

Pain Extent and Diagnosis: Development and Validation of the Regional Pain Scale in 12,799 Patients with Rheumatic Disease

FREDERICK WOLFE

ABSTRACT. Objective. To develop and validate a pain scale that measures the extent of body pain.

Methods. A total of 12,799 patients with rheumatoid arthritis (RA), osteoarthritis (OA), and fibromyalgia (FM) completed a mailed survey regarding the location and intensity of their pain in 38 articular and nonarticular regions. The data were analyzed using item response theory (IRT) by nonparametric Mokken analysis followed by Rasch analysis. The resultant scale was examined for its association with clinical severity variables and its ability to distinguish patients diagnosed with and without FM.

Results. The resultant 19 item regional pain scale (RPS) was composed primarily of nonarticular regions. The scale had strong scalability as measured by the Mokken H statistic ($H = 0.52$), and satisfied the Mokken monotonicity and double monotonicity criteria. The RPS also fit the Rasch model and had satisfactory reliability and separation statistics. Of all clinical variables assessed by survey, the RPS best identified patients diagnosed with FM. In addition, the scale correlated with measures of clinical severity, regardless of diagnosis, and predicted measures of utilization.

Conclusion. The RPS is a valid scale of pain extent. It can be useful to identify patients with FM or can be used to develop a new definition of FM, even among patients with concomitant illnesses such as RA and OA. In addition, it is a measure of pain extent that is disease independent, and works as well in RA and OA as in FM to identify patients with increased severity and resource utilization. (J Rheumatol 2003;30:369–78)

Key Indexing Terms:

REGIONAL PAIN SCALE
PAIN

FIBROMYALGIA

RHEUMATOID ARTHRITIS
PAIN EXTENT

Pain is one of several key variables in fibromyalgia (FM), but it is also central to other rheumatic disorders such as rheumatoid arthritis (RA) and osteoarthritis (OA). In the clinic and in clinical trials, pain is most often measured using a visual analog or similar scale to capture pain intensity, ignoring, for reasons of feasibility and simplicity, other complex elements of pain that are assessed by instruments such as the McGill Pain Scale¹. Yet along with pain intensity, one of the most common clinical assessments is that of pain location and extent (“Tell me where are you having pain?”). Pain extent is also important in research and diagnosis. Widespread pain, for example, is one of the 2 criteria for classification of FM by the American College of Rheumatology (ACR)², and widespread pain and/or pain extent has been the subject of many investigations^{3–20}.

Almost all studies that have measured pain extent have done so with pain diagrams or drawings (Figure 1). Although simple to administer, pain diagrams are difficult to code and score, and quantification of pain extent is, perhaps consequently, a rarely performed measure except in the quantification of joint pain by specific yes/no questions. It is common in arthritis care to count the number of specific tender joints. In addition, self-report measures of joint tenderness of specific joints are also available^{21–27}. None of these specific measures, however, includes nonarticular regions. As noted, the extent of nonarticular region pain is important not only in the identification of FM and FM-like syndromes, but also in the identification of distress related features of medical illnesses.

We investigated a series of questions: (1) whether it is possible to develop a valid and psychometrically sound index of widespread pain; (2) whether such an index should include articular and nonarticular regions; and (3) whether a quantitative measure of widespread pain adds knowledge regarding diagnosis and characterization of illness, including assessment of severity and outcome. We report here on the development and preliminary validation of the Regional Pain Scale (RPS).

From the National Data Bank for Rheumatic Diseases — Arthritis Research Center Foundation and University of Kansas School of Medicine, Wichita, Kansas, USA.

F. Wolfe, MD, Director, National Data Bank for Rheumatic Diseases.

Address correspondence to Dr. F. Wolfe, Arthritis Research Center Foundation, 1035 N. Emporia, Suite 230, Wichita, KS 67214.

E-mail: fwolfe@arthritis-research.org

Submitted February 19 2002; revision accepted July 19, 2002.

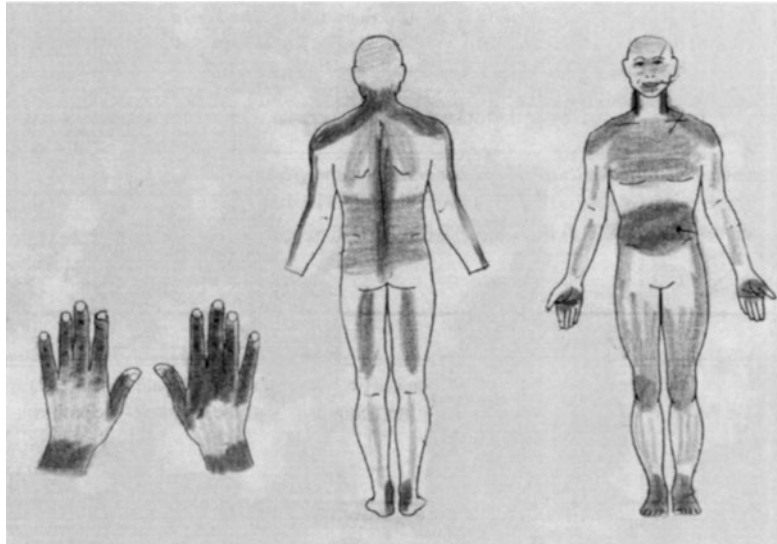


Figure 1. Pain diagram from a patient with fibromyalgia. Figure displays pain in all sites as well as intensity of pain represented by darkness of penciled-in areas.

MATERIALS AND METHODS

Patients. We surveyed 12,799 patients with rheumatic disease at 6 month intervals concerning their illnesses during the period from January 1998 through December 1999. Of these patients, 9576 (74.8%) had RA, 2689 (21.0%) had OA of the knee or hip, and 543 (4.2%) had FM. Diagnoses were made by the patients' rheumatologists ($n = 641$). Participants completed between one and 4 surveys, and one survey per participant was randomly selected for analysis. Patients in these surveys were participants in the National Data Bank (NDB) rheumatic disease outcome project, and had been referred to the NDB by their rheumatologist.

Questionnaires. Participants were asked to complete semiannual detailed 28 page questionnaires about all aspects of their illness. At each questionnaire assessment, demographic variables were recorded, including sex, age, ethnic origin, education level, and current marital status. Study variables included the Stanford Health Assessment Questionnaire functional disability index (HAQ disability)^{28,29}, a visual analog (VAS) pain scale, a VAS global disease severity, VAS sleep and fatigue scales³⁰, the Arthritis Impact Measurement Scales (AIMS)^{31,32}, anxiety and depression scales^{33,34}, the SF-36 mental and physical component scales (MCS and PCS)³⁵, the Western Ontario and McMaster University Osteoarthritis Index (WOMAC) pain, stiffness and function scale^{36,37}, and Likert scales that assessed current satisfaction with health and current perceived health. To measure health quality of life, we used the VAS scale from the EuroQol^{38,39}.

In addition, a specific set of questions relating to articular and non-articular sites was included (Appendix). These sites are identified in Tables 1 and 2. The identifying question read: "Please indicate below the amount of pain and/or tenderness you have had over the past 7 days in each of the joint and body areas listed below. Please make an X in the box that best describes your pain or tenderness. Please be sure to mark both right and left sides separately." Below these instructions, a series of site descriptions (e.g., Shoulder, Lt. and Shoulder, Rt.) were followed by 4 boxes labeled none, mild, moderate, and severe.

Statistical methods. Psychometrics is concerned with the scaling properties of questionnaires⁴⁰ and therefore with issues of validity, reliability, and ultimately, clinical applicability and usefulness. Studies of the HAQ in RA^{41,42} and FM⁴³, and the SF-36 in a number of conditions including arthritis^{44,45}, have shown departures from item response theory (IRT) or modern psychometrics. IRT has been used in the development of questionnaires in rheumatology^{43,46}.

Mokken analysis⁴⁷ and Rasch analysis⁴⁸ were used for scale development in this study. Mokken analysis is a nonparametric approach to IRT and has been used to create or analyze various scales^{47,49-54}. Mokken analysis was performed using MSP5 for Windows⁴⁷.

Using the assumptions of IRT, one seeks to measure a latent or unobserved variable. In this study, for example, we attempted to measure the latent variable, extent/severity of body pain. One assumption of IRT is that for each item of a scale the probability of a positive response is a monotonically nondecreasing function of the latent variable or trait θ . This is known as the monotonicity assumption (M). The relationship between the latent trait and the probability of a positive response is the item characteristic curve (ICC). Each item has its own characteristic curve. In a given scale, when the ICC do not intersect, they satisfy the condition of double monotonicity (DM). Scales should also be unidimensional, in that they measure only one concept. An additional important requirement of IRT is that the items be locally independent. This means that once the underlying value of θ has been accounted for, the probability of a response to one item is unrelated to the probability of response to any other item.

The extent to which a scale conforms to the perfect data structure described above can be expressed by the Mokken H statistic (scalability). Scalability is considered moderate if $H \geq 0.4$, and strong if $H \geq 0.5$. Various other statistics are available to characterize aspects of a Mokken scale. Restscore crit is a check on monotonicity. Values of restscore crit > 80 represent serious violations of monotonicity, and values < 40 may represent sampling variation and are unimportant. Restsplit and Pmatrix tests are checks for nonintersection (DM). Pmatrix test values have a tendency to be high with longer scales, and their main value may be in comparisons of different items for DM.

Rasch analysis, a form of analysis based on item response theory, was developed in the 1950s by George Rasch, a Danish mathematician⁵⁵, and has been used in a number of analyses of functional assessment questionnaires in rheumatology^{42,46,56,57}, orthopedics⁵⁸, and physical medicine⁵⁹⁻⁶³.

Rasch analysis is a method for obtaining objective, fundamental, linear measures (qualified by standard errors and quality control fit statistics) from stochastic observations of dichotomous or ordered category responses⁶⁴. The details of this method are available in a number of conceptually and mathematically simple^{48,65} and more complex texts^{40,55,66,67}.

The Rasch model develops theory and methods regarding a number of characteristics of questionnaire scales. These include unidimensionality, scale length, assessment of difficulty of items, positioning of items along

Please indicate below the amount of pain and/or tenderness you have had over THE PAST 7 DAYS in each of the joint and body areas listed below. Please make an X in the box that best describes your pain or tenderness. Be sure to mark both right side and left side separately. If you have had no pain or tenderness in a particular joint or body part, mark "None." There should be an answer for every joint or body part listed.

JOINTS	None	Mild	Mod	Severe	OTHER BODY AREAS	None	Mild	Mod	Severe
Shoulder, Lt. Shoulder, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Jaw, Lt. Jaw, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Elbow, Lt. Elbow, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Lower Back Upper Back	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Wrist, Lt. Wrist, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Neck	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hand knuckles, Lt. Hand knuckles, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Upper arms, Lt. Upper arms, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Finger knuckles, Lt. Finger knuckles, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Lower arms, Lt. Lower arms, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hip, Lt. Hip, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Upper leg, Lt. Upper leg, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Knee, Lt. Knee, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Lower leg, Lt. Lower leg, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ankle, Lt. Ankle, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Headache Chest	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ball of foot, Lt. Ball of foot, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Abdomen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Heel, Lt. Heel, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
Foot arch, Lt. Foot arch, Rt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					

APPENDIX. The test questionnaire.

the length of the scale, statistical fit, scale linearity, floor and ceiling effects, and reliability.

An important concept to understand for this study is item difficulty. The term is derived from the use of IRT in educational assessment. In clinical use, difficulty can be thought of as a measure of severity. For example, it is easier (more common) to have low back pain and more difficult to have jaw pain. Stochastically, the presence of jaw pain implies the presence of back pain, such that on a severity scale back pain has the least severity (difficulty) and jaw pain the most severity (difficulty). Item severity or difficulty can be seen in Table 1 and Figure 2. Examination of the figure and table allows one to gauge other Rasch characteristics, including length and linearity of the scale, positioning of the items, and floor and ceiling effects.

Rasch analysis produces 2 fit statistics. The mean square INFIT and mean square OUTFIT statistics are measures of signal to noise that allow one to determine how well an item or an individual step of an item fits the Rasch model⁶⁴. An item that has a high INFIT or OUTFIT statistic (> 1.3) may not fit the model because it is noisy (not read well or read with error) or is measuring a 2nd dimension. A high INFIT statistic indicates that unexpected responses are found at this level of the person's ability. A high outlier sensitive fit statistic (OUTFIT) indicates that unexpected responses are found for the person's ability. In general, abnormal INFIT statistics are more important than abnormal OUTFIT statistics.

We used Mokken analysis in the initial phases of questionnaire devel-

opment and analysis and used Rasch analysis for confirmation and determination of item difficulties.

Data were also analyzed by linear and logistic regression using Stata 7.0⁶⁸ and by classification and regression tree methodology (CART)^{69,70}. Statistical differences in the ability of the variables to detect differences in diagnostic groups in univariate analyses were assessed by logistic regression followed by the determination of the Bayesian Information Criterion (BIC or BIC⁷¹). The BIC is a measure of overall fit and a means to compare nested and nonnested models.

Statistical significance was set at the 0.05 level. In some tables, patients with OA are omitted for simplicity of viewing 2 groups instead of 3 groups.

RESULTS

Study participants. The demographic characteristics of the 3 groups were as follows for the 9576 RA patients, the 2689 OA patients, and the 534 FM patients, respectively: median age 58.7, 65.7, 54.6 years; median disease duration 9.9, 11.2, 17.2 years; percentage male 23.1, 18.8, 5.4%; and percentage high school graduates 88.9, 91.2, 89.4%.

Development of the regional pain scale (RPS). Initially, we were not sure if nonarticular or nonarticular plus articular

Table 1. Regional pain scale (RPS) item characteristics sorted in descending order of Rasch difficulty.

Region	Fibromyalgia n = 543 (%)	RA n = 9576 (%)*	Mokken H	Mokken Restscore	Rasch Difficulty	OR	Lower 95% CI	Upper 95% CI
Jaw-L	31.0	12.7	0.47	20	1.62	3.08	2.67	3.56
Chest	38.1	12.4	0.54	42	1.58	4.35	3.78	4.99
Jaw-R	31.0	13.3	0.44		1.55	2.93	2.54	3.39
Abdomen	43.0	16.9	0.45	44	1.02	3.71	3.25	4.25
Forearm-L	42.0	20.5	0.58	30	0.68	2.80	2.45	3.20
Forearm-R	44.5	21.7	0.58		0.56	2.90	2.54	3.31
Upper leg-L	51.3	21.6	0.57	20	0.51	3.83	3.35	4.36
Upper leg-R	53.5	22.8	0.56		0.40	3.91	3.42	4.45
Lower leg-L	55.5	25.6	0.55		0.15	3.62	3.17	4.13
Lower leg-R	55.6	26.2	0.55		0.11	3.53	3.10	4.02
Upper arm-L	53.1	27.9	0.55		0.00	2.93	2.57	3.34
Upper arm-R	55.3	29.2	0.55		-0.11	3.00	2.64	3.42
Upper back	73.2	32.9	0.50	19	-0.45	5.56	4.81	6.43
Hip-L	67.4	34.4	0.47	21	-0.55	3.94	3.43	4.52
Hip-R	66.9	35.5	0.47		-0.62	3.68	3.21	4.22
Shoulder-L	70.2	44.7	0.51		-1.23	2.91	2.53	3.35
Shoulder-R	70.7	46.7	0.51		-1.36	2.76	2.40	3.18
Neck	84.0	53.7	0.54		-1.88	4.53	3.81	5.39
Low back	87.3	54.6	0.50		-1.96	5.72	4.73	6.92

Mokken H and Mokken restscore is Mokken H and restscore statistic. Odds ratio for FM versus RA/OA. *Patients with pain in region.

Table 2. Joint and regional pain items not included in the regional pain scale (RPS) following Mokken analysis. Data are percentages of patients with pain in region.

Region	Fibromyalgia (%)	RA (%)
Head	53.7	20.0
Wrist-L	52.5	48.1
Wrist-R	55.9	51.5
Elbow-L	42.0	25.9
Elbow-R	43.9	28.1
Ankle-L	47.9	38.7
Ankle-R	48.9	39.6
Foot arch-L	40.2	24.9
Foot arch-R	42.4	25.6
PIP-L	65.6	59.4
PIP-R	70.6	63.6
MTP-L	46.8	38.3
MTP-R	47.8	39.6
MCP-L	64.5	58.1
MCP-R	67.8	61.1
Heel-L	36.9	21.2
Heel-R	37.8	21.7
Knee-L	71.1	52.6
Knee-R	72.3	54.2

pain sites should be included in the scale. Therefore all joint and regional (or nonarticular) pain variables (Tables 1 and 2) were analyzed with Mokken analysis. Analyses consisted of using the variables at their full range (0–3) and rescored to 0–1, where values of 2 and 3 were recoded to a value of 1. Using all of the variables, an adequately scaled, unidimensional scale based on a total H score of 48 was identified, but departures from monotonicity and general nonintersec-

tion were noted. Iterative runs of the Mokken analysis program⁴⁷ with inspection of output suggested that the best-scaled items within a single dimension were the nonarticular items. Conceptually, we felt that hips and shoulders should be included because they are commonly interpreted by patients as nonarticular regions. In addition, jaw or facial pain is also a part of the nonarticular syndromes. Addition of these items and elimination of articular regions and the feet and head yielded a generally monotonic scale. This scale was compared in its full 0–3 range to a 0/1 range (Table 1). The 0/1 scale was perfectly monotonic. Restscore criteria values, a measure of nonintersection, were all below 80 and only 5 of the 19 scales had values above 40. Restsplit test values were satisfactory. Pmatrix values were uniformly high but, as indicated above, there is a tendency to find higher test crit values even when there are no serious violations of nonintersection. The overall scale reliability was 0.94, and the overall coefficient of scalability (H) was 0.52, representing a strong scale. Scale reliability using the conventional Cronbach's alpha was 0.91.

Full range joint and region scores are conceptually different from 0/1 scales as they include a severity component, while the 0/1 scales record only a location component. We used both measures in subsequent analyses, but only the 0/1 scale was retained in the RPS because the 0/3 model was more difficult to use, had reduced fit and reliability, and did not increase diagnostic accuracy (data not shown).

The 0/1 scales were also studied using Rasch analysis. We considered that the nonparametrically based RPS might not fit the parametric Rasch model, but employed Rasch analysis to document and understand the deviations from

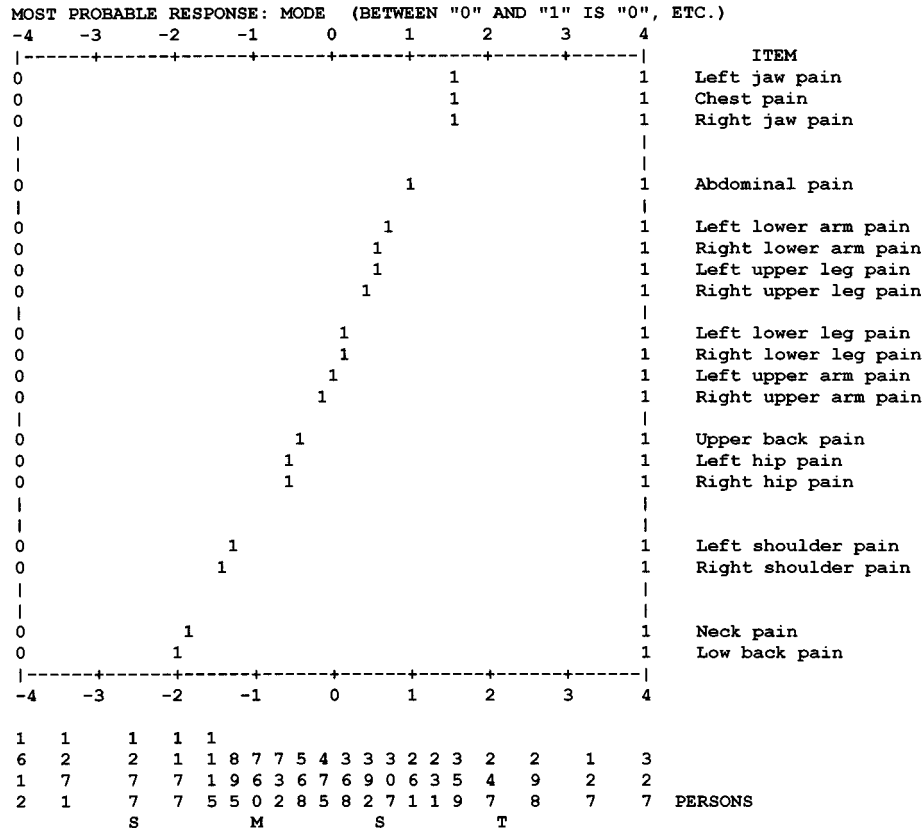


Figure 2. Item and person difficulties on the same logit scale (x axis). The y axis represents individual RPS items. Persons (below x axis) represent the distribution of patients according to their position on the logit difficulty scale. The mean difficulty of the item scale is defined by the 0 logit position. Note that the mean difficulty of patients occurs at -1 logits. The figure shows that low back pain is the least difficult item and left jaw pain is the most difficult item. Overlapping item difficulties are primarily a function of bilateral regions.

that model, and to obtain measures of item difficulty that can only be approximated in the nonparametric Mokken analyses. However, Figure 2 and Table 1 confirm the monotonicity and scalability of the RPS. The easiest or most common items (low back pain, neck pain, shoulder pain are at the bottom of the Figure, and the most difficult items (jaw pain and chest pain) are at the top. Figure 2 also shows the general separation of items by item difficulty, except for the expected agreement of bilateral areas (e.g., left jaw, right jaw).

Although INFIT statistics were satisfactory, OUTFIT statistics of 1.62 to 1.72 were noted for jaw pain, and 1.50 was noted for abdominal pain. These values most likely represent deviations from unidimensionality since the pain sites are not nonarticular sites as are the other scale items. Because of their overall importance in FM and the normal INFIT statistics, these items were retained in the scale. The Rasch model separation statistic was 2.09, indicating an adequate but somewhat restricted range under the Rasch model.

Table 1 also displays the percentage of patients with FM and RA reporting pain in the various regions. The odds

ratios for the identification of FM compared to RA are greatest for the neck, upper back, and lower back regions (as predicted by the ACR FM criteria study²).

Associations of the RPS with clinical variables. To understand how the 17 item RPS (range 0–17) compared with the conventional VAS pain intensity scale, we obtained Pearson correlation coefficients, as shown in Table 3. In general, pain intensity was more strongly correlated with clinical variables than was the sum of the pain locations as measured by the RPS. The combination of both variables, however, explained more variance than was explained by pain intensity alone. For example, the R-square for the regression of the WOMAC function on pain was 0.473, 0.335 for RPS, and 0.536 for RPS and VAS pain together. The stronger correlation of the VAS pain scale was expected, since intensity of pain rather than just presence of pain is more strongly related to clinical variables.

We also found that the RPS was stable over time in the 9582 patients who had a paired survey observation 6 months previously, with a mean RPS difference between surveys of 0.102. As the RPS is correlated with VAS pain severity, we would expect changes in VAS pain to be weakly related to

Table 3. Correlation of regional pain scale (RPS) and VAS pain with clinical variables.

Variable	RPS	VAS Pain
VAS Pain	0.532	1.000
Patient global severity	0.502	0.663
Helplessness	0.482	0.611
Anxiety	0.401	0.387
HAQ	0.512	0.605
Depression	0.397	0.423
Fatigue	0.493	0.613
GI severity scale	0.409	0.413
Sleep scale	0.439	0.524
Quality of life	-0.429	-0.472
Rheumatology distress index	0.541	0.645
WOMAC pain	0.580	0.720
WOMAC stiffness	0.522	0.653
WOMAC function	0.579	0.688
SF-36 physical component score	-0.434	-0.553
SF-36 mental component score	-0.469	-0.487
Lifetime comorbidity	0.316	0.259

changes in RPS. Analysis of change data did indicate that the Pearson correlation of change scores for these variables was 0.191, and the correlation between HAQ change and RPS change was 0.160.

Associations of RPS and other variables with FM. Since widespread pain is a diagnostic criterion for FM classification, we examined it and other variables that are key characteristics of FM to determine their ability to identify FM patients compared to RA/OA patients (Figure 3 and Tables 4–6). Figure 3 shows that the means and general distribution of the RPS score are similar for RA and OA. Therefore in the analyses in Tables 4–6, we combined RA and OA patients. As shown Figure 3, the mean RPS score for FM

was substantially higher (10.9) than the scores for RA and OA (5.6, 6.3).

We also analyzed fatigue, pain, and lifetime comorbidity scores, since analyses of these variables had shown them to be more strongly related to FM versus RA/OA than other variables. Table 4 indicates, as shown by the pseudo R-square, that the RPS and the lifetime comorbidity score were the best predictors of FM, and that the RPS was significantly and substantially better than the VAS pain scale. Patients with a RPS score ≥ 8 had an odds ratio of almost 10 for FM, and being in the 3rd tertile for comorbidity (4–11) was associated with an odds ratio of 11.28. Using comparative logistic regression models, we compared the fit of the continuously scaled RPS and the lifetime comorbidity scores using the Bayesian Information Criterion (BIC')⁷¹. A difference of 26.46 in the BIC' for the 2 models provided very strong support for the better fit of the RPS model compared to the comorbidity model. These data indicate that the RPS is the best self-report measure to distinguish FM from other diagnoses.

Since fatigue, pain, and reported comorbidity (perhaps as a measure of somatization) are central to FM, we used all of the variables together in a classification and regression tree (CART) analysis^{69,70}. In this analysis, using cross-validation, 72.8% of patients were correctly identified, assuming equal numbers in each group. The CART variable importance scores were lifetime comorbidity 100.00, RPS 79.50, fatigue 56.47, and VAS pain 38.82.

The predictive value of this model is quite good, considering that as many as 20% (or more) of RA/OA patients may also have FM⁷², but are not properly classified^{2,73,74}. Further insight into the overlap in diagnosis and symptoms can be seen in Table 5. The RPS is dichotomized at a score of 8 and the VAS fatigue scale is dichotomized at a score of 6, as

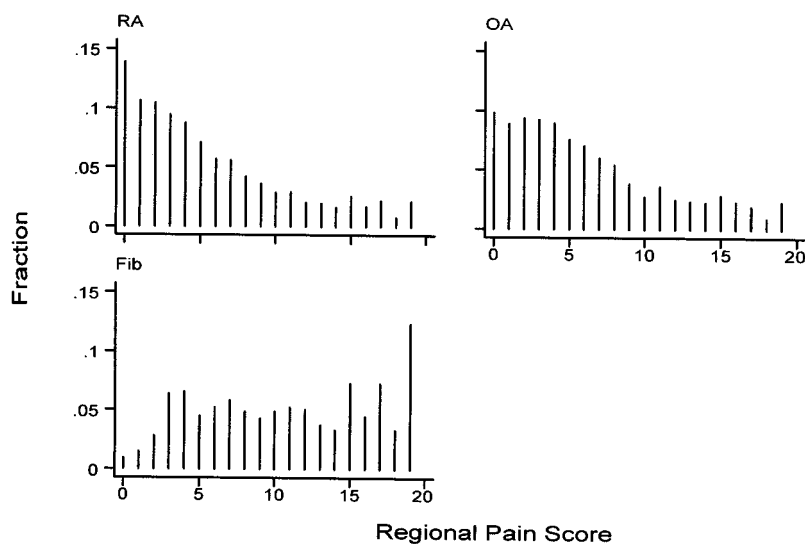


Figure 3. Distribution curves of regional pain score (RPS) by diagnostic grouping. Mean joint and regional pain scores: RA 5.6, OA 6.3, and FM 10.9. Fib: fibromyalgia.

Table 4. Predictive ability of regional pain scale and other variables for the identification of FM among 12,799 patients with RA, OA, and FM.

Variable	OR	Lower 95% CI	Upper 95% CI	Pseudo R-Square
VAS pain				0.05
1st tertile (0–2)	Comparison group			
2nd tertile (2.5–5.5)	2.46	2.01	3.02	
3rd tertile (6–10)	4.97	4.10	6.02	
Lifetime comorbidity score				0.10
1st tertile (0–1)	Comparison group			
2nd tertile (2–3)	3.41	2.73	4.26	
3rd tertile (4–11)	11.28	9.09	13.98	
Fatigue score				0.06
1st tertile (0–2.5)	Comparison group			
2nd tertile (3–6)	1.88	1.52	2.33	
3rd tertile (6.5–10)	5.47	4.53	6.60	
Regional pain scale score				0.10
1st tertile (0–3)	Comparison group			
2nd tertile (4–7)	3.57	2.82	4.51	
3rd tertile (8–19)	9.98	8.09	12.32	

Table 5. Diagnostic classification at levels of the regional pain scale (RPS) and VAS fatigue score.

Category	RA n (%)	OA n (%)	FM n (%)	Total n (%)
RPS < 8				
Fatigue < 6	5078 (53.0)	1410 (52.4)	102 (19.1)	6590 (51.5)
RPS ≥ 8				
Fatigue ≥ 6	1761 (18.4)	386 (14.4)	78 (14.6)	2225 (17.4)
RPS < 6	1104 (11.5)	363 (13.5)	74 (13.9)	1541 (12.0)
RPS ≥ 8				
Fatigue ≥ 6	1633 (17.1)	530 (19.7)	280 (52.4)	2443 (19.1)
Total	9576 (100.0)	2689 (100.0)	534 (100.0)	12799 (100.0)

Table 6. Comparison of RA/OA and FM patients by tertile of regional pain scale.

Variable	RA/OA		FM	
	Mean	SD	Mean	SD
1st tertile				
HAQ	0.68	0.62	0.72	0.60
VAS quality of life	73.50	18.61	69.37	20.78
VAS fatigue	2.99	2.64	4.37	2.78
VAS pain scale	2.53	2.33	3.31	2.63
2nd tertile				
HAQ	1.11	0.65	0.97	0.60
VAS quality of life	64.02	19.42	63.06	19.66
VAS fatigue	4.57	2.73	5.57	2.83
VAS pain scale	4.18	2.51	4.60	2.42
3rd tertile				
HAQ	1.55	0.60	1.39	0.61
VAS quality of life	53.66	20.94	49.38	22.66
VAS fatigue	6.17	2.49	7.16	2.33
VAS pain scale	5.80	2.44	6.43	2.33

suggested by previous analyses³⁰. There is a clustering of low fatigue/low RPS patients in the non-FM group (52–53%) and a clustering of high fatigue/high RPS patients in the FM group (52%). Apparently misclassified patients include RA (17.1%), OA (19.7%), and FM (19.1%). The misclassification of RA and OA patients might reflect estimates of FM occurring in RA and OA patients generally^{2,72-74}.

These observations are also extended by tertile analysis of the RPS in 4 clinical variables in RA/OA and FM patients, as shown in Table 6. In general, FM patients have slightly more severe scores for the clinical variables, except for the HAQ in tertiles 2 and 3. The scores in each tertile are progressively more abnormal. This suggests that the RPS can be used to identify FM-like severity among patients regardless of disease classification.

DISCUSSION

We have developed and provided preliminary validation for a regional pain scale (RPS). It is not only among the most effective tools for identifying FM patients in clinical surveys, but also, in the univariate analyses we performed, was the best overall variable in distinguishing FM from non-FM diagnosed patients. When the RPS was used together with lifetime comorbidity scores and fatigue, it identified 73% of patients correctly (CART), a value that is quite good given the likely presence of undiagnosed (up to 20%) FM in RA and OA. It should be remembered that CART predictions are conservative, given the cross-validation methodology.

The RPS is one approach to the analysis of the extent of body pain. For research purposes it has much to recommend it compared to pain diagrams (Figure 1), which are difficult

and time consuming to score, and are not completed by all patients in the same way. On the other hand, clinicians can get a quick and useful picture of a patient's pain by just looking at a drawing, a method far simpler than scoring a questionnaire such as the RPS. It would appear that there is a place for both methods.

There have been relatively few studies of body pain, perhaps because of the difficulty in scoring. A few studies have addressed body pain in FM (using pain diagrams)^{3,75,76}, and a number of studies have assessed the meaning of pain extent as it applies to diagnosis, function, and psychological status⁷⁷⁻⁸³. Our results are in agreement with previous studies that were done using pain diagrams. As shown in Table 3, the RPS is correlated with major severity and outcome variables across all rheumatic disorders. When used together with a simple VAS pain intensity scale, the combined information is greater than is available with either method separately.

One might reasonably ask, what is the value of having the RPS? First, the RPS might help to develop an alternative definition of FM. For example, the use of the RPS, comorbidity, and fatigue scores correctly classified 73% of patients in this study. As indicated above, however, it is likely that many patients with RA and OA also have FM⁷², and that the correct classification rate is considerably higher than our analyses indicate. Therefore one might make an alternative definition of FM (or FM criteria) based on RPS, level of VAS fatigue, and comorbidity or the number of somatic symptoms reported. Such a definition would allow FM to be identified in survey research, something that is not possible using the current ACR criteria.

An additional use of the scale is to give insight, perhaps with the use of other covariates, into the continuum between perfect health and FM that is clearly present in rheumatic disease patients⁸⁴ and other illnesses^{9,75,85-87}, but is ignored by the ACR criteria². Finally, it offers another validated way to study the effect of the extent of pain and various outcomes. As a simple example, RA patients with a score of 8 or greater on the RPS have an incidence rate ratio (IRR) for the number of physician visits per 6 months that is significantly greater than for those with an RPS score < 8 (1.59, 95% confidence interval 1.53–1.65). In addition, the ability to predict outpatient utilization is better with the RPS than the HAQ by the BIC test statistic. All of this suggests that a measure of extent of pain may be a useful adjunct to current assessment tools regardless of diagnosis.

The RPS was developed in a survey sample. Survey participants are systematically different from clinical sample participants in that survey participants tend to be slightly better educated, older, female, and to have slightly greater clinical severity. These differences tend to be small¹⁸, and survey patients look much like those seen in the clinic¹⁸. Even so, validation of the RPS in a clinical sample is required, and this is particularly so if the RPS is to be used

as an element of an alternative definition of FM. In the latter instance, the RPS should be studied in combination with fatigue and comorbidity scales. Future studies that include the RPS should provide additional data on the worth of this measure in the diagnosis of FM and in FM-like pain states that may occur in illnesses such as RA and OA.

REFERENCES

- Melzack R. The McGill pain questionnaire: major properties and scoring methods. *Pain* 1975;1:277-99.
- Wolfe F, Smythe HA, Yunus MB, et al. The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia: Report of the Multicenter Criteria Committee. *Arthritis Rheum* 1990;33:160-72.
- McBeth J, Macfarlane GJ, Benjamin S, Silman AJ. Features of somatization predict the onset of chronic widespread pain: results of a large population-based study. *Arthritis Rheum* 2001;44:940-6.
- Mikkelsen M, Kaprio J, Salminen JJ, Pulkkinen L, Rose RJ. Widespread pain among 11-year-old Finnish twin pairs. *Arthritis Rheum* 2001;44:481-5.
- Benjamin S, Morris S, McBeth J, Macfarlane GJ, Silman AJ. The association between chronic widespread pain and mental disorder — A population-based study. *Arthritis Rheum* 2000;43:561-7.
- White KP, Harth M, Speechley M, Ostbye T. A general population study of fibromyalgia tender points in noninstitutionalized adults with chronic widespread pain. *J Rheumatol* 2000;27:2677-82.
- Hunt IM, Silman AJ, Benjamin S, McBeth J, Macfarlane GJ. The prevalence and associated features of chronic widespread pain in the community using the 'Manchester' definition of chronic widespread pain. *Rheumatology* 1999;38:275-9.
- Macfarlane GJ. Generalized pain, fibromyalgia and regional pain: an epidemiological view. *Best Pract Res Clin Rheumatol* 1999;13:403-14.
- Macfarlane GJ, Morris S, Hunt IM, et al. Chronic widespread pain in the community: The influence of psychological symptoms and mental disorder on healthcare seeking behavior. *J Rheumatol* 1999;26:413-9.
- McBeth J, Silman AJ. Unraveling the association between chronic widespread pain and psychological distress: An epidemiological approach. *J Psychosom Res* 1999;47:109-14.
- Mikkelsen M, Sourander A, Salminen JJ, Kautiainen H, Piha J. Widespread pain and neck pain in schoolchildren. A prospective one-year follow-up study. *Acta Paediatr* 1999;88:1119-24.
- White KP, Harth M. The occurrence and impact of generalized pain. *Best Pract Res Clin Rheumatol* 1999;13:379-89.
- Ruiz MR, Munoz AM, Perula dT, Aguayo GM. Biopsychosocial features of patients with widespread chronic musculoskeletal pain in family medicine clinics. *Fam Pract* 1997;14:242-8.
- Andersson HI, Ejlertsson G, Leden I, Rosenberg C. Characteristics of subjects with chronic pain, in relation to local and widespread pain report — A prospective study of symptoms, clinical findings and blood tests in subgroups of a geographically defined population. *Scand J Rheumatol* 1996;25:146-54.
- Jacobsson LT, Nagi DK, Pillemer SR, et al. Low prevalences of chronic widespread pain and shoulder disorders among the Pima Indians. *J Rheumatol* 1996;23:907-9.
- Macfarlane GJ, Thomas E, Papageorgiou AC, Schollum J, Croft PR, Silman AJ. The natural history of chronic pain in the community: a better prognosis than in the clinic? *J Rheumatol* 1996;23:1617-20.
- Macfarlane GJ, Croft PR, Schollum J, Silman AJ. Widespread pain: is an improved classification possible? *J Rheumatol* 1996; 23:1628-32.

18. Wolfe F, Ross K, Anderson J, Russell IJ, Hebert L. The prevalence and characteristics of fibromyalgia in the general population. *Arthritis Rheum* 1995;38:19-28.
19. Russell TS, Percy JS. The presence of chronic widespread pain in the general population. *J Rheumatol* 1994;21:579-80.
20. Croft P, Rigby AS, Boswell R, Schollum J, Silman AJ. The prevalence of widespread pain in the general population. *J Rheumatol* 1993;20:710-3.
21. Taal E, Abdel Nasser AM, Rasker JJ, Wiegman O. A self-report Thompson articular index: What does it measure? *Clin Rheumatol* 1998;17:125-9.
22. Stewart MW, Palmer DG, Knight RG, Highton J. A self-report articular index — relationship to variations in mood and disease activity measures. *Br J Rheumatol* 1993;32:631-2.
23. Mason JH, Anderson JJ, Meenan RF, Haralson KM, Lewis-Stevens D, Kaine JL. The rapid assessment of disease activity in rheumatology (RADAR) questionnaire — validity and sensitivity to change of a patient self-report measure of joint count and clinical status. *Arthritis Rheum* 1992;35:156-62.
24. Calvo FA, Calvo A, Berrocal A, et al. Self-administered joint counts in rheumatoid arthritis: Comparison with standard joint counts. *J Rheumatol* 1999;26:536-9.
25. Alarcon GS, Tilley BC, Li SH, Fowler SE, Pillemer SR. Self-administered joint counts and standard joint counts in the assessment of rheumatoid arthritis. *J Rheumatol* 1999;26:1065-7.
26. Franssen J, Hauselmann H, Michel BA, Caravatti M, Stucki G. Responsiveness of the self-assessed rheumatoid arthritis disease activity index to a flare of disease activity. *Arthritis Rheum* 2001;44:53-60.
27. Stucki G, Liang MH, Stucki S, Bruhlmann P, Michel BA. A self-administered rheumatoid arthritis disease activity index (RADAI) for epidemiologic research: Psychometric properties and correlation with parameters of disease activity. *Arthritis Rheum* 1995;38:795-8.
28. Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. *J Rheumatol* 1982;9:789-93.
29. Fries JF, Spitz PW, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23:137-45.
30. Wolfe F, Hawley DJ, Wilson K. The prevalence and meaning of fatigue in rheumatic disease. *J Rheumatol* 1996;23:1407-17.
31. Meenan RF, Gertman PM, Mason JH, Dunaif R. The Arthritis Impact Measurement Scales. *Arthritis Rheum* 1982;25:1048-53.
32. Meenan RF. The AIMS approach to health status measurement: conceptual background and measurement properties. *J Rheumatol* 1982;9:785-8.
33. Hawley DJ, Wolfe F. Depression is not more common in rheumatoid arthritis: a 10 year longitudinal study of 6608 rheumatic disease patients. *J Rheumatol* 1993;20:2025-31.
34. Hawley DJ, Wolfe F. Anxiety and depression in patients with RA: A prospective study of 400 patients. *J Rheumatol* 1988;15:932-41.
35. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). 1. Conceptual framework and item selection. *Med Care* 1992;30:473-83.
36. Wolfe F. Determinants of WOMAC function, pain and stiffness scores: evidence for the role of low back pain, symptom counts, fatigue and depression in osteoarthritis, rheumatoid arthritis and fibromyalgia. *Rheumatology* 1999;38:355-61.
37. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;15:1833-40.
38. Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: Validity, responsiveness and reliability of EuroQol (EQ-5D). *Br J Rheumatol* 1997;36:551-9.
39. Wolfe F, Hawley DJ. Measurement of the quality of life in rheumatic disorders using the EuroQol. *Br J Rheumatol* 1997;36:786-93.
40. Nunnally JC, Bernstein IH. *Psychometric theory*. 3rd ed. New York: McGraw-Hill; 1994.
41. Tennant A, Ryser L, Stucki G, Wolfe F. An 8-item international version of the HAQ: The Inter-HAQ. 1999. Presented at the 1999 meeting of the European League Against Rheumatism (EULAR).
42. Tennant A, Hillman M, Fear J, Pickering A, Chamberlain MA. Are we making the most of the Stanford Health Assessment Questionnaire? *Br J Rheumatol* 1996;35:574-8.
43. Wolfe F, Hawley DJ, Goldenberg DL, Russell IJ, Buskila D, Neumann L. The assessment of functional impairment in fibromyalgia (FM): Rasch analyses of 5 functional scales and the development of the FM Health Assessment Questionnaire. *J Rheumatol* 2000;27:1989-99.
44. Mchorney CA, Haley SM, Ware JE Jr. Evaluation of the MOS SF-36 Physical Functioning Scale (PF-10): II. Comparison of relative precision using Likert and Rasch scoring methods. *J Clin Epidemiol* 1997;50:451-61.
45. Haley SM, Mchorney CA, Ware JE Jr. Evaluation of the MOS SF-36 physical functioning scale (PF-10): I. Unidimensionality and reproducibility of the Rasch item scale. *J Clin Epidemiol* 1994;47:671-84.
46. Wolfe F. Which HAQ is best? A comparison of the HAQ, MHAQ and RA-HAQ, a difficult 8 item HAQ (DHAQ), and a rescored 20 item HAQ (HAQ20): analyses in 2491 rheumatoid arthritis patients following leflunomide initiation. *J Rheumatol* 2001;28:982-9.
47. Molenaar IW, Sijtsma K. *MSP5 for Windows: A program for Mokken Scale Analysis for polytomous items*. Groningen: ProGamma, 2000.
48. Bond TG, Fox CM. *Applying the Rasch model: fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum Associates; 2001.
49. Hanson EK, Schaufeli W, Vrijkotte T, Plomp NH, Godaert GL. The validity and reliability of the Dutch Effort-Reward imbalance questionnaire. *J Occup Health Psychol* 2000;5:142-55.
50. Schmitz N, Hartkamp N, Kiuse J, Franke GH, Reister G, Tress W. The Symptom Check-List-90-R (SCL-90-R): a German validation study. *Qual Life Res* 2000;9:185-93.
51. Ringdal K, Ringdal GI, Kaasa S, et al. Assessing the consistency of psychometric properties of the HRQoL scales within the EORTC QLQ-C30 across populations by means of the Mokken Scaling Model. *Qual Life Res* 1999;8:25-43.
52. Roorda LD, Roebroeck ME, Lankhorst GJ, van Tilburg T, Bouter LM. Measuring functional limitations in rising and sitting down: development of a questionnaire. *Arch Phys Med Rehabil* 1996;77:663-9.
53. Suurmeijer TP, Doeglas DM, Moum T, et al. The Groningen activity restriction scale for measuring disability: its utility in international comparisons. *Am J Public Health* 1994;84:1270-3.
54. DeJong A, Molenaar IW. An application of Mokken's model for stochastic, cumulative scaling in psychiatric research. *J Psychiatr Res* 1987;21:137-49.
55. Andrich D. *Rasch models for measurement*. Quantitative applications in the social sciences. Newbury Park, NJ: Sage Publications; 1988:1-94.
56. Nordenskiöld U. Daily activities in women with rheumatoid arthritis. Aspects of patient education, assistive devices and methods for disability and impairment assessment. *Scand J Rehabil Med* 1997;37 Suppl:1-72.
57. Stucki G, Daltroy L, Katz JN, Johannesson M, Liang MH. Interpretation of change scores in ordinal clinical scales and health status measures: the whole may not equal the sum of the parts.

- J Clin Epidemiol 1996;49:711-7.
58. Tesio L, Granger CV, Fiedler RC. A unidimensional pain/disability measure for low-back pain syndromes. *Pain* 1997;69:269-78.
 59. Grimby G, Andren E, Holmgren E, Wright B, Linacre JM, Sundh V. Structure of a combination of Functional Independence Measure and Instrumental Activity Measure items in community-living persons: a study of individuals with cerebral palsy and spina bifida. *Arch Phys Med Rehabil* 1996;77:1109-14.
 60. Nordenskiöld U, Grimby G, Hedberg M, Wright B, Linacre JM. The structure of an instrument for assessing the effects of assistive devices and altered working methods in women with rheumatoid arthritis. *Arthritis Care Res* 1996;9:358-67.
 61. Chang WC, Chan C. Rasch analysis for outcomes measures: some methodological considerations. *Arch Phys Med Rehabil* 1995;76:934-9.
 62. Granger CV, Ottenbacher KJ, Baker JG, Sehgal A. Reliability of a brief outpatient functional outcome assessment measure. *Am J Phys Med Rehabil* 1995;74:469-75.
 63. Silverstein F, Kilgore DJ, Harman JG, Harvey T. Applying psychometric criteria to functional assessment in medical rehabilitation, I: exploring unidimensionality. *Arch Phys Med Rehabil* 1991;72:631-7.
 64. A user's guide to BIGSTEPS: Rasch model computer program. Chicago: Mesa Press; 1997.
 65. McNamara T. Measuring second language performance. London: Longman; 1996.
 66. Fischer GH, Molenaar IW. Rasch models: foundations, recent developments, and applications. New York: Springer-Verlag; 1995.
 67. Wright BD, Masters GN. Rating scale analysis: Rasch measurement. Chicago: Mesa Press; 1982.
 68. Stata Corp. Stata statistical software: Release 7.0. College Station, TX: Stata Corp.; 2001.
 69. Steinberg D, Colla P. CART: Tree-structured non-parametric data analysis. San Diego, CA: Salford Systems; 1999.
 70. Breiman L, Friedman J, Olshen R, Stone C. Classification and regression trees. Pacific Grove: Wadsworth; 1984.
 71. Raftery AE. Bayesian model selection in social research. In: Marsden PV, editor. *Sociological methodology*. Oxford: Basil Blackwell; 2000:111-63.
 72. Wolfe F, Cathey MA. Prevalence of primary and secondary fibrositis. *J Rheumatol* 1983;10:965-8.
 73. Buskila D, Langevitz P, Gladman DD, Urowitz S, Smythe HA. Patients with rheumatoid arthritis are more tender than those with psoriatic arthritis. *J Rheumatol* 1992;19:1115-9.
 74. Wolfe F, Cathey MA, Kleinheksel SM. Fibrositis (fibromyalgia) in rheumatoid arthritis. *J Rheumatol* 1984;11:814-8.
 75. McBeth J, Macfarlane GJ, Hunt IM, Silman AJ. Risk factors for persistent chronic widespread pain: a community-based study. *Rheumatology* 2001;40:95-101.
 76. Wigers SH, Skrondal A, Finset A, Gotestam KG. Measuring change in fibromyalgic pain: The relevance of pain distribution. *J Musculoskel Pain* 1997;5:29-41.
 77. Bjorksten MG, Talback M. A follow-up study of psychosocial factors and musculoskeletal problems among unskilled female workers with monotonous work. *Eur J Public Health* 2001; 11:102-8.
 78. Dahl B, Gehrchen PM, Kiaer T, Blyme P, Tondevold E, Bendix T. Nonorganic pain drawings are associated with low psychological scores on the preoperative SF-36 questionnaire in patients with chronic low back pain. *Eur Spine J* 2001;10:211-4.
 79. Ohnmeiss DD, Vanharanta H, Estlander AM, Jansen A. The relationship of disability (Oswestry) and pain drawings to functional testing. *Eur Spine J* 2000;9:208-12.
 80. Lindal E, Bergmann S, Thorlacius S, Stefansson JG. The localization of pain in chronic fatigue syndrome on a pain drawing according to grid areas. *Percept Mot Skills* 1996;83:508-10.
 81. Toomey TC, Mann JD, Abashian S, Thompson-Pope S. Relationship of pain drawing scores to ratings of pain description and function. *Clin J Pain* 1991;7:269-74.
 82. Tait RC, Chibnall JT, Margolis RB. Pain extent: Relations with psychological state, pain severity, pain history, and disability. *Pain* 1990;41:295-301.
 83. Vlaeyen JW, Geurts SM, Kole-Snijders AM, Schuerman JA, Groenman NH, van Eek H. What do chronic pain patients think of their pain? Towards a pain cognition questionnaire. *Br J Clin Psychol* 1990;29:383-94.
 84. Wolfe F. The relation between tender points and fibromyalgia symptom variables: evidence that fibromyalgia is not a discrete disorder in the clinic. *Ann Rheum Dis* 1997;6:268-71.
 85. McBeth J, MacFarlane GJ, Hunt IM, Silman AJ. Risk factors for persistent chronic widespread pain: a community-based study. *Rheumatology* 2001;40:95-101.
 86. Croft P, Burt J, Schollum J, Thomas E, Macfarlane G, Silman A. More pain, more tender points: is fibromyalgia just one end of a continuous spectrum? *Ann Rheum Dis* 1996;55:482-5.
 87. Croft P, Schollum J, Silman A. Population study of tender point counts and pain as evidence of fibromyalgia. *BMJ* 1994;309:696-9.