrix and robust standard errors and a mean and variance adjusted chi-square statistic was used with delta parameterization³⁵. Modification indices were used to identify pairs of items within scales for which model fit would improve if error estimates were freed to covary and for which there appeared to be theoretically justifiable shared method effects³⁶. Chi-square goodness-of-fit and 3 fit indices were used to assess model fit, including the Tucker-Lewis Index (TLI)³⁷, the comparative fit index (CFI)³⁸, and the root mean-square error of approximation (RMSEA)³⁹. Since the chi-square test is highly sensitive to sample size and can lead to the rejection of well-fitting models, practical fit indices were emphasized⁴⁰. Guidelines proposed by Hu and Bentler⁴¹ have suggested that models with TLI and CFI close to 0.95 or higher and the RMSEA close to 0.06 or lower are representative of good-fitting models. A CFI ≥ 0.90 ⁴² and RMSEA ≤ 0.08 ⁴³ may also be considered to represent reasonably acceptable model fit.

We used the Mplus DIFFTEST procedure to assess the differences in fit between 2-factor and single-factor MPQ-SF models. A negligible difference between the fit of the 2 models would indicate that the 2-factor model does not improve substantively upon the single-factor model. The DIFFTEST is a chi-square-based procedure, however, so it is sensitive to sample size. Cheung and Rensvold⁴⁴ recommended comparing the change in goodness-of-fit indices, which are not affected by sample size, between 2 models to determine whether there are substantive differences in model fit. Consistent with

Cheung’s recommendations, we compared the CFI between the single-factor and 2-factor models with a difference of ≤ 0.01 indicative of substantively similar models⁴⁴.

Pearson’s bivariate correlations with 95% CI were also computed to assess whether the sensory and affective subscales of the MPQ-SF were differentially associated with other relevant outcome measures, including measures of depressive symptoms (CES-D), mental and physical functioning (SF-36-MCS, SF-36-PCS), and disability (HAQ-DI). In addition to 95% CI for the correlations, 95% CI for the difference between the sensory and the affective subscale correlations (Δr) with each of the other outcome measures were calculated too.

Many patients in the sample had low pain scores, and their inclusion in analyses could potentially affect results. Therefore, we conducted a sensitivity analysis (available from the authors upon request) by repeating all analyses using only data from patients who both reported a discomfiting to excruciating level of pain on the PPI (score 2–5) and endorsed at least 1 item on the MPQ-SF.

RESULTS

Sample characteristics. A total of 1091 patients completed all items of the MPQ-SF. Mean age was 55.3 years (SD 7.4, range 18–88, $n = 1089$ with age data). About 92% ($n = 908$ of 986) were white, 86% ($n = 938$ of 1091) were women, 70% ($n = 704$ of 1012) were married or living as married, 27% ($n = 274$ of 1008) had college education or higher, and 41% ($n = 414$ of 1003) had a full-time or part-time job. Mean duration since diagnosis of SSc was 10.9 years (SD 9.5, range 0–53.8, $n = 1047$), and 37.6% of patients ($n = 402$ of 1068) had diffuse SSc. Mean total body skin score was 10.2 (SD 9.6, range 0–48.0, $n = 1052$). Patients reported a mean total score of 6.1 on the MPQ-SF (SD 7.4, range 0–42), including 4.7 on the MPQ-SF sensory subscale (SD 5.8, range 0–33) and 1.4 on the MPQ-SF affective subscale (SD 2.1, range 0–12). Mean score was 3.6 (SD 2.8, range 0–10, $n = 1025$) on the pain NRS. Means and SD for other outcome variables are presented in Table 1. Of 1091 patients who completed the MPQ-SF, 985 also had complete data for the NRS. A total of 459 patients with PPI score ≥ 2 endorsed at least 1 MPQ-SF item and were included in the sensitivity analysis (supplementary data available from the authors upon request).

Comparison of MPQ-SF correlations with other outcomes versus NRS correlations. As shown in Table 1, correlations of the MPQ-SF and the NRS with the CES-D, HAQ, and SF-36 component scores were moderate, ranging from 0.32 to 0.59. The differences between MPQ-SF correlations and the NRS (Δr) and each of the other measures were small and varied between -0.05 and 0.15 . The NRS had significantly higher correlations than the MPQ-SF with the physical component of the SF-36 ($\Delta r = 0.15$, 95% CI 0.12 to 0.29). Results were similar for the sensitivity analysis (available from the authors upon request).

Confirmatory factor analysis. CFA was used to test the hypothesized 2-factor (sensory and affective) structure for the MPQ-SF. Four pairs of item error covariances were freed based on modification indices. In each case, both members of the pair demonstrated shared method or format features. Error

Table 1. Correlations of the short-form McGill Pain Questionnaire (MPQ-SF), numerical rating scale (NRS), and Present Pain Intensity scale (PPI) to other outcome measures based on total sample (n = 985).

Outcome Measures	Mean (SD)	MPQ-SF r*	r*	NRS Δr^\dagger (95% CI)
CES-D	14.3 (10.5)	0.46	0.41	0.05 (-0.03, 0.15)
HAQ-DI	0.8 (0.7)	0.46	0.51	-0.05 (-0.15, 0.02)
SF-36-MCS	48.3 (11.7)	-0.35	-0.32	-0.03 (-0.12, 0.06)
SF-36-PCS	37.4 (11.2)	-0.44	-0.59	0.15 (0.12, 0.29)

* All correlations $p < 0.05$ (2-tailed). \dagger Difference between MPQ-SF and NRS correlation coefficients. CES-D: Center for Epidemiologic Studies Depression scale; HAQ-DI: Health Assessment Questionnaire-Disability Index; SF-36: Short-form 36 Health Survey Questionnaire; MCS: mental component summary; PCS: physical component summary.

variances were freed to covary for (1) throbbing with shooting, (2) shooting with sharp, (3) shooting with stabbing, and (4) stabbing with sharp. Model fit for the 2-factor model in the total sample was good [chi-square (55) = 198.3, $p < 0.001$, CFI = 0.97, TLI = 0.99, RMSEA = 0.05]. The correlation between the sensory and affective factors was 0.92. Model fit was also good for the single-factor model [chi-square (54) = 234.2, $p < 0.001$, CFI = 0.96, TLI = 0.99, RMSEA = 0.06]. As expected, given the large sample size, the chi-square-based DIFFTEST procedure that compared the fit of the 2 models was statistically significant [chi-square (1) = 26.1, $p < 0.001$]. The differences in goodness-of-fit indices between the 2-factor and single-factor models, however, were negligible (Δ CFI = 0.01, Δ TLI = 0.00, Δ RMSEA = -0.01; Table 2). For patients with significant levels of pain who were included in the sensitivity analysis, results were similar, and the 2 factors correlated at 0.86 with nonsubstantive differences in fit between 1- and 2-factor models (supplementary data available from the authors upon request).

Correlations of sensory and affective subscale scores with NRS and other outcome measures. The raw correlation between the sensory and affective subscale scores was 0.70 (95% CI 0.68 to 0.74, $p < 0.001$). Similar to analyses with the MPQ-SF total score, correlations between the sensory and affective scores and the CES-D, HAQ, and SF-36 component scores were moderate (0.30 to 0.49). Bivariate correlations between the sensory and affective subscales, the NRS, and other outcome measures were similar (Table 3), and all differences between sensory and affective subscale correlations (Δr) with other outcome measures were < 0.10 . A statistically significant difference was found only for the MCS ($\Delta r = 0.09$, 95% CI 0.01 to 0.18). Results from sensitivity analyses were similar (available from the authors upon request).

DISCUSSION

We found that the sensory and affective factors of the MPQ-SF were highly correlated in a model that included all patients ($r = 0.92$) and a model of data only from patients who

Table 2. Item means and factor loadings of MPQ-SF sensory and affective subscales for total sample (n = 1091).

Items	Mean (SD)	2-Factor Model	Single-Factor Model
Sensory factor			
Throbbing	0.43 (0.80)	0.69	0.68
Shooting	0.30 (0.71)	0.70	0.70
Stabbing	0.31 (0.74)	0.73	0.72
Sharp	0.39 (0.80)	0.77	0.76
Cramping	0.47 (0.80)	0.61	0.61
Gnawing	0.27 (0.67)	0.74	0.73
Hot — burning	0.43 (0.82)	0.73	0.72
Aching	0.91 (1.00)	0.80	0.79
Heavy	0.32 (0.74)	0.84	0.74
Tender	0.74 (0.96)	0.74	0.73
Splitting	0.20 (0.60)	0.79	0.78
Affective factor			
Tiring — exhausting	0.82 (1.02)	0.84	0.80
Sickening	0.22 (0.63)	0.88	0.85
Fearful	0.20 (0.58)	0.76	0.73
Punishing — cruel	0.13 (0.50)	0.90	0.87
Correlation of sensory and affective factors		0.92	—

MPQ-SF: short-form McGill Pain Questionnaire.

Table 3. Correlations of MPQ-SF subscales with NRS and outcome measures for the total sample (n = 1091).

Outcome Measures	Mean (SD)	MPQ-SF	MPQ-SF	Difference
		Sensory r*	Affective r*	Δr (95% CI)
CES-D**	14.3 (10.5)	0.42	0.49	-0.07 (-0.17, 0.00)
HAQ-DI***	0.8 (0.7)	0.45	0.39	0.06 (-0.01, 0.16)
SF-36-MCS†	48.3 (11.8)	-0.30	-0.39	0.09 (0.02, 0.19)
SF-36-PCS†	37.3 (11.2)	-0.43	-0.37	-0.06 (-0.16, 0.02)
Single-item measure NRS††	3.6 (2.8)	0.57	0.45	-0.12 (0.08, 0.24)

* All correlations $p < 0.05$ (2-tailed). ** n = 1005; *** n = 1024; † n = 1012; †† n = 1025. For definitions see Table 1.

reported at least “discomforting” pain and endorsed at least 1 item on the MPQ-SF ($r = 0.86$). Both the very high intercorrelations of these 2 factors and the virtually identical fit indices that were obtained when 2-factor and single-factor models of the MPQ-SF were compared provide evidence that the sensory and affective factors of the MPQ-SF are not sufficiently distinguishable to make a practical difference among patients with SSc. When correlations of the sensory and affective subscales with measures of mental health, physical health, and disability were compared, the affective subscale of the MPQ-SF tended to be somewhat more closely associated with mental health outcome measures, but did not differ from the sensory subscale in comparisons to measures of physical function or disability.

When the correlations of the MPQ-SF total score with mental health, physical health, and disability outcome measures were compared to similar correlations with the pain NRS, there were few significant or substantially meaningful differences. Notably, when differences were detected, the single-item measure tended to produce more robust correlations, which is not what one would expect if the single-item measure was less valid due to poor reliability compared to the 15-item MPQ-SF.

Three previous studies used CFA methods to assess the fit of the standard 2-factor model of the MPQ-SF^{22,23,24} in non-rheumatology samples. Two of these^{22,24} were studies of 188 and 373 patients with chronic back pain, and each found that the sensory and affective subscales were highly intercorrelated (0.88 and 0.89), similar to the results of our study. A third study²³, conducted in a sample of 338 patients with major burn injury, reported a somewhat lower correlation of 0.77 and somewhat better fit indices for the 2-factor model compared to the single-factor model. Although the burn patients in that study had higher mean MPQ-SF scores (15.1) than those in our study (6.1 in total sample, 11.1 in sample of patients with pain), mean scores were similar to the 2 chronic back pain samples (14.9 and 16.5). This suggests that different intensities of pain across the samples would not likely explain the somewhat different model fit among burn patients. Regardless, the consistency of findings across samples of patients with chronic pain, including patients with SSc in our

study and patients with chronic back pain from previous studies^{22,24}, suggests that the sensory and affective subscales are not sufficiently different in these patient groups as to produce incremental benefit compared to the use of a single-item pain NRS.

On the whole, there were no consistent patterns to suggest that the MPQ-SF would be substantively advantageous for measuring pain among patients with SSc compared to a single-item measure of pain. This finding supports recommendations for the use of single-item measures of pain in SSc¹⁶. Indeed, without evidence for additional measurement benefits of the MPQ-SF versus single-item measures, single-item measures are preferable because they limit the response burden of patients and are more cost- and time-effective for researchers and clinicians. It is possible that the adjectives used to describe pain in the MPQ-SF may be useful descriptively or could be helpful in determining pain treatment for patients with SSc, but this should be examined in future research. It is also possible that despite their high correlations, the MPQ-SF and NRS may measure somewhat different components of pain, but this also would need to be demonstrated.

A number of limitations should be considered in evaluating these results. Our study was based on a convenience sample of patients with SSc. The sample included here tended to have a stable pattern of disease, as indicated by the mean duration of 10.9 years since time of diagnosis. Patients who are not treated by a rheumatologist, who have more severe SSc so as to limit participation in the CSRG Registry, or who die early in the course of their disease may be undersampled in the CSRG Registry. Therefore, our sample may have included an overrepresentation of healthier patients with SSc. On the other hand, our patient sample was drawn from a large number of centers across Canada, and the demographic and disease characteristics are consistent with other outpatient samples⁴.

Another limitation is that the MPQ-SF has not been validated in SSc. Related to this, given the cross-sectional design of this study, we were unable to conduct test-retest measures of reliability or to test for responsiveness to treatment of the MPQ-SF, its subscales, and the single-item pain measures. Other measures of pain or sleep, for instance, which may have been informative, were not available. This study did not assess

the relative reliability, extent of measurement error, or responsiveness of the MPQ-SF compared to a single-item measure of pain in SSc. The study was observational and did not include a pain treatment component, so data could not be stratified based on treatment offered. Finally, the study did not account for different sources of pain, and it is possible that patients with pain caused by reflux, joint pain, or other sources might have different profiles.

The results of our study showed that the MPQ-SF did not perform substantively differently than a single-item pain NRS compared to other important patient-reported outcomes. Further, the sensory and affective factors of the MPQ-SF were too highly correlated to meaningfully distinguish separate constructs among patients with SSc. These results support recommendations for the use of single-item pain measures in SSc and suggest that the MPQ-SF would add to patient burden without improving measurement substantively. These results have implications for data collection in cohort studies and clinical trials, as well as for routine pain assessment in clinical settings.

APPENDIX

List of study collaborators. Canadian Scleroderma Research Group recruiting rheumatologists: J. Pope, London, Ontario; M. Baron, Montreal, Quebec; J. Markland, Saskatoon, Saskatchewan; N.A. Khalidi, Hamilton, Ontario; A. Masetto, Sherbrooke, Quebec; E. Sutton, Halifax, Nova Scotia; N. Jones, Edmonton, Alberta; D. Robinson, Winnipeg, Manitoba; E. Kaminska, Hamilton, Ontario; P. Docherty, Moncton, New Brunswick; C.D. Smith, Ottawa, Ontario; J-P. Mathieu, Montreal, Quebec; S. LeClercq, Calgary, Alberta; M. Hudson, Montreal, Quebec; S. Ligier, Montreal, Quebec; T. Grodzicky, Montreal, Quebec; C. Thorne, Newmarket, Ontario; S. Mittoo, Winnipeg, Manitoba; M. Fritzler, Advanced Diagnostics Laboratory, Calgary, Alberta.

REFERENCES

- Benrud-Larson LM, Haythornthwaite JA, Heinberg LJ, Boling C, Reed J, White B, et al. The impact of pain and symptoms of depression in scleroderma. *Pain* 2002;95:267-75.
- Steen VD. Clinical manifestations of systemic sclerosis. *Semin Cutan Med Surg* 1998;17:48-54.
- Medsker Jr TA. Natural history of systemic sclerosis and the assessment of disease activity, severity, functional status, and psychologic well-being. *Rheum Dis Clin North Am* 2003;29:255-73.
- Mayes MD. Scleroderma epidemiology. *Rheum Dis Clin North Am* 2003;29:239-54.
- Bernatsky S, Hudson M, Panopalis P, Clarke AE, Pope J, Leclercq S, et al. The cost of systemic sclerosis. *Arthritis Rheum* 2009;61:119-23.
- Clements PJ, Wong WK, Hurwitz EL, Furst DE, Mayes M, White B, et al. Correlates of the Disability Index of the Health Assessment Questionnaire: a measure of functional impairment in systemic sclerosis. *Arthritis Rheum* 1999;42:2372-80.
- Danieli E, Airo P, Bettoni L, Cinquini M, Antonioli CM, Cavazzana I, et al. Health-related quality of life measured by the Short-Form 36 (SF-36) in systemic sclerosis: correlations with indexes of disease activity and severity, disability, and depressive symptoms. *Clin Rheumatol* 2005;24:48-54.
- Hudson M, Steele R, Lu Y, Thombs BD, Baron M. Work disability in systemic sclerosis. *J Rheumatol* 2009;36:2481-6.
- Hudson M, Steele R, Taillefer S, Baron M. Quality of life in systemic sclerosis: psychometric properties of the World Health Organization Disability Assessment Schedule II. *Arthritis Rheum* 2008;59:270-8.
- Hudson M, Thombs BD, Steele R, Panopalis P, Newton E, Baron M. Health-related quality of life in systemic sclerosis: a systematic review. *Arthritis Rheum* 2009;61:1112-20.
- Ioannidis JP, Vlachoyiannopoulos PG, Haidich AB, Medsker TA Jr, Lucas M, Michet CJ, et al. Mortality in systemic sclerosis: an international meta-analysis of individual patient data. *Am J Med* 2005;118:2-10.
- Poole JL, Steen VD. The use of the Health Assessment Questionnaire (HAQ) to determine physical disability in systemic sclerosis. *Arthritis Care Res* 1991;4:27-31.
- Wilson L. Cost-of-illness of scleroderma: the case for rare diseases. *Semin Arthritis Rheum* 1997;27:73-84.
- Malcarne VL, Hansdotter I, McKinney A, Upchurch R, Greenbergs HL, Henstorf GH, et al. Medical signs and symptoms associated with disability, pain, and psychosocial adjustment in systemic sclerosis. *J Rheumatol* 2007;34:359-67.
- Schieir O, Thombs BD, Hudson M, Boivin JF, Steele R, Bernatsky S, et al. Prevalence, severity, and clinical correlates of pain in patients with systemic sclerosis. *Arthritis Care Res* 2010;62:409-17.
- Merkel PA, Clements PJ, Reveille JD, Suarez-Almazor ME, Valentini G, Furst DE. Current status of outcome measure development for clinical trials in systemic sclerosis. Report from OMERACT 6. *J Rheumatol* 2003;30:1630-47.
- Edwards RR, Goble L, Kwan A, Kudel I, McGuire L, Heinberg L, et al. Catastrophizing, pain, and social adjustment in scleroderma: relationships with educational level. *Clin J Pain* 2006;22:639-46.
- Fernandez E, Turk DC. Sensory and affective components of pain: separation and synthesis. *Psychol Bull* 1992;112:205-17.
- Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975;1:277-99.
- Melzack R. The Short-Form McGill Pain Questionnaire. *Pain* 1987;30:191-7.
- Price DD. Psychological and neural mechanisms of the affective dimension of pain. *Science* 2000;288:1769-72.
- Beattie PF, Dowda M, Feuerstein M. Differentiating sensory and affective-sensory pain descriptions in patients undergoing magnetic resonance imaging for persistent low back pain. *Pain* 2004;110:189-96.
- Mason ST, Arceneaux LL, Abouhassan W, Lauterbach D, Seebach C, Fauerbach JA. Confirmatory factor analysis of the Short-Form McGill Pain Questionnaire with burn patients. *Eplasty* 2008;8:e54.
- Wright KD, Asmundson GJ, McCreary DR. Factorial validity of the Short-Form McGill Pain Questionnaire (SF-MPQ). *Eur J Pain* 2001;5:279-84.
- LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsker TA Jr, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. *J Rheumatol* 1988;15:202-5.
- Clements P, Lachenbruch P, Siebold J, White B, Weiner S, Martin R, et al. Inter and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. *J Rheumatol* 1995;22:1281-5.
- Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. *Pain* 1986;27:117-26.
- Radloff L. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 1977;7:189-98.
- Thombs BD, Hudson M, Schieir O, Taillefer SS, Baron M. Reliability and validity of the Center for Epidemiologic Studies Depression scale in patients with systemic sclerosis. *Arthritis Rheum* 2008;59:438-43.

30. Khanna D, Furst DE, Clements PJ, Park GS, Hays RD, Yoon J, et al. Responsiveness of the SF-36 and the Health Assessment Questionnaire Disability Index in a systemic sclerosis clinical trial. *J Rheumatol* 2005;32:832-40.
31. Poole JL, Williams CA, Bloch DA, Hollak B, Spitz P. Concurrent validity of the Health Assessment Questionnaire Disability Index in scleroderma. *Arthritis Care Res* 1995;8:189-93.
32. Steen VD, Medsger TA Jr. The value of the Health Assessment Questionnaire and special patient-generated scales to demonstrate change in systemic sclerosis patients over time. *Arthritis Rheum* 1997;40:1984-91.
33. Ware JJ, Snow K, Kosinski M, Gandek B. SF-36 Health Survey: manual and interpretation guide. Boston: The Health Institute, New England Medical Center; 1993.
34. Ware JJ, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). Conceptual framework and item selection. *Med Care* 1992;30:473-83.
35. Muthén LK, Muthén BO. Mplus user's guide version 3.11. Los Angeles: Muthén LK and Muthén BO; 1998-2004.
36. McDonald RP, Ho MH. Principles and practice in reporting structural equation analyses. *Psychol Methods* 2002;7:64-82.
37. Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika* 1973;38:1-10.
38. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull* 1990;107:238-46.
39. Steiger JH. Structural model evaluation and modification: An interval estimation approach. *Multivar Behav Res* 1990;25:173-80.
40. Reise SP, Widaman KF, Pugh RH. Confirmatory factor analysis item response theory: two approaches for exploring measurement invariance. *Psychol Bull* 1993;114:552-66.
41. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling* 1999;6:131-52.
42. Kline RB. Principles and practice of structural equation modeling. New York: Guilford Press; 2005.
43. Browne MW, Cudeck R, eds. Alternative ways of assessing fit. Newbury Park, CA: Sage Publications; 1993.
44. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Modeling* 2002;9:233-55.