Patient Self-Administered Joint Tenderness Counts in Rheumatoid Arthritis Are Reliable and Responsive to Changes in Disease Activity

FERNANDO FIGUEROA, YOLANDA BRAUN-MOSCOVICI, DINESH KHANNA, ECHING VOON, LY GALLARDO, DONNA LUINSTRA, XIOMARA PINA, GRETCHEN HENSTORF, SANDY LAURENCE, RICHARD NEIMAN, and DANIEL FURST

ABSTRACT. ***Objective.*** To examine whether self-assessment of tender and swollen joints by patients with rheumatoid arthritis (RA) can be used to evaluate changes in disease activity instead of joint counts by physicians.

Methods. Eighty-two patients with RA taking part in controlled studies were recruited for investigation. The patient's self-assessment of joint tenderness and swelling was completed both before and 30 minutes after examination by a physician. Examinations of tender and swollen joints by a rheumatologist were performed at baseline and 3 months later. The correlations and verification of agreement of these clinical assessments were analyzed.

Results. Within-patient and patient-physician correlations for joint tenderness counts were high (r = 0.96 and 0.78, respectively). Patient-physician correlation for joint swelling counts was still significant, although much lower (r = 0.34). Patients' and physicians' estimations of the change in disease activity over 3 months did not differ (p > 0.76 for all comparisons).

Conclusion. Joint tenderness counts were consistent when comparing intra-patient and patient–physician assessments, while joint swelling counts were poorly correlated. Patient and physician assessments of change over 3 months were parallel and similar for joint tenderness count. Self-administered tender joint counts might be a useful tool to evaluate the response to therapy in RA. (First Release Nov 1 2006; J Rheumatol 2007;34:54–6)

Key Indexing Terms:
RHEUMATOID ARTHRITIS DISEASE ACTIVITY SELF-ASSESSMENT
JOINT COUNT OUTCOME MEASURES

There has been considerable debate concerning disease activity assessments and their usefulness in predicting outcomes in rheumatoid arthritis (RA)[1-3]. Although the physician-performed joint evaluation is regarded as the gold standard in the assessment of patients with RA, there is increasing interest in patient self-reported measures[4]. The purpose of this study was to examine whether the patient's self-assessment of tender and swollen joint counts correlates with the physician's evaluation, and whether it reliably estimates the change in disease activity over time.

From the Department of Medicine, Universidad de los Andes, Santiago de Chile, Chile; Department of Rheumatology, Rambam Medical Center, Technion, Haifa, Israel; Division of Immunology, Section of Rheumatology, University of Cincinnati, Cincinnati, Ohio, USA; and Department of Rheumatology, University of California at Los Angeles, Los Angeles, California, USA.

F.E. Figueroa, Professor of Medicine, Facultad de Medicina, Universidad de los Andes; Y. Braun-Moscovici, MD, B. Rappaport Faculty of Medicine, Rambam Medical Center; D. Khanna, MD, Assistant Professor of Medicine, University of Cincinnati; E. Voon, Physician Assistant, University of Washington Medical Center; L. Gallardo, BS; D. Luinstra, BS, MBA; X. Pina, BA; G. Henstorf, BA, CCRC, Fred Hutchinson Cancer Research Center, Seattle, Washington; S. Laurence, CMA; R. Neiman, MD; D.E. Furst, MD, Professor of Rheumatology, University of California at Los Angeles.

Address reprint requests to Dr. D.E. Furst, Department of Rheumatology, University of California at Los Angeles, Room 32-59, 1000 Veteran Ave., Los Angeles, CA 90095-1670. E-mail: defurst@mednet.ucla.edu

Accepted for publication August 17, 2006.

MATERIALS AND METHODS

Patients and study design. Eighty-two patients with RA fulfilling the 1987 American College of Rheumatology criteria were recruited; they were already enrolled in controlled studies of an anti-CD4 agent and an anti-tumor necrosis factor agent that were ultimately found not to be effective. After signing fully informed, voluntary consent, patients completed a self-administered joint count form (SAJ). The following joints were evaluated bilaterally for tenderness and swelling (1 = present; 0 = absent): temporomandibular, shoulder, sternal, elbow, wrist, each metacarpophalangeal and proximal interphalangeal, hip (tenderness only), knee, ankle, tarsus, metatarsophalangeal (MTP; as a unit), and toes (as a unit). The maximum count was 42. The SAJ was completed prior to the physician's examination, and 30–60 minutes later, with patients uninformed of the initial score. Three trained physicians repeatedly assessed their own patients, performing a similar joint count that included acromioclavicular joints and individual scoring of MTP. The maximum count was 50.

Methods and analysis. With an anticipated moderate effect size and a physician's joint tenderness count coefficient of variation of 0.2, 64 patients were required to achieve an alpha-2 of 0.05 and a power of 0.8. Eighty-two patients were enrolled. The SAJ counts for tenderness and swelling were compared to the physician's joint counts at the same visit. We purposely did not train the

patients in order to more closely approximate what might (and probably would) occur in clinical practice. The differences between patient and physician joint counts were corrected by multiplying the patient count by 1.19 to adjust for the differences between the patient's and physician's maximum count. Spearman and intraclass correlation coefficients were calculated. Agreement between patient and physician joint counts and test-retest reliability of patient self-reported joint counts were computed. Signed-ranks tests were done to examine potential differences between measurements. Sensitivity to change was tested by measuring the differences in amount of paired changes for tender and swollen joint counts, assuming linearity.

RESULTS

Of the 82 patients enrolled (age 54.1 ± 12 yrs, disease duration 10.4 ± 7.5 yrs, 82% female), 79 completed at least one set of evaluations, and 61 completed the 3-month followup visit (26% dropout). Rheumatoid factor was positive in 70% and 67% used prednisone (mean dose 6.2 ± 2.9 mg/day). Data regarding treatment were unavailable to us but this did not influence our study, which was aimed at comparing 2 methods of assessing disease activity and was not concerned with drug efficacy. A large decrease in mean joint tenderness count (JTC) and a smaller decrease in joint swelling count (JSwC) were seen in both physician and patient assessments (data not shown). The correlation coefficient between the patient's and the physician's JTC was 0.78 ($p < 0.0007$) at baseline, explaining 60.8% of the variance. A high correlation was also found between the patient pre- and post-visit JTC ($r = 0.9$) and JSwC ($r = 0.96$, explaining 92% of the variance). The correlation between physician and patient JSwC was weak, although the p value was still significant ($r = 0.34$, $p = 0.018$, explaining only 11.5% of the variance). The intraclass correlation coefficients followed the same behavior as the Spearman correlation coefficients (Table 1). Agreement among patient and physician assessments is shown in Table 2. Patient's pre- and post-visit JTC as well as the physician's and the patient's pre-visit JTC were similar, although a significant difference was detected between the physician's and the patient's pre-visit JSwC ($p < 0.0001$). The standard deviation of the differences was high in all cases, rendering the prediction of one variable unreliable when using the other (Table 2). However, the changes in paired patient and physician JTC were significantly correlated, with no statistically significant differences between the patients' and the physicians' evaluations (Table 3).

DISCUSSION

The use of self-administered questionnaires to assess disease indicators has been increasing in recent years[4]. Such questionnaires differ with regard to actual measurements (tenderness, swelling, limitation of motion, etc.), joints assessed, and the grading of abnormalities by joint size or severity[5]. Self-administered forms exhibit adequate reliability, reproducibility and construct validity[6-10], but the correlation with the assessors' joint examination — considered the gold standard — is only moderate[4,6,7,11]. We examined whether the patient's self-assessment of joint counts can be used instead of the physician's joint count to judge changes in RA disease activity. JTC were reliable with respect to intrapatient variability and patient–physician agreement, but JSwC of patients and physicians correlated poorly, although still remaining close to the intrarater reliability of JSwC reported for physicians in RA (intraclass correlation coefficient 0.47)[12]. While the mean difference between the physician and the patient JTC are small, the standard deviation around those differences is high (Table 2), indicating that it would be impossible to predict the patient JTC from the physician's JTC, or vice versa. However, an interesting finding was that changes in JTC and JSwC from baseline to Month 3 were reliably estimated by the patient, and were not statistically different from the physician's estimation (Table 3). Although patient JTC are not predictive of investigators' counts per se, they could be used to measure onset of effect, as patients could repeat these counts at home on a frequent basis. The usefulness of the patient JTC is supported by the finding that the degree of change in this measurement was not different from the investigator's. On the other hand, assuming that the investigator's ability to measure swollen joints represents the "gold standard"[10,11], patients should not be asked to measure joint swelling, since those correlations were quite low.

Since these studies showed no treatment effects, we were unable to determine sensitivity to change; this is a weakness of the study that must be left for future investigations.

Regression analysis showed a relationship between physicians' swollen joint counts and patient tender joint counts (coefficient 0.116, $p = 0.006$), but no direct relationships of physician tender joint counts with patient tender joint counts. This may indicate that the swollen counts of patients and

Table 1. Correlations among clinical assessments.

		Number of Paired Observations	Spearman Correlation Coefficient	Correlated p*	ICC
Joint Tender Count	Physician and patient (pre-visit)	79	0.78	< 0.0007	0.77
	Patient pre-and post-visit	66	0.90	< 0.0007	0.90
Joint Swelling Count	Physician and patient (pre-visit)	79	0.34	0.018	0.43
	Patient pre-and post-visit	66	0.96	< 0.0007	0.89

* Corrected for repeated testing (multiplied by 7). ICC: intraclass correlation coefficient.

Table 2. Agreement among clinical assessments.

Differences at Baseline	N	Mean Difference (SD)	95% CI	p (signed-rank)
Patient pre-and post-visit joint tenderness count	66	0.14 (5.2)	–1.12 to 1.40	0.68
Physician and patient pre-visit joint tenderness count	79	1.02 (7.3)	–0.60 to 2.64	0.27
Patient pre-and post-visit joint swelling count	66	0.38 (4.0)	–0.58 to 1.34	0.22
Physician and patient pre-visit joint swelling count	79	8.7 (8.6)	6.8 to 10.6	< 0.0001

Table 3. Changes during 3 months and correlation among clinical assessments.

	Differences (3 mo-baseline)	p (difference compared to 0)	Spearman correlation coefficient*	p (correlation)	p** (difference)
Joint Tenderness Counts					
Physician	–4.2 (8.6)***, n = 61	0.0003	0.527, n = 58	< 0.0001	0.89
Patient pre-visit	–4.6 (11.2), n = 59	0.002			
Joint Swelling Count					
Physician	–1.8 (6.3), n = 61	0.03	0.314, n = 58	0.016	0.76
Patient pre-visit	–1.2 (8.6), n = 59	0.28			

* Correlation with corresponding patient difference; ** signed-rank test for differences assessed by the patient and the physician; *** mean (SD).

physicians reflect differences in ability to measure swollen joints or differences in the perception of the meaning of "swollen joint" (e.g., a feeling of being swollen rather than ability to actually measure the swelling) or something completely different for physician and patient. In addition, the deliberate lack of patient training in how to do swollen joint counts may account for the differences in this aspect of the study. However, tracking these counts in the same direction over 3 months seemed to indicate similarities, and patient tender joint count paralleled physician tender joint count, indicating ability to show change.

The different numbers and patterns between patient and physician joint counts were chosen deliberately, since preliminary testing had shown that patients had difficulty separating tenderness or pain for individual MTP and toes, so we incorporated them as single joints. The same was true for the acromioclavicular and shoulder joints. The patients enrolled were familiar with the purpose and methods of joint counting, but were purposely not trained in how to perform joint counts, to approximate what would occur in clinical practice. Despite this, the self-administered JTC correlated well with the physicians' counts.

In summary, in the context of clinical trials in RA, a patient's self-assessment of joint tenderness seems to verify a patient's response to treatment as well as the physician's and might be a useful tool to evaluate change in RA activity over time.

REFERENCES

1. Furst DE. Predictors of worsening clinical variables and outcomes in rheumatoid arthritis. Rheum Dis Clin North Am 1994;20:309-19.
2. Smythe HA, Helewa A, Goldsmith CH. Selection and combination of outcome measures. J Rheumatol 1982;9:770-4.
3. Kirwan JR. Outcome measures in rheumatoid arthritis clinical trials: assessing improvement. J Rheumatol 1993;20:543-5.
4. Pincus T, Strand V, Koch G, et al. An index of the three core data set patient questionnaire measures distinguishes efficacy of active treatment from that of placebo as effectively as the American College of Rheumatology 20% response criteria (ACR20) or the Disease Activity Score (DAS) in a rheumatoid arthritis clinical trial. Arthritis Rheum 2003;48:625-30.
5. Boers M, van Riel PL, Felson DT, Tugwell P. Assessing the activity of rheumatoid arthritis. Baillieres Clin Rheumatol 1995;9:305-17.
6. Stewart MW, Palmer DG, Knight RG. A self-report articular index measure of arthritic activity: investigations of reliability, validity and sensitivity. J Rheumatol 1990;17:1011-5.
7. 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.
8. Abraham N, Blackmon D, Jackson JR, Bradley LA, Lorish CD, Alarcon GS. Use of self-administered joint counts in the evaluation of rheumatoid arthritis patients. Arthritis Care Res 1993;6:78-81.
9. Wong AL, Wong WK, Harker J, et al. Patient self-report tender and swollen joint counts in early rheumatoid arthritis. Western Consortium of Practicing Rheumatologists. J Rheumatol 1999;26:2551-61.
10. 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.
11. Hewlett SE, Haynes J, Shepstone L, Kirwan JR. Rheumatoid arthritis patients cannot accurately report signs of inflammatory activity. Br J Rheumatol 1995;34:547-53.
12. Hernandez-Cruz B, Cardiel MH. Intra-observer reliability of commonly used outcome measures in rheumatoid arthritis. Clin Exp Rheumatol 1998;16:459-62.