

# Assessment of Pain in Rheumatoid Arthritis: Minimal Clinically Significant Difference, Predictors, and the Effect of Anti-Tumor Necrosis Factor Therapy

FREDERICK WOLFE and KALEB MICHAUD

**ABSTRACT.** *Objective.* To compare a visual analog pain scale (VAS) with the Medical Outcomes Study Short Form-36 Health Survey (SF-36) bodily pain; to define the minimal clinically important change (MCIC) for pain in observational studies; to define clinically useful cutpoints for pain; to quantify the predictors of pain; and to estimate the effect of anti-tumor necrosis factor (TNF) therapy on pain.

*Methods.* Over 6 years we studied 12,090 patients with rheumatoid arthritis (RA). Pain was assessed by VAS and SF-36 pain scales.

*Results.* Compared with the SF-36 scale, the 0–10 VAS pain scale was better correlated with all clinical variables. The mean VAS score was 3.4 (standard deviation 2.8), and the best cutpoint for an “acceptable” level of pain was  $\leq 2.0$ . The MCIC for pain was approximately 0.5 units by one measure and 1.1 by another. Pain increased slightly with the duration of RA, 0.03 (95% confidence interval 0.02–0.03) and decreased with age, 0.01 (95% CI 0.01–1.02) units per year. Pain was greater in ethnic minorities [0.78 (95% CI 0.63–0.93)] and women [0.31 (95% CI 0.23–0.39)] and was lower in college graduates [–0.88 (95% CI –1.00 to –0.76)]. Self-reported joint and nonarticular pain at 16 bilateral sites explained 44% of VAS pain scores. Anti-TNF therapy reduced pain by 0.59 to 0.53 units and EuroQol utility by 0.02 (95% CI 0.02–0.02) units.

*Conclusion.* Anti-TNF therapy improved pain by 0.53 to 0.70 units. The MCIC for improvement and worsening of pain is about 0.5 to 1.1 units. Pain levels are almost constant over RA duration, and are increased in women, ethnic minorities, smokers, and those with less education. (First Release July 1 2007; J Rheumatol 2007;34:1674–83)

## Key Indexing Terms:

PAIN

MINIMAL CLINICALLY IMPORTANT CHANGE

ANTI-TUMOR NECROSIS FACTOR

RHEUMATOID ARTHRITIS

Pain and function are the key clinical variables in rheumatoid arthritis (RA). Pain assessment tools have been many, including multi-item and multidimensional tools<sup>1–10</sup> as well as a simple visual analog scale (VAS)<sup>11–14</sup>. In RA clinical trials, observational studies, and clinical practice, however, pain assessment most often comes down to 2 questionnaires in almost all settings, the VAS and the bodily pain subscale that is part of the Medical Outcomes Study Short Form-36 Health Survey (SF-36) assessment<sup>7</sup>.

Randomized clinical trials (RCT) and observational studies and clinical practice differ in a number of ways with respect

to pain. RCT select patients because of severity and therefore select for high levels of pain. The RCT is usually performed to compare the effect of treatment on outcomes such as pain.

For observational studies and clinical practice, there are different questions, such as, what causes pain? how much pain is present? why do people differ with respect to pain? and how effective are treatments in relieving pain? Although these seem to be simple questions, they are not easy ones because pain may be related to genetic differences and biologic, psychosocial, and demographic factors<sup>15–27</sup>, as well as to illness, expectations, and treatment. When so many factors work together to influence pain, it is difficult to disentangle covariate effects and to establish a direction of causality.

Here, we use a large data bank to quantitatively answer questions about pain in RA that should be useful to clinicians and researchers. Specifically, we measure the effect of ethnicity, sex, education, age, and duration of RA on pain, keeping in mind the direction of causality. We also describe the extent to which specific joints contribute to pain scores, and we describe the minimal clinically important difference (MCID) for pain improvement and worsening, as well as levels of pain that appear to be acceptable. Finally, we compare pain as measured by the SF-36 and VAS as to their clinical relevance.

From the National Data Bank for Rheumatic Diseases, Wichita, Kansas, USA.

Supported by a grant from Bristol-Myers-Squibb.

F. Wolfe, MD, National Data Bank for Rheumatic Diseases and University of Kansas School of Medicine; K. Michaud, MS, National Data Bank for Rheumatic Diseases and University of Nebraska, Omaha, Nebraska, USA.

Address reprint requests to Dr. F. Wolfe, National Data Bank for Rheumatic Diseases, Arthritis Research Center Foundation, 1035 N. Emporia, Suite 230, Wichita, KS 67214.

E-mail: fwolfe@arthritis-research.org

Accepted for publication April 3, 2007.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2007. All rights reserved.

## MATERIALS AND METHODS

Patients in our study were participants in the National Data Bank for Rheumatic Diseases (NDB) longitudinal study of RA outcomes. NDB participants are recruited on an ongoing basis from the practices of United States rheumatologists, and are followed prospectively with semiannual, detailed, 28-page questionnaires, as described<sup>28-31</sup>. The diagnosis of RA was made by the patient's rheumatologist, and there are no exclusions for enrollment. Approximately 8% of patients discontinue participation per year. We studied 12,090 patients with RA who had completed at least one semiannual questionnaire between January 1, 1999, and December 30, 2004, and who were not participants in a safety registry. Safety registry patients were excluded because their recruitment methods may select for a subset of patients with more severe RA than is ordinarily found in RA clinical practice, and their inclusion might bias the study toward patients with more severe RA. For the cross-sectional analyses in this report (Tables 1, 2, 4, Figures 1-4), we used a random number generator to select a single questionnaire from the longitudinal data of each patient in the event a patient had completed more than one survey.

At each assessment we recorded demographic variables including age, sex, ethnic origin, marital status, smoking status, household income, and all treatments. Patients reported functional status using the Health Assessment Questionnaire (HAQ)<sup>32,33</sup> and the original SF-36 function scale<sup>7</sup>. We also determined pain, global disease severity, and fatigue by VAS<sup>34</sup>. The VAS pain scale measures 21 points, from 0 to 10 at 0.5 unit intervals. The line was not marked, and patients were unable to see their previous responses. The second

Table 1. Demographic and clinical characteristics of 12,090 RA study participants. Values are from a randomly selected observation.

Variable	Mean	SD
Age, yrs	60.09	13.61
Sex, % male	23.17	
Ethnic origin		
Caucasian, not of Hispanic origin, %	90.58	
African American, not of Hispanic origin, %	4.12	
Asian or Pacific Islander, %	1.00	
American Indian or Alaskan Native, %	0.96	
Hispanic, %	2.83	
Other, %	0.51	
Education category, %		
0-8 yrs	2.76	
8-11 yrs	7.75	
12 yrs	35.76	
13-15 yrs	26.26	
≥ 16	27.46	
Household income, \$US	45038	29073
Disease duration, yrs	14.27	11.12
Symptom intensity scale (0-9.8)	3.61	2.36
VAS Pain (0-10)	3.84	2.78
Pain total (SF-36)	6.18	1.68
HAQ (0-3)	1.04	0.74
VAS Global severity (0-10)	3.49	2.52
Regional Pain Score (0-19)	5.65	5.06
Patient joint count (0-16)	7.41	4.65
Patient joint score (0-48)	12.10	9.44
VAS QOL (linear transformation) (0-1)	0.78	0.09
EuroQol (US) (0-1)	0.73	0.19
SF-6D utility (0-1)	0.62	0.10

HAQ: Health Assessment Questionnaire disability index; HAQ-II: Health Assessment Questionnaire disability index-II; Global: patient global severity; VAS Qol: linearly transformed VAS QOL utility scale; EuroQol: EuroQol utility score; SF-6D: SF-6D utility score.

Table 2. Correlation of pain related variables. The number of observations was 12,090 except for EuroQol (N = 4,170) and depression and anxiety (N = 9,600) that was not administered during all question assessments.

Variable	VAS Pain	SF-36 Pain
Clinical variables		
VAS pain	1	0.58
Symptom intensity scale	0.69	0.49
VAS global severity	0.68	0.48
VAS fatigue	0.64	0.42
HAQ	0.62	0.44
Patient joint score	0.61	0.48
Pain (SF-36)	0.58	1
Satisfaction with health	0.58	0.45
Physical function (SF-36)	-0.58	-0.44
Regional pain score	0.54	0.43
Vitality (SF-36)	-0.53	-0.43
Patient joint count	0.51	0.42
Depression	0.46	0.35
Mood	0.44	0.35
Anxiety	0.42	0.34
Mental health (SF-36)	-0.39	-0.32
Quality of life utilities		
EuroQol	-0.66	-0.55
VAS QOL (linear transformation)	-0.51	-0.40
SF-6D utility	-0.48	-0.34

pain scale was the bodily pain subscale from the SF-36<sup>7</sup> and is derived from one 6-item question (pain intensity) and one 5-item question (interference with work because of pain). The SF-36 scales inquire about status over the previous 4 weeks, the other clinical scales about status over the last week.

To assess quality of life (QOL) by utilities, we administered the EuroQol<sup>35-38</sup>, utilizing the newly developed US tariffs<sup>39</sup>, the SF-6D<sup>40</sup>, and a VAS-based QOL scale utilizing linear transformation<sup>41</sup>. The EuroQol contains 5 questions, 3 of which are about function, one about pain, and one about psychological status. The VAS QOL scale was anchored at one end with "death" and the other end with "perfect health." It was transformed using the algorithm,  $0.44 \times (\text{VAS QOL}/100) + 0.49$ <sup>41</sup>. VAS QOL scales have higher utilities than multi-item scales such as the EuroQol, and are closer in that respect to values obtained by the standard gamble<sup>41</sup>.

A self-report joint count and joint score was obtained using the rheumatoid arthritis disease activity index scales<sup>42</sup>, and the joint regions from this scale were also evaluated separately as individual scores and counts. The joint score allows designations of no pain, mild, moderate, and severe pain, and is coded as 0 through 3. The joint count compresses this scale to 0 and 1 by counting all scores greater than 0 as 1. The Regional Pain Scale (RPS) is a self-report count of nonarticular regions<sup>43,44</sup>. The Symptom Intensity (SI) scale is derived from 2 separate scales, a VAS for fatigue<sup>45</sup> and the Regional Pain Scale (RPS)<sup>44</sup>. The SI scale uses these 2 measures in continuous form according to the following formula:  $[\text{VAS fatigue} + (\text{RPS}/2)]/2$ . This yields a scale with a 0 to 9.75 range.

Satisfaction with health was evaluated with a 5-point scale<sup>46</sup>. The categories were very dissatisfied, somewhat dissatisfied, neither satisfied nor dissatisfied, somewhat satisfied, and very satisfied. Minimal improvement was defined as 1 category change in health satisfaction.

The mood scale we used represents the normalized Arthritis Impact Measurement Scales (AIMS) anxiety and depression scales if available<sup>47</sup>; otherwise, it represents the SF-36 mental health subscale<sup>7</sup>. Both scales are transformed to a 0-10 scale, with higher values representing greater mood abnormality.

*Statistical methods.* The graphs of Figure 2 were generated by running line

Table 3. Association of demographic factors with pain intensity. Analysis by generalized estimating equations (n = 12,090, observations = 64,090).

Variable	Multivariable Analyses			Univariable Analyses		
	Coefficient	p	95% CI	Coefficient	p	95% CI
Sex Female	0.41	< 0.001	0.31, 0.50	0.48	< 0.001	0.38, 0.58
Education category, yrs						
0-8	0.80	< 0.001	0.53, 1.08	0.90	< 0.001	0.62, 1.17
8-11	0.70	< 0.001	0.53, 0.87	0.81	< 0.001	0.64, 0.97
12 (referent)	1.00			1.00		
13-15	-0.16	0.003	-0.26, -0.05	-0.16	0.003	-0.26, -0.05
≥ 16	-0.74	< 0.001	-0.84, -0.64	-0.81	< 0.001	-0.91, -0.70
Non-Hispanic Caucasian	0.78	< 0.001	0.63, 0.93	0.94	< 0.001	0.79, 1.10
Married	-0.24	< 0.001	-0.31, -0.16	-0.35	< 0.001	-0.42, -0.27
Current smoker	0.31	< 0.001	0.22, 0.40	0.40	< 0.001	0.31, 0.49
Age, yrs	-0.01	< 0.001	-0.01, -0.00	-0.01	< 0.001	-0.02, -0.01

smooths of VAS pain on duration and age simultaneously, adjusting for sex<sup>48</sup>. Using a simple type of backfitting, the resulting smoother is a locally linear function of the predictors for each observation.

To determine the minimal clinically significant change (MCIC), we estimated the mean change in pain over 6 months that was associated with a 1 category change in health satisfaction at each of the 21 levels of baseline pain (Figure 4), adjusting for pain changes at baseline levels that were associated with no change in health status. We also used linear regression to estimate the change in pain that was associated with a 0.22 unit change in HAQ over a 6-month period. A 0.22 change in HAQ is thought to be the MCIC for HAQ. We used this method to corroborate the previously described method of MCIC estimation. The limitations of this method are discussed below. Finally, we estimated the MCIC based on the standard error of the mean (SEM), using the formula:

$$SEM = SD * \sqrt{1 - \text{reliability}}$$

and a reliability of 0.85<sup>49,50</sup>. The SEM is consistent with the MCIC across most studies<sup>51-53</sup>:

$$SEM = SD * \sqrt{1 - \alpha}$$

The univariable and multivariable analyses of Tables 3 and 5 were performed using all observations from study subjects and generalized estimating

equations (GEE). The GEE analyses specified an exchangeable correlation structure and robust standard errors.

Correlation coefficients were calculated by the Pearson method. Data were analyzed using Stata (Stata Corp., College Station, TX, USA) version 9.1.

## RESULTS

**Demographic and clinical characteristics.** The demographic and clinical characteristics of the 12,090 patients with RA in this study are shown in Table 1. The SF-36 pain scores have been reversed and rescaled to make them comparable to the VAS scale, where 0 is no pain and 10 is maximum pain. The level of pain in this cohort was 3.84 (SD 2.78) using the VAS scale and 6.18 (SD 1.68) using the SF-36 pain scale. The scores for the SF-36 are based on questions about pain intensity (6 categories) and interference with work because of pain (5 categories). By contrast, the VAS scale is purely a pain intensity scale. The distributions of VAS pain scores and SF-36 pain scores are shown in Figure 1. In addition to the sparseness of the SF-36 scale, its distribution differs from that of the VAS scale in being relatively normally distributed compared with the VAS scale, which is skewed to the left. We also iden-

Table 4. Levels of pain by measures of health status.

Quintile/Category	By Quintiles of Pain, mean (IQR)	By Quintiles of QOL, mean (IQR)*	By Categories of Health Satisfaction, mean (IQR)	By Categories of Health Satisfaction, mean (IQR) (≤ 1 comorbid condition)
1st, very satisfied	0.6 (0.0, 1.0)	1.5 (0.5, 2.0)	1.5 (0.5, 2.0)	1.3 (0.0, 1.5)
2nd, somewhat satisfied	1.7 (1.5, 2.0)	2.6 (1.0, 3.5)	3.0 (1.0, 4.5)	2.7 (1.0, 4.0)
3rd, intermediate	3.1 (2.5, 3.5)	3.6 (1.5, 5.5)	4.3 (2.0, 6.0)	3.8 (2.0, 5.5)
4th somewhat dissatisfied	5.5 (5.0, 6.0)	4.5 (2.5, 6.5)	5.3 (3.5, 7)	5.0 (3.0, 7.0)
5th, very dissatisfied	8.0 (7.7, 8.5)	5.6 (3.5, 7.5)	6.9 (5.5, 8.5)	6.5 (5.0, 8.5)

\* QOL was assessed by VAS QOL assessment. IQR: Interquartile range; QOL: Quality of life.

Table 5. Effect of anti-TNF therapy on pain intensity scores. Adjusted for age, sex, education level, minority status, marital status, baseline Health Assessment Questionnaire, baseline pain, and baseline prednisone use.

Treatment	Coefficient	Z	p	95% CI
No MTX, infliximab, etanercept	1.00			
Infliximab, no MTX, no etanercept	-0.53	-4.73	< 0.001	-0.74 to -0.31
Etanercept, no MTX	-0.62	-7.50	< 0.001	-0.78 to -0.46
Infliximab, MTX, no etanercept	-0.58	-8.76	< 0.001	-0.71 to -0.45
Etanercept, MTX, no infliximab	-0.70	-9.80	< 0.001	-0.81 to -0.56

MTX: methotrexate.

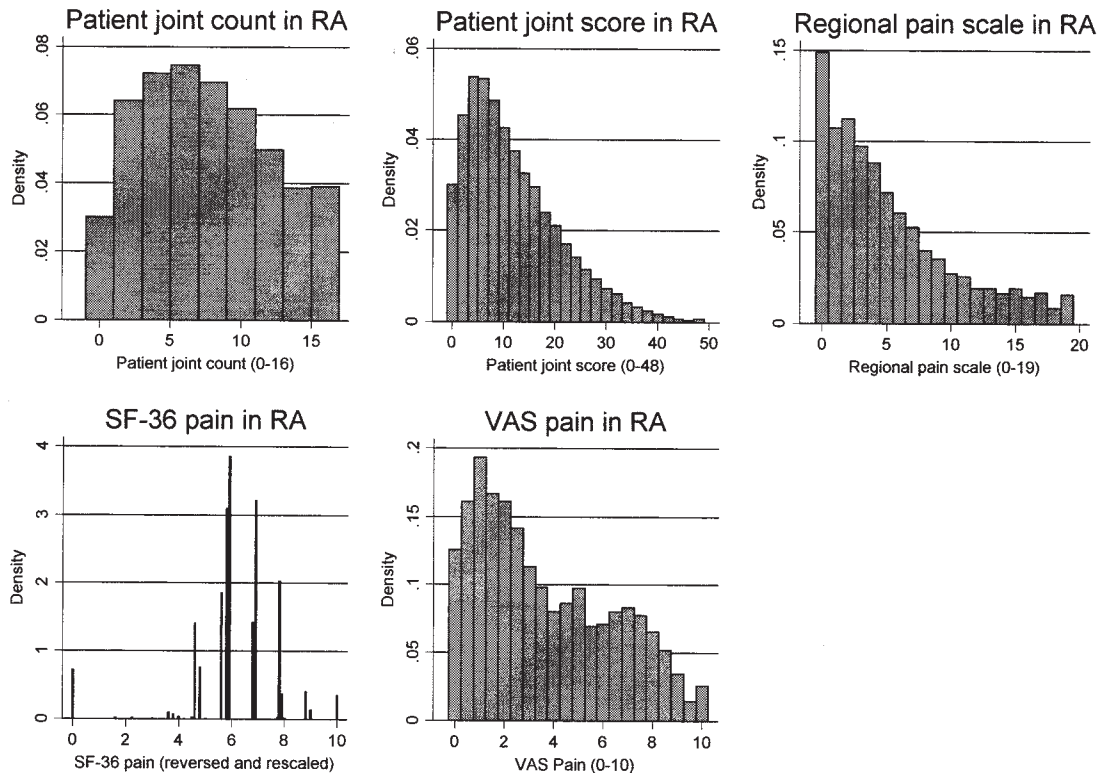


Figure 1. The distributional characteristics of 5 measures of pain in 12,090 patients with RA.

tified several other pain related scales, the patient self-reported joint count, a 0–16 scale; the patient joint score, a 0–48 scale; and the regional pain score, a 0–19 scale (Table 1).

*Association of pain with clinical variables and QOL.* To place pain, generally, and the 2 scales, specifically, into perspective, we examined the degree to which the pain scales were associated with clinical measures and with QOL utility measures (Table 2). For every clinical and QOL measure, the VAS pain scale was more strongly correlated with the clinical or utility measure than was the SF-36 pain score, and we use it alone for all analyses that follow.

The strongest correlate of pain is the Symptom Intensity scale, which is a composite measure of fatigue and nonarticular pain. This is followed by global, fatigue, HAQ, and the patient joint score, a variable that expresses the extent of joint

pain and its severity. The VAS pain scale is also more strongly correlated with other VAS scales than with SF-36 measures of pain and function. The weakest associations are with the psychological scales. The VAS utility scale is correlated with the EuroQol at 0.57. The EuroQol is correlated with VAS pain at -0.66. It is difficult to judge the association of pain with utility scales because the EuroQol and SF-6D contain pain questions. However, when the VAS QOL scale is studied, its correlation with pain is -0.51, indicating that QOL is concerned with factors in addition to pain. Not shown in the table of clinical association, but of interest, is that VAS pain and the SF-36 bodily pain were correlated with household income at r values of -0.23 and -0.16, respectively.

*Association of pain with demographic characteristics.* Age, duration of RA, and pain are associated (Figure 2). However,

## The relation of age and RA duration to pain intensity

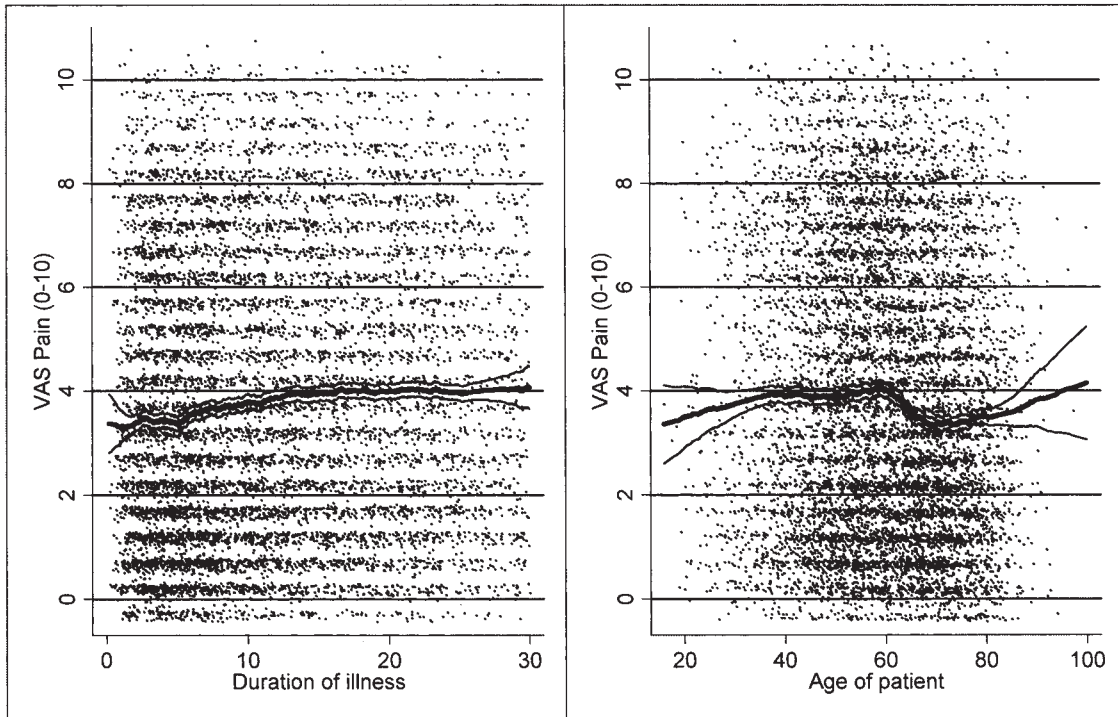


Figure 2. Cross-sectional estimation of the relationship between pain and RA duration, adjusted for age and sex (left panel) and pain and age, adjusted for duration of RA and sex (right panel). The association with pain is very weak. Adjusted for duration, age, and sex,  $R^2 = 0.016$ .

## Prevalence and effect of joint and regional pain

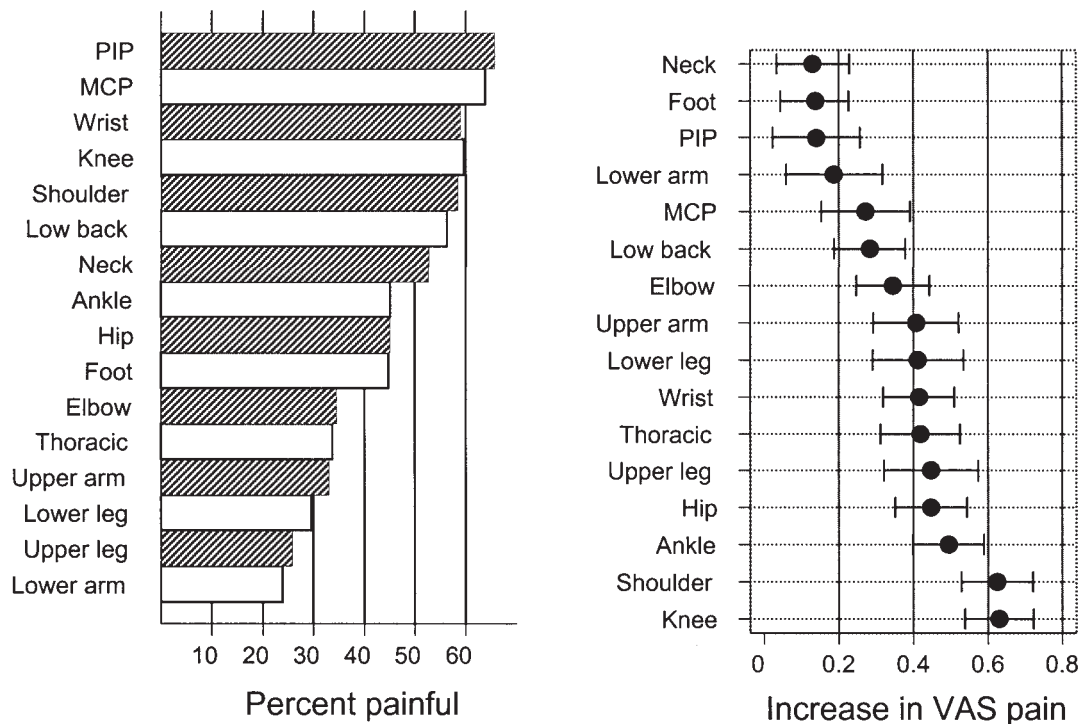


Figure 3. *Left panel.* The percentage of joints and regions with pain. Pain is present if either the right or left side is reported as being painful, except for neck, thoracic, and low back regions, which are not bilateral regions. *Right panel.* Multivariable increase in VAS pain associated with pain in joints or body regions.

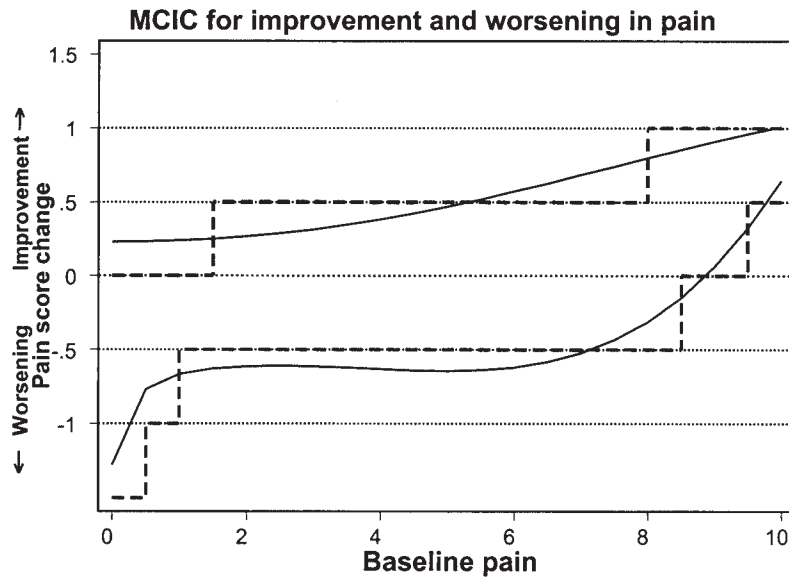


Figure 4. Minimal clinically important change (MCIC) in pain based on baseline pain scores. The MCIC for improvement (solid line) is shown above the zero line. The broken line represents rounding of the solid line to the nearest 0.5. Lines below the zero line represent MCIC for worsening. Over most of the range of pain, the MCIC for improvement or worsening is closest to 0.5.

this association is weak and the increase in pain as a function of RA duration is very small, 0.03 (95% CI 0.02 to 0.3) units per year. By contrast, pain falls slightly, beginning at about age 62 years. This small alteration is noted during the years when retirement from work activities ordinarily occurs. The separate figures allow the relationship of pain with age or duration to be seen, disentangling age, duration, and sex.

We studied the association of selected demographic factors and pain intensity in longitudinal analyses, as shown in Table 3. In multivariable analyses (all variables analyzed simultaneously), women had pain levels that were 0.41 (95% CI 0.31–0.50) units higher than those in men. Similarly, current smokers had pain levels that were 0.31 (95% CI 0.22–0.40) units greater than those in nonsmokers. Ethnic minorities had higher levels compared with non-Hispanic Caucasians [0.78 (95% CI 0.63–0.93)] and married persons had lower levels [–0.24 (95% CI –0.31 to –0.16)] compared with unmarried persons. Finally, in comparison to persons with a high school level education, more education was associated with lower pain scores and less education with higher scores. Notably, college graduates had pain scores that were 0.74 (95% CI 0.71–0.91) units lower than those in high school graduates. Despite the significant difference in pain according to demographic characteristics, the overall effect of demographics on pain was small, and the R-square of a regression model of pain on the demographic variables was 0.05.

*The relation of joint and regional pain to overall pain intensity.* In RA, inflammation of peripheral joints, muscular groups, and the axial skeleton represents the driving force for the pain intensity measured by the VAS scale. In our surveys of pain,

patients describe which joints and regions are painful (Figure 3, left). As these are self-reported painful regions, it is not possible to accurately separate shoulder girdle pain from shoulder joint pain and buttock region pain from hip joint pain. Areas that are painful in more than 50% of patients include proximal interphalangeal joints, metacarpophalangeal joints, wrist, knee, shoulder, and lower back.

To understand the extent to which painful regions were related to VAS pain intensity, we performed a multivariable regression analysis of VAS pain on the joints and regions in Figure 3, right. The multivariate increases in VAS pain associated with the presence of pain in each joint or region are shown in the figure. The largest contributions to the VAS pain score came from shoulder, knee, upper leg, ankle, and hip pain. The R-square for this model was 0.33. When individual joint pain severity (joint score) was incorporated into the model (none, mild, moderate, severe) instead of simply a joint count, the R-square increased to 0.44. Addition of age, sex, ethnicity, marital status, education, and duration increased the R-square to 0.45.

*How much change in pain is clinically significant? The MCIC (or MCID).* Pain is significantly correlated with health satisfaction,  $r = 0.58$ , and with HAQ,  $r = 0.62$  (Table 2). To ascertain the MCIC for pain at a level of patient importance, we determined the change in pain score that accompanied a minimal change in patient satisfaction with health. Satisfaction categories include very dissatisfied, somewhat dissatisfied, neither satisfied nor dissatisfied, somewhat satisfied, and very satisfied. Minimal improvement was defined as 1 category change in health satisfaction. As the extent of change in pain

scores varies as a function of baseline satisfaction, we determined the MCIC at each level of baseline pain. As shown in Figure 4, the MCIC for improvement or worsening is approximately 0.5 across most levels of baseline pain. We also estimated MCIC a second way, by regressing changes in pain on changes in HAQ, and then estimating the pain change at a HAQ score of 0.22, which is considered the MCIC for HAQ. By this method, the average MCIC was 0.41 (95% CI 0.31–0.50) for improvement and 0.54 (95% CI 0.45–0.63) for worsening, in general agreement with the data from Figure 3. However, the estimated MCIC for HAQ of 0.22 does not consider MCIC as a function of starting values. We found that the ratio of pain changes to HAQ changes increases by approximately 0.28 (95% CI 0.16–0.39) per baseline pain unit. Therefore, this measure of MCIC cannot be used to extrapolate beyond the mean HAQ and pain values. Finally, we calculated the MCIC by estimating the standard error of the mean for the VAS pain scale, and noted it to be 1.1.

We also attempted to define “acceptable” levels of pain (Table 4). To put pain scores in perspective, columns 2 and 3 show pain scores according to quintiles of pain and QOL. By overall health satisfaction, 75% of patients in the “very satisfied with their health” category had levels of pain < 2.0 (interquartile range 0.5 to 2.0), and these values were in agreement with the first quintile of the VAS QOL scale. When the analysis was restricted to those with 0 or 1 comorbid conditions, in order to eliminate the contribution of excess comorbidity to health satisfaction, the 75th percentile declined to 1.5. In general these data suggest that the goal of therapy should aim at reduction of pain scores to  $\leq 2.0$ .

*The relation of treatment variables to pain intensity.* To estimate the effect of treatment, we excluded all patients who were enrolled in the NDB as part of treatment registries, as those registries were started immediately after the release of the therapies and selected for more severe RA. In addition, we did not analyze patients who were taking anti-TNF therapy at their first observation in the NDB. For the remaining 10,180 patients, who each had an average of 5.6 observations, we obtained baseline scores for HAQ and VAS pain, and baseline use of prednisone. We used those data as well as age, sex, education, minority status, and marital status as covariates in a cross-sectional time series analysis (GEE) to test whether pain scores were lower in anti-TNF treated patients after controlling for baseline covariates. Because anti-TNF therapy may be given with methotrexate (MTX), we formed dummy variables of the important treatment groups (Table 4). Therefore the comparison for the treatment groups is patients not taking TNF or MTX, adjusted for the model covariates. Adalimumab was infrequently taken and therefore was not included in the groupings.

Table 5 shows that the effect of anti-TNF therapy was to reduce pain by at least 0.53 units. Overall, anti-TNF therapy in this model reduced pain by 0.51 (95% CI 0.43–0.59 units). There were no statistically significant differences between any

combinations of the 4 anti-TNF groups. Specifically, the difference in pain scores for the 2 etanercept groups compared with the 2 infliximab groups was 0.22 (95% CI –0.12 to 0.54;  $p = 0.202$ ). QOL differences in patients receiving anti-TNF therapy were 0.02 (95% CI 0.02–0.02) for the EuroQol and 0.01 (95% CI 0.01–0.01) for the transformed VAS QOL scale.

## DISCUSSION

Pain is the dominant symptom in RA. Pain is also a complex experience. Clinically, it has been described in terms of sensory, affective, and evaluative components<sup>6</sup>, and it can be measured in all of these domains, quantified in terms of intensity, or measured in many other complex ways. In RA, there have only been 2 pain questionnaires that have achieved widespread, indeed almost universal use, the VAS pain intensity scale and the SF-36 bodily pain scale. The SF-36 pain scale is based on a 6-category pain intensity and 5-category pain interference scale. Because the SF-36 is the most widely used assessment tool in all medical disorders, including RA, the pain scale is always available in clinical trials and in many observational studies. The data from our report, however, suggest that the SF-36 pain scale is not nearly as well correlated with clinical measures as is the VAS scale. We therefore performed the bulk of the analyses in this report using only the VAS pain scale.

The interpretation of pain scores is primarily related to the understanding of the direction of causal pathways. We found that pain intensity increases only slightly with duration of RA, an amount equal to 0.3 units every 10 years, or an essentially negligible increase. The behavior of age is somewhat different. Pain reaches its maximum between ages 50 and 62 years, after which it falls from about 4.1 to 3.4 by age 67. This may be related to changes in physical activity associated with retirement. However, there are not enough participants to test the possibility. Overall, however, we can conclude that neither age nor duration alters clinical pain intensity meaningfully.

However, certain demographic factors do affect pain along its causal pathway. Pain scores are greater in women, minorities, and those with less education. The causal pathway is less secure for smoking, marital status, and household income, all of which are associated with varying levels of pain. But for these latter variables, the direction of the causal pathway is not clear, and is likely to be bidirectional. For example, people with higher levels of pain might be more likely to continue smoking, be divorced, or earn less income. Studies that try to explain pain levels on the basis of variables like these, and by psychological variables, become mired in the problem of causality.

In RA, however, a theoretical model adds to the understanding of pain. In this instance, inflammation in peripheral joints causes local pain and that pain is reflected in the VAS pain intensity scale. In the models of this study, peripheral joint and regional pain explain 44% of the variance in the VAS pain scale scores, which a sizable amount for self-report vari-

ables. The addition of demographic variables, which explain 5% of the variance by themselves, adds only 1% to the joint and regional pain model explanation in a combined model. Therefore, if 45% of pain variance is explained by sex, education, ethnicity, smoking, and involved joints and regional pain, and 5% by the demographic factors alone, then, on average, 11% of pain score variance is explainable by demographic factors. The effect of demographic features can be seen most clearly in the extreme case: a male, Caucasian, nonsmoking college graduate will have, on average, a pain score that is 2.66 (95% CI -2.90 to -2.41) units lower than the score of a female, ethnic minority, smoking non-college graduate. While such cases are rare (7.2% in the current data set), the direction of the 11% variance is observable in clinical practice.

We also tried to understand what an “acceptable” level of pain is. In this instance we judged acceptable as the level occurring in patients who were “very satisfied” with their health. Table 4 suggests that 75% of patients with a “very satisfied” classification will have a pain score of 2.0 or less. If we restrict the classification to include only those with one or fewer comorbid conditions, the 75th percentile falls to 1.5. Considering all of the data in Table 4, we suggest that a level of 2.0 or less seems to be an appropriate cutpoint. This level has been suggested by others on the basis of consensus.

Current interest in MCIC derives from the desire to interpret clinical trial results. Empirically, the American College of Rheumatology improvement criteria take as the minimal level of improvement a 20% change in pain<sup>54</sup>. However, this change level is based on consensus rather than patients’ preferences. Clinically important changes are those that are important to the patient. MCIC, however, is a function of the starting pain level (Figure 4), whether the change is toward improvement or toward worsening, whether the pain is acute or chronic, and whether the setting is a clinical trial, a clinical practice setting, or an observational study, among other factors<sup>55</sup>. In this report we approached the MCIC in 2 ways. First, we determined the level of pain that was associated with a one-unit change in health status. As can be seen in the upper curved line of Figure 4, MCIC is lowest at low baseline levels of pain and highest at high baseline levels. As a practical compromise, we rounded the MCIC to the nearest 0.5. That resulted in an MCIC for improvement of 0.5. The MCIC for worsening also works out to about 0.5. However, an asymmetry between improvement and worsening that has been noted by others<sup>56</sup> can be seen in the upper and lower curved lines. On average, the MCIC represents a change of approximately 13%. We also estimated the MCIC to be 1.1 by determining the SEM. The SEM is close to a half a standard deviation change, which is close to the average MCIC in many studies<sup>52</sup>.

MCIC for pain improvement and worsening was determined by interview by Wells, *et al*<sup>56</sup>, and was 6% for improvement and 16% for worsening. In an osteoarthritis clinical trial, the minimal clinically important improvement for pain was found to be between -1.9 units (-40.8%) and -1.5 units

(-32.0%) for pain, using a 5-point Likert scale at the final visit as a method of determining pain at the minimal response level<sup>57</sup>. In the setting of acute pain, 2 emergency room studies found a change of 1.3 units to be the minimal change in acute pain that was clinically significant<sup>58,59</sup>. Lee, *et al* also found in an emergency room setting that a 3.0-unit change indicated adequate pain relief<sup>60</sup>. Because clinical trials may have inflated responses, we advise that the MCIC be interpreted by subtracting the comparison group response. However, it is extremely important to understand that MCIC from RCT and observational studies are likely to be quite different, as responsiveness in RCT is considerably greater than in observational studies owing to the nature of RCT design.

Finally, we were able to apply the MCIC information to the result of anti-TNF therapy. Overall, the average reduction in pain of 0.51 units indicated that this therapy satisfied the MCIC criterion developed from NDB data in the setting of clinical practice.

In our study we used a simple horizontal VAS. However, different types of VAS pain scales can yield different results, depending on the marking, anchors, and orientation of the scale. Paul-Dauphin and colleagues studied simple, middle-marked, graphic rating, graduated, graduated-numbered, and numerical rating scales in horizontal and vertical orientations<sup>61</sup>. They noted that scale characteristics influenced the proportion of zero and low values, although not mean scores. In addition, these differences influenced cross-sectional precision, but not the precision of measurement of change over time. VAS pain can also differ depending on whether patients are allowed to see their previous scores. In our study, patients did not see their previous scores.

The VAS pain scale is superior to the SF-36 scale. The MCIC for pain is approximately 0.5 to 1.1 in observational studies. The level of pain does not change importantly as a function of age or RA duration. Joint and regional pain sites, together with demographic characteristics, explain 45% of pain variance, and demographics explain 5%, or 11% of explainable variance. “Acceptable” levels of pain are approximately  $\leq 2.0$ .

## REFERENCES

1. Gloth FM III, Scheve AA, Stober CV, Chow S, Prosser J. The Functional Pain Scale: reliability, validity, and responsiveness in an elderly population. *J Am Med Dir Assoc* 2001;2:110-4.
2. Jensen MP, Turner LR, Turner JA, Romano JM. The use of multiple-item scales for pain intensity measurement in chronic pain patients. *Pain* 1996;67:35-40.
3. Vlaeyen JW, Pernot DF, Kole-Snijders AM, Schuerman JA, van Eek H, Groenman NH. Assessment of the components of observed chronic pain behavior: the Checklist for Interpersonal Pain Behavior (CHIP). *Pain* 1990;43:337-47.
4. Turk DC, Rudy TE. Towards a comprehensive assessment of chronic pain patients. *Behav Res Ther* 1987;25:237-49.
5. Kerns RD, Turk DC, Rudy TE. The West Haven-Yale Multidimensional Pain Inventory (WHYMPI). *Pain* 1985;24:345-56.
6. Melzack R. The McGill pain questionnaire: major properties and



- scoring methods. *Pain* 1975;1:277-99.
7. Stewart AL, Hays RD, Ware JE. The MOS Short-form General Health Survey. *Med Care* 1988;26:724-35.
  8. McCracken LM, Zayfert C, Gross RT. The Pain Anxiety Symptoms Scale — Development and validation of a scale to measure fear of pain. *Pain* 1992;50:67-73.
  9. Callahan LF, Brooks RH, Summey JA, Pincus T. Quantitative pain assessment for routine care of rheumatoid arthritis patients, using a pain scale based on activities of daily living and a visual analog pain scale. *Arthritis Rheum* 1987;30:630-6.
  10. Meenan RF, Gertman PM, Mason JH. Measuring health status in arthritis: the Arthritis Impact Measurement Scales. *Arthritis Rheum* 1980;23:146-52.
  11. Huskisson EC. Measurement of pain. *Lancet* 1974;2:1127-31.
  12. Joos E, Peretz A, Beguin S, Famaey JP. Reliability and reproducibility of visual analogue scale and numeric rating scale for therapeutic evaluation of pain in rheumatic patients. *J Rheumatol* 1991;18:1269-70.
  13. Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. *Pain* 1983;16:87-101.
  14. Dixon JS, Bird HA. Reproducibility along a 10 cm vertical visual analogue scale. *Ann Rheum Dis* 1981;40:87-9.
  15. Brown SC, Glass JM, Park DC. The relationship of pain and depression to cognitive function in rheumatoid arthritis patients. *Pain* 2002;96:279-84.
  16. Moldofsky H. Sleep and pain. *Sleep Med Rev* 2001;5:385-96.
  17. Lefebvre JC, Keefe FJ, Affleck G, et al. The relationship of arthritis self-efficacy to daily pain, daily mood, and daily pain coping in rheumatoid arthritis patients. *Pain* 1999;80:425-35.
  18. Fifield J, Reisine S, Sheehan TJ, Mcquillan J. Gender, paid work, and symptoms of emotional distress in rheumatoid arthritis patients. *Arthritis Rheum* 1996;39:427-35.
  19. Bradley LA. Pain measurement in arthritis. *Arthritis Care Res* 1993;6:178-86.
  20. Anderson KO, Keefe FJ, Bradley LA, et al. Prediction of pain behavior and functional status of rheumatoid arthritis patients using medical status and psychological variables. *Pain* 1988;33:25-32.
  21. Hawley DJ, Wolfe F. Anxiety and depression in patients with rheumatoid arthritis: a prospective study of 400 patients. *J Rheumatol* 1988;15:932-41.
  22. Parker J, Frank R, Beck N, et al. Pain in rheumatoid arthritis: relationship to demographic, medical, and psychological factors. *J Rheumatol* 1988;15:433-7.
  23. Skevington SM. Psychological aspects of pain in rheumatoid arthritis: a review. *Soc Sci Med* 1986;23:567-75.
  24. Anderson KO, Bradley LA, Young LD, McDaniel LK, Wise CM. Rheumatoid arthritis: review of psychological factors related to etiology, effects, and treatment. *Psychol Bull* 1985;98:358-87.
  25. Bradley LA. Psychological aspects of arthritis. *Bull Rheum Dis* 1985;35:1-12.
  26. DeRoos AJ, Callahan LF. Differences by sex in correlates of work status in rheumatoid arthritis patients. *Arthritis Care Res* 1999;12:381-91.
  27. Meana M. The meeting of pain and depression: Comorbidity in women. *Can J Psychiatry* 1998;43:893-9.
  28. Wolfe F, Michaud K, Choi HK, Williams R. Household income and earnings loss among 6,396 persons with rheumatoid arthritis. *J Rheumatol* 2005;32:1875-83.
  29. Wolfe F, Michaud K. A brief introduction to the National Data Bank for Rheumatic Diseases. *Clin Exp Rheumatol* 2005;23:S168-71.
  30. Wolfe F, Michaud K. Heart failure in rheumatoid arthritis: rates, predictors, and the effect of anti-tumor necrosis factor therapy. *Am J Med* 2004;116:305-11.
  31. Wolfe F, Michaud K. Lymphoma in rheumatoid arthritis: the effect of methotrexate and anti-tumor necrosis factor therapy in 18,572 patients. *Arthritis Rheum* 2004;50:1740-51.
  32. 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.
  33. Fries JF, Spitz PW, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23:137-45.
  34. Wolfe F, Hawley DJ, Wilson K. The prevalence and meaning of fatigue in rheumatic disease. *J Rheumatol* 1996;23:1407-17.
  35. Picavet HS, Hoeymans N. Health related quality of life in multiple musculoskeletal diseases: SF-36 and EQ-5D in the DMC3 study. *Ann Rheum Dis* 2004;63:723-9.
  36. 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.
  37. Wolfe F, Hawley DJ. Measurement of the quality of life in rheumatic disorders using the EuroQol. *Br J Rheumatol* 1997;36:786-93.
  38. EuroQol Group. EuroQol — a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199-208.
  39. Shaw JW, Johnson JA, Coons SJ. U.S. Valuation of the EQ-5D Health States: Development and testing of the D1 valuation. *Med Care* 2005;3:203-20.
  40. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;21:271-92.
  41. Mrus JM, Yi MS, Freedberg KA, et al. Utilities derived from visual analog scale scores in patients with HIV/AIDS. *Med Decis Making* 2003;23:414-21.
  42. 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.
  43. Wolfe F, Michaud K. Severe rheumatoid arthritis (RA), worse outcomes, comorbid illness, and sociodemographic disadvantage characterize RA patients with fibromyalgia. *J Rheumatol* 2004;31:695-700.
  44. Wolfe F. Pain extent and diagnosis: development and validation of the regional pain scale in 12,799 patients with rheumatic disease. *J Rheumatol* 2003;30:369-78.
  45. Wolfe F. Fatigue assessments in rheumatoid arthritis: comparative performance of visual analog scales and longer fatigue questionnaires in 7760 patients. *J Rheumatol* 2004;31:1896-902.
  46. Wolfe F, Michaud K, Strand V. Expanding the definition of clinical differences: from minimally clinically important differences to really important differences. Analyses in 8931 patients with rheumatoid arthritis. *J Rheumatol* 2005;32:583-9.
  47. Meenan RF, Gertman PM, Mason JH, Dunaif R. The Arthritis Impact Measurement Scales. *Arthritis Rheum* 1982;25:1048-53.
  48. Royston P, Cox NJ. A multivariable scatterplot smoother. *The Stata Journal* 2005;5:405-12.
  49. Wolfe F, Michaud K, Li T. Sleep disturbance in patients with rheumatoid arthritis: evaluation by medical outcomes study and visual analog sleep scales. *J Rheumatol* 2006;33:1942-51.
  50. Daltroy LH, Cats-Baril WL, Katz JN, Fossel AH, Liang MH. The North American Spine Society Lumbar Spine Outcome Assessment Instrument: reliability and validity tests. *Spine* 1996;21:741-9.
  51. Wyrwich KW. Minimal important difference thresholds and the standard error of measurement: is there a connection? *J Biopharm Stat* 2004;14:97-110.
  52. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a

- standard deviation. *Med Care* 2003;41:582-92.
53. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol* 1999;52:861-73.
  54. Felson DT, Anderson JJ, Boers M, et al. American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;38:727-35.
  55. Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. *Curr Opin Rheumatol* 2002;14:109-14.
  56. Wells GA, Tugwell P, Kraag GR, Baker PR, Groh J, Redelmeier DA. Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. *J Rheumatol* 1993;20:557-60.
  57. Tubach F, Ravaud P, Baron G, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. *Ann Rheum Dis* 2005;64:29-33.
  58. Gallagher EJ, Liebman M, Bijur PE. Prospective validation of clinically important changes in pain severity measured on a visual analog scale. *Ann Emerg Med* 2001;38:633-8.
  59. Todd KH, Funk KG, Funk JP, Bonacci R. Clinical significance of reported changes in pain severity. *Ann Emerg Med* 1996;27:485-9.
  60. Lee JS, Hobden E, Stiell IG, Wells GA. Clinically important change in the visual analog scale after adequate pain control. *Acad Emerg Med* 2003;10:1128-30.
  61. Paul-Dauphin A, Guillemin F, Virion JM, Briancon S. Bias and precision in visual analogue scales: a randomized controlled trial. *Am J Epidemiol* 1999;150:1117-27.