Validation of Patient-Reported Joint Counts in Rheumatoid Arthritis and the Role of Training

GERALD LEVY, CRAIG CHEETHAM, ALLISON CHEATWOOD, and RAOUL BURCHETTE

ABSTRACT. ***Objective.*** To demonstrate the effectiveness of simple training on improving the ability of patients with rheumatoid arthritis (RA) to assess joint swelling, and to validate the use of a computerized questionnaire, the Health Assessment Questionnaire (HAQ-ulous), to collect patient-reported tender and swollen joint counts.

Methods. Sixty patients completed the HAQ-ulous, reporting pain and swelling of the 28 joints included in the Disease Activity Score-28. A rheumatologist blinded to the patients' responses assessed each joint for the presence of tenderness and swelling. At followup visits, 30 patients received training in distinguishing a swollen joint from a chronically enlarged joint, completed the HAQ-ulous again, and were reassessed by the physician.

Results. At the initial visit, a strong correlation was shown between patient- and clinician-reported tender joints [Pearson correlation coefficient (r_p) = 0.79; $p < 0.0001$]. Correlation between patient- and clinician-reported swollen joints was less robust ($r_p = 0.41$; $p = 0.001$). Following training at the second visit, agreement between patients and the clinician improved for both tender joints ($r_p = 0.94$; $p < 0.0001$) and swollen joints ($r_p = 0.93$; $p < 0.0001$).

Conclusion. With simple training in distinguishing swollen joints from chronically enlarged joints, the majority of patients are able to accurately assess joint swelling. Objective tools, such as the HAQ-ulous, that incorporate patient-reported outcomes are a valuable and reliable addition to standard clinical practice for monitoring patients with RA. (First Release April 15 2007; J Rheumatol 2007;34:1261–5)

Key Indexing Terms:
RHEUMATOID ARTHRITIS SELF-ASSESSMENT TRAINING JOINT COUNTS
HEALTH ASSESSMENT QUESTIONNAIRE DISEASE ACTIVITY SCORE

Instruments such as the Health Assessment Questionnaire (HAQ) and the Disease Activity Score (DAS) have been shown to have great clinical value in evaluating and monitoring patients with rheumatoid arthritis (RA). Nevertheless, only a small percentage of rheumatologists have incorporated these tools into their standard, everyday clinical practice[1,2]. This is likely due to the time required to administer a questionnaire, assess the patient's joint pain and swelling, score the results, and record the information in a readily retrievable format.

Patient-reported outcomes (PRO) are an attractive option in a busy medical practice, as the time burden is transferred from the clinician to the patient. Still, questionnaires completed on paper must be scored by hand or scanned using specialized equipment and computer programs. These methods may not be practical for physicians trying to incorporate PRO into their routine patient visits. Also, monitoring patients' disease activity and responses to treatment over time using paper PRO is difficult and time-consuming. To fully realize the benefit of using PRO to monitor a patient's progress, data from all previous visits should be readily available. To address this problem, we developed a computerized questionnaire (HAQ-ulous) that incorporates the traditional HAQ, a homunculus to record tender and swollen joints, and results of laboratory tests for inflammation (either erythrocyte sedimentation rate or C-reactive protein). After the patient completes the HAQ-ulous, the responses are automatically summarized and displayed on a table that includes the patient's previous data. This allows the physician to quickly assess patient response to the current treatment.

The validity and usefulness of PRO in measuring disease activity in patients with RA have been well documented[3]. In 1992, Mason, *et al*[4] validated the Rapid Assessment of Disease Activity in Rheumatology (RADAR), a 2-page patient questionnaire, demonstrating good agreement between patient and physician assessments of RA activity and level of function. Similarly, Hanly, *et al*[5] reported that a simple self-report questionnaire is a reliable assessment of disease activity. Several studies have also shown that patient-reported joint counts correlate well with physician assessments, with less robust correlation for joint swelling than joint tenderness

From the Department of Rheumatology, Southern California Kaiser Permanente, Bellflower, California, USA.

Supported by funding from the Regional Research Committee of Southern California Kaiser Permanente.

G.D. Levy, MD, MBA, Rheumatologist, Southern California Permanente Medical Group; C.T. Cheetham, PharmD, MS, Pharmacist, Pharmacy Analytical Service; A.P. Cheatwood, MS, Research Assistant, Department of Rheumatology; R.J. Burchette, MA, MS, Data Consultant, Research and Evaluation, Southern California Kaiser Permanente.

Address reprint requests to Dr. G.D. Levy, Department of Rheumatology, Kaiser Permanente, 9449 East Imperial Hwy, Downey, CA 90242. E-mail: gerald.d.levy@kp.org

Accepted for publication January 10, 2007.

(Table 1)[5-11]. Patient training has been suggested by a number of authors as a method to improve the reliability of patient reports[11-15].

Our primary purpose was to validate the use of the HAQ-ulous to collect patient-reported joint counts for both tender and swollen joints. A second goal was to determine if the correlation of patient- and physician-reported joint counts improves following simple training to distinguish swollen joints from joints that are chronically enlarged secondary to osteoarthritis without active inflammation.

MATERIALS AND METHODS

Sixty patients with RA at the Kaiser Permanente Bellflower Department of Rheumatology agreed to participate in this study. The Southern California Kaiser Permanente Institutional Review Board approved this study and informed consent was obtained from each subject prior to enrollment.

As part of their routine clinic visit, a nurse or research assistant logged the patient into the HAQ-ulous program and reviewed their current medications. The patient had the option to complete the HAQ-ulous in either English or Spanish. For patients who were computer literate or had used the program before, the HAQ-ulous is self-explanatory and instruction was minimal. Patients who were "computer illiterate" were instructed in how to use a computer, a computer mouse, and the HAQ-ulous program. These patients were generally able to master the necessary skills in a few minutes. Occasionally a patient was not able to operate the mouse, in which case a family member or clinic employee assisted the patient in filling out the questionnaire without giving input as to which joints were tender or swollen.

We use a modified HAQ[16], which includes the standard 10 HAQ functional questions, 3 mental health questions, and visual analog scales for pain, fatigue and general health. The computer program also has 2 homunculi on which the patient can indicate painful or swollen joints. A patient's response indicating painful and swollen joints was used to calculate the DAS28[17]. The DAS28 joints are the shoulders, elbows, wrists, knees, 10 metacarpophalangeal joints, and 10 proximal interphalangeal joints. In this study, we utilized the DAS28 as opposed to the DAS44 joint formula, for patient convenience and because of the number of studies validating the DAS28 as a useful measure of disease activity[7,14,17,18].

Once the patient completed the HAQ-ulous, a clinician, blinded to the patient-reported joint counts, assessed the patient for tender and swollen joints. The research assistant recorded the clinician-reported joint counts.

Sixty patients completed an initial visit for the primary analysis. In addition, the first 30 patients to return for a followup visit repeated the HAQ-ulous immediately after being trained to identify swollen joints and were then reassessed by the clinician. The training consisted of informing the patient that a swollen joint should be warm and puffy with excess fluid and that a bony enlargement secondary to osteoarthritis should not be considered swollen. The entire training, including instruction on use of the HAQ-ulous, training on distinguishing a swollen joint from a chronically enlarged joint, and time for questions, generally lasted less than 5 minutes.

Comparisons were done to determine if patients improved over time in assessing painful and swollen joints, and specifically, if simple training improved a patient's ability to distinguish between a swollen joint with active inflammation and a joint that is chronically enlarged.

Descriptive statistical analyses were performed for patient demographic data and joint counts. Agreement between patient-reported and clinician-reported tender and swollen joint counts was determined by calculating both Pearson (r_p) and Spearman (r_s) correlation coefficients and associated p values (p_p and p_s, respectively) at each visit. Comparability between patients with one visit and patients with 2 visits was tested by t tests and Wilcoxon rank-sum tests on the first visit data.

RESULTS

The mean age of the patients in the study was 54.1 ± 14.8 years and 46 were female (76.7%). The mean disease duration for the group was 7.4 ± 5.7 years (range 0–27 yrs). At the initial assessment, patients reported an average of 3.5 ± 3.3 tender and 2.5 ± 3.3 swollen joints (Table 2). The clinician reported an average of 3.9 ± 3.7 tender and 2.0 ± 2.0 swollen joints. Pearson correlation showed strong agreement between patient and clinician reported tender joints (r_p = 0.79, r_s = 0.83, $p_{p,s} < 0.0001$; Figure 1) with less robust agreement between patient- and clinician-reported swollen joints (r_p = 0.41, p_p = 0.001; r_s = 0.64, $p_s < 0.0001$; Figure 2). One of the patients may not have understood the instructions, indicating 19 swollen joints compared to only 1 swollen joint reported by the clinician. Removal of this single patient from the data set changed the results dramatically. The revised calculation showed a Pearson correlation of r_p = 0.60 (r_s = 0.67, $p_{p,s} <$ 0.0001) for the swollen joint counts.

A second evaluation was completed for 30 patients an average of 49.73 ± 16.06 days after their first visit. Patients reported an average of 3.60 ± 4.56 tender and 2.20 ± 3.89 swollen joints compared to the clinician report of 3.93 ± 5.04

Table 1. Review of references reporting comparisons between patient- and-clinician-reported joint tenderness and/or swelling.

Reference	Year of Study	Number of Subjects	Correlation Coefficients	
			Joint Tenderness	Joint Swelling
Abraham[6]	1993	32	0.89[P]	—
Hanly[5]	1996	61	0.57[P], 0.31[I]	0.16[P], –0.02[I]
Prevoo[7]	1996	236	0.60–0.65[P]	0.51–0.64[P]
Escalante[8]	1998	110	0.78[I]	0.31[I]
Calvo[9]	1999	60	0.75–0.77[S]	—
Wong[10]	1999	60	0.56–0.65[I]	0.61–0.64[I]
Houssien[11]	1999	100	0.88[S]	0.63[S]
Current study (1st visit)	2006	60	0.79[P], 0.83[S]	0.41[P], 0.64[S]
Current study (2nd visit)	2006	30	0.94[P], 0.89[S]	0.93[P], 0.52[S]

P, Pearson Correlation Coefficient; S, Spearman Correlation Coefficient; I, Intraclass Correlation Coefficient.

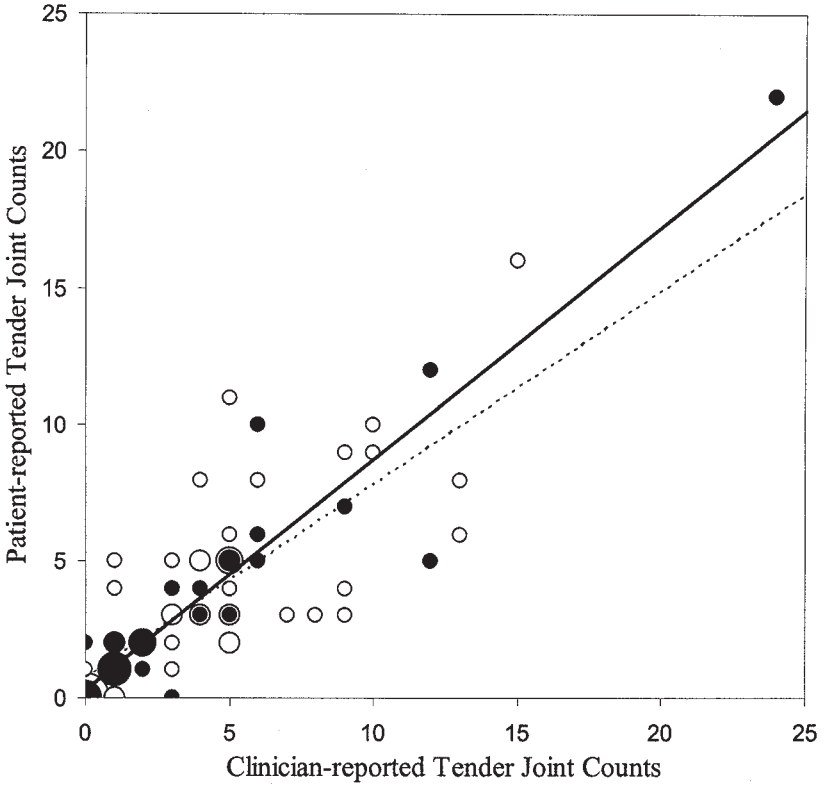

Figure 1. Patient-reported and clinician-reported tender joint counts for the original 60 patients at first visit (○) and for the 30 patients reevaluated at second visit (●). Broken line represents trendline for first visit; solid line represents trendline for second visit. Area of each data point is proportional to frequency of the observed value.

tender and 1.60 ± 4.08 swollen joints. Pearson correlation showed a strong agreement between the patients and clinician for both tender ($r_p = 0.94$, $p_p < 0.0001$; $r_s = 0.89$, $p_s < 0.0001$; Figure 1) and swollen ($r_p = 0.93$, $p_p < 0.0001$; $r_s = 0.52$, $p_s = 0.003$; Figure 2) joint counts. The Pearson correlation for swollen joints increased dramatically, whereas the Spearman correlation decreased slightly between the first and second visits because of one patient with a higher swollen joint count in the second visit (Figure 2).

As the 30 return patients were seen in sequential order based on their normal followup appointment schedule, no significant difference was found between the first group of 60 and the followup group of 30 for age, sex, duration of disease, and joint counts.

DISCUSSION

Our study demonstrates a strong correlation between patient- and clinician-reported tender joint counts and with training, a correlation between patient- and clinician-reported joint swelling. On first exposure to the HAQ-ulous program, patient-clinician correlation values were similar to those found in previous studies. Following simple training, the correlation values improved from 0.79 to 0.94 for tender joints and from 0.41 to 0.93 for swollen joints. The Spearman values for swollen joints decreased due to the paucity of involved joints and the lack of variation among the study population (Figure 2). The Pearson correlation increased dramatically between the first and second visits because of the patient with a high swollen joint count. Because the Spearman correlation is based on rank, the positive effect of this single patient on the R value is considerably reduced. We expect that if more patients with elevated joint counts were included, the Spearman correlation for swollen joint counts would increase.

One potential weakness of our study is the incomplete spectrum of disease of our patients with RA, as most patients had minimal to moderate disease. Inclusion of patients with more active disease would allow us to generalize results of this study to the entire spectrum of patients with RA. Further work is required to show if the HAQ-ulous correlation is true across a broad range of disease activity; however, the single patient with a very high swollen joint count that compared well with the physician-reported joint count (21 vs 22, respectively) is encouraging. Also, some clinicians prefer to use the 44 joint DAS, which includes feet and ankles, as opposed to the DAS28 used in this study. Future versions of the HAQ-ulous may include a "switch" to allow the rheumatologist to select the preferred DAS scale appropriate for their practice.

Taking the first 30 patients that returned for treatment followup may have introduced some unknown biases, especially if sick patients required more frequent visits. We do not

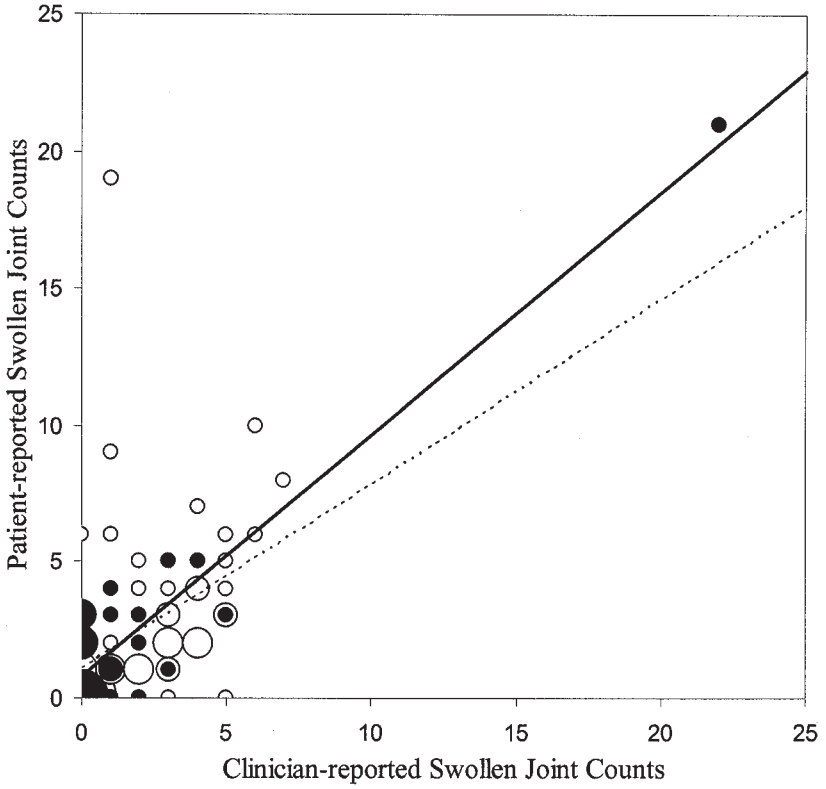

Figure 2. Patient-reported and clinician-reported swollen joint counts for the original 60 patients at first visit (○) and for the 30 patients reevaluated at second visit (●). Broken line represents trendline for first visit; solid line represents trendline for second visit. Area of each data point is proportional to frequency of the observed value.

Table 2. Patient demographic and clinical data.

	First Visit	Second Visit
Sample size, n	60	30
Women (%)	46 (77)	22 (73)
Age, yrs, mean ± SD	54.1 ± 14.8	54.0 ± 13.7
Disease duration, years, mean ± SD	7.4 ± 5.7	7.4 ± 6.1
Patient-reported tender joint count, mean ± SD	3.5 ± 3.3	3.6 ± 4.6
Clinician-reported tender joint count, mean ± SD	3.9 ± 3.7	3.9 ± 5.0
Patient-reported swollen joint count, mean ± SD	2.5 ± 3.3	2.2 ± 3.9
Clinician-reported swollen joint count, mean ± SD	2.0 ± 2.0	1.6 ± 4.1

believe this to be the case, however, because comparisons between the 1-visit and 2-visit groups showed no significant differences in demographics or joint counts. Further, results on the difference between doctor- and patient-reported outcomes (not shown) indicated that the training for properly identifying swollen joints did not reduce the variance of the difference between these 2 measures, while this variance was significantly reduced for tender joints. This may indicate a need to enhance or revise the training material for swollen joints. As more centers begin to use the HAQ-ulous the study could be repeated to see how well results hold up with different clinician evaluators in different clinical settings.

Although we did not specifically test the permanence of the training, to prevent decay and to improve the quality of future responses, the physician normally reviews with the patient his/her responses during the office visit. We anticipate this reinforcement of training will improve correlations over time and should be tested in the future.

Tender and swollen joint counts are the traditional "gold standard" method for measuring disease activity in patients with RA. Joint counts are time-consuming and are generally performed only by trained clinicians participating in clinical research studies. Rheumatologists in a busy practice setting infrequently do joint counts. Cush reported that only 12% of rheumatologists collect and score a HAQ and only 6% calculate DAS as part of their routine clinic visit[1].

Patient-reported joint counts were first described in 1992 as an element of the RADAR questionnaire[4]. Results of this study showed excellent agreement between patient- and clinician-reported joint tenderness (intraclass correlation coefficient for total joint pain/tenderness score was 0.81). Since then, many studies have validated the use of patient-reported tender and swollen joint counts with correlation coefficients between patient- and clinician-reported joint counts ranging from 0.31 to 0.89 for tender joints and from –0.02 to 0.64 for swollen joints (Table 1). Remarkably, only a small percentage of rheumatologists have incorporated these tools into their

standard, everyday clinical care. One consideration may be the lower correlation seen in joint swelling compared to joint tenderness.

The HAQ and DAS are objective tools that can be used to more precisely monitor disease activity. With simple training, the majority of patients are able to grasp the difference between swollen joints and joints that are chronically enlarged without active inflammation and are able to accurately report joint tenderness and swelling. Our study suggests that the use of the HAQ-ulous can be confidently integrated into routine clinical practice and that simple patient training can allow patients to distinguish between chronically enlarged and swollen joints. Plans are under way to make the HAQ-ulous available on the Internet as a service to practicing rheumatologists.

REFERENCES

1. Cush JJ. Biological drug use: US perspectives on indications and monitoring. Ann Rheum Dis 2005;64:iv18-iv23.
2. Wolfe F, Pincus T, Thompson AK, Doyle J. The assessment of rheumatoid arthritis and the acceptability of self-report questionnaires in clinical practice. Arthritis Rheum 2003;49:59-63.
3. Wolfe F, Pincus T. Listening to the patient: a practical guide to self-report questionnaires in clinical care. Arthritis Rheum 1999;42:1797-808.
4. 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.
5. Hanly JG, Mosher D, Sutton E, Weerasinghe S, Theriault D. Self-assessment of disease activity by patients with rheumatoid arthritis. J Rheumatol 1996;23:1531-8.
6. 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.
7. Prevoo ML, Kuper IH, van't Hof MA, van Leeuwen MA, van de Putte LB, van Riel PL. Validity and reproducibility of self-administered joint counts. A prospective longitudinal followup study in patients with rheumatoid arthritis. J Rheumatol 1996;23:841-5.
8. Escalante A. What do self-administered joint counts tell us about patients with rheumatoid arthritis? Arthritis Care Res 1998;11:280-90.
9. 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.
10. 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.
11. Houssien DA, Stucki G, Scott DL. A patient-derived disease activity score can substitute for a physician-derived disease activity score in clinical research. Rheumatology Oxford 1999;38:48-52.
12. Hart LE, Tugwell P, Buchanan WW, Norman GR, Grace EM, Southwell D. Grading of tenderness as a source of interrater error in the Ritchie Articular Index. J Rheumatol 1985;12:716-7.
13. Lewis PA, O'Sullivan MM, Rumfeld WR, Coles EC, Jessop JD. Significant changes in Ritchie scores. Br J Rheumatol 1998;27:32-6.
14. Scott DL, Choy EH, Greeves A, et al. Standardising joint assessment in rheumatoid arthritis. Clin Rheumatol 1996;15:579-82.
15. Thompson PW, Hart LE, Goldsmith CH, Spector TD, Bell MJ, Ramsden MF. Comparison of four articular indices for use in clinical trials in rheumatoid arthritis: patient, order, and observer variation. J Rheumatol 1991;18:661-5.
16. Pincus T, Summey JA, Soraci SA Jr, Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. Arthritis Rheum 1983;26:1346-53.
17. Prevoo ML, van't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum 1995;38:44-8.
18. Smolen JS, Breedveld FC, Eberl G, et al. Validity and reliability of the twenty-eight-joint count for the assessment of rheumatoid arthritis activity. Arthritis Rheum 1995;38:38-43.