Patient-derived Joint Counts Are a Potential Alternative for Determining Disease Activity Score

ARTHUR KAVANAUGH, SUSAN J. LEE, HAOLING H. WENG, YUN CHON, XING-YUE HUANG, and SHAO-LEE LIN

ABSTRACT. ***Objective.*** To investigate the correlation between the Disease Activity Score using a 28-joint count (DAS28) based on physician-derived joint counts and the DAS28 based on patient-derived joint counts (Pt-DAS28) in rheumatoid arthritis (RA).

Methods. Data from a multicenter, open-label study investigating the immunogenicity of etanercept (ETN) were analyzed. ETN-naive patients with active RA received ETN 50 mg once weekly alone or with methotrexate (MTX). Joint counts were performed at baseline, Week 12, and Week 24 by the physician and patient independently. Patients received instruction in performing joint assessments.

Results. Of 447 patients enrolled (ETN, n = 218; ETN + MTX, n = 229), most were women (79%) and the mean age was 54.5 years. Correlation coefficients between DAS28 and Pt-DAS28 were ≥ 0.57 at baseline, Week 12, and Week 24. At Week 24, 48%, 39%, and 12% of patients could be classified as having low, moderate, or high disease activity, respectively, using DAS28. Using Pt-DAS28, 43%, 39%, and 18% were similarly classified. Agreement in the category of disease activity classification occurred in 72% of patients (kappa = 0.55). At Week 24, 78% of patients using DAS28 and 72% of patients using Pt-DAS28 were classified as moderate or good European League Against Rheumatism responders.

Conclusion. These results support the possible use of patient-derived tender and swollen joint counts to aid in the assessment of disease activity and clinical response in patients with RA. (First Release Feb 15 2010; J Rheumatol 2010;37:1035–41; doi:10.3899/jrheum.090704)

Key Indexing Terms:
RHEUMATOID ARTHRITIS PATIENT REPORTED OUTCOMES
DISEASE ACTIVITY SCORE EULAR RESPONSE ETANERCEPT

Controlling disease activity is important in the management of rheumatoid arthritis (RA)[1]. Frequent assessments of RA can be beneficial in improving patient outcomes by allowing treatment adjustments to minimize disease activity and improve disease outcome[2]. The Disease Activity Score (DAS) and the shorter 28-jount count DAS (DAS28) are well validated measures of RA disease activity[3,4]. Tender and swollen joint counts are the key components of the DAS score, which also includes a measure of the acute-phase response and a patient global assessment of disease activity. DAS scores correlate with important outcomes such as joint damage, and changes in DAS scores are widely used to assess responses to treatment[5,6].

Quantitative measures of disease activity are increasingly monitored by healthcare providers during office visits to assess the need for treatment changes. Joint assessments have traditionally been performed in-office by physicians or trained healthcare personnel. However, it is possible that patient self-assessment could be performed at home, which would allow more frequent assessment of disease activity and may be a reliable alternative when clinical resources are limited.

We evaluated the consistency between physician- and patient-derived tender and swollen joint counts and corresponding DAS28 scores over time in a cohort of patients with RA participating in a 24-week open-label trial of etanercept (ETN) alone versus ETN + methotrexate (MTX). Patient-derived DAS28 (Pt-DAS28) scores were calculated using the same formula used to calculate DAS28, except that patient joint counts were used in place of physician joint

From the Division of Rheumatology, Allergy and Immunology; and Department of Medicine, University of California, San Diego, La Jolla, California; Veterans Affairs San Diego Healthcare System, San Diego, California; and Amgen Inc., Thousand Oaks, California, USA.

Supported by Immunex Corporation, a wholly owned subsidiary of Amgen Inc. and by Wyeth Research. Drs. Weng, Chon, Huang, and Lin are Amgen employees and own Amgen stock and stock options. Dr. Kavanaugh received research grants from Amgen. Writing assistance supported by Amgen, Inc.

A. Kavanaugh, MD, Professor of Medicine, Division of Rheumatology, Allergy and Immunology, University of California at San Diego (UCSD); S.J. Lee, MD, Veterans Affairs San Diego Healthcare System, Assistant Professor of Medicine, Department of Medicine, UCSD; H.H. Weng, MD, MHS, Clinical Research Medical Director; Y. Chon, PhD, Biostatistics Senior Manager; X-Y. Huang, PhD, Health Economics Senior Manager; S-L. Lin, MD, PhD, Executive Medical Director Global Development, Amgen Inc.

Address correspondence to Dr. A. Kavanaugh, Division of Rheumatology, Allergy and Immunology, University of California, 9500 Gilman Drive, La Jolla, CA 92037-0943. E-mail: akavanaugh@ucsd.edu

Accepted for publication November 13, 2009.

See related editorial in this issue.

counts. Our aim was to determine whether patient self-assessment of tender and swollen joints correlated with physician assessments and whether patient-derived joint counts were a reliable substitution for physician assessments in determining DAS28 status in RA patients undergoing treatment.

MATERIALS AND METHODS

Patients. Data were derived from ETN-naive patients (n = 447) with active RA (≥ 18 yrs old) participating in an open-label study assessing the immunogenicity of ETN[7]. Patients were randomized to receive subcutaneous ETN 50 mg or ETN 50 mg + MTX once weekly for 24 weeks. All patients provided written informed consent, and all study procedures were approved by the appropriate institutional ethics committee.

Assessments. Joint evaluations by the patient and his or her physician were performed during the study visit. The patient's evaluation was performed first, followed by the physician's evaluation during the same study visit. The physicians were blinded to the results of the patients' evaluations and were not involved in the treatment strategies of the clinical trial. Neither patient nor physician was aware of the other's evaluation. Joint assessments included the 28-joint count for tenderness and swelling as established by the European League Against Rheumatism (EULAR)[8]. Assessments were performed at baseline, Week 12, and Week 24 or termination of the study.

To train patients in performing tender joint self-assessments, instructions were given to "Please put an 'X' on each joint you see on a mannequin that corresponds to the joints that hurt you. There may be pain when you press or move your joints." For swollen joint counts, patients were instructed to "Please put an 'X' on each joint you see on a mannequin that corresponds to the joints you feel are swollen. The joints may feel swollen or look swollen"[9]. DAS28 and Pt-DAS28 were calculated from these joint counts using the formula:

$$\text{DAS28} = 0.56 \times \sqrt{\text{TEN28 (Physician or Patient)}} + 0.28 \times \sqrt{\text{SW28 (Physician or Patient)}} + 0.36 \times \ln(\text{CRP} + 1) + 0.014 \times \text{PtGA} + 0.96$$

DAS28 and Pt-DAS28 scores range from 0 to 9.4. Disease activity can be categorized as low (≤ 3.2), moderate (> 3.2 to ≤ 5.1), or high (> 5.1) based on the DAS28 score[10]. Baseline and posttreatment DAS28 and Pt-DAS28 scores were further used to identify EULAR responders, that is, those with "good" or "moderate" responses to treatment (Table 1)[11].

Patients also assessed physical function at baseline, Week 12, and Week 24 using the Health Assessment Questionnaire Disability Index (HAQ-DI)[12,13]. Two other composite scores, the Clinical Disease Activity Index (CDAI)[14] and Simplified Disease Activity Index (SDAI)[15], were each calculated using either physician joint counts or patient joint counts at baseline, Week 12, and Week 24.

Statistics. Patients who received at least 1 dose of ETN or ETN + MTX were included in the analyses. Convergent validity was assessed by Spearman correlation using rank scores and Pearson product-moment correlation. Intraclass correlation coefficients were used to measure agreement between the continuous variables of physician and patient tender and swollen joint counts. Kappa statistics were used to determine agreement between the noncontinuous variables of physician and patient individual joint assessments. Further, Bland-Altman plots were included to visually assess the agreement between physician and patient tender and swollen joint counts across different levels of joint involvement. Agreement between DAS28 and Pt-DAS28 in classifying patients' disease as low, moderate, or high was measured by kappa statistics[16]. EULAR responses based on DAS28 scores were used as a standard to measure the sensitivity and specificity of the EULAR response based on Pt-DAS28 scores. Logistic regression analysis, adjusting for age, sex, and treatment, was used to determine differences in physician- and patient-derived joint counts between baseline, Week 12, and Week 24.

RESULTS

Patients. A total of 447 patients were enrolled; 218 received ETN alone and 229 received ETN + MTX. Baseline demographics were similar between treatment groups (Table 2). Most patients were women (79%), white (81%), and < 65 years of age (77%; mean 54.5 yrs old; range 19–88 yrs). At baseline, mean DAS28 (5.4 vs 5.0) and Pt-DAS28 (5.5 vs 5.1) were similar between the ETN and ETN + MTX arms. Mean RA disease duration was 8.5 years in ETN-treated patients compared with 7.2 years in ETN + MTX-treated patients.

Tender and swollen joint assessments. The results of physician-derived TJC and SJC were strongly correlated with the results of patient-derived TJC and SJC (Table 3). The proportion of patients with differences of ≤ 3 joints between physician and patient TJC was 52.6% at baseline, 67.6% at Week 12, and 65.4% at Week 24. Similarly, the proportions of patients with SJC differences of ≤ 3 were 46.6%, 58.6%, and 61.7% at baseline, Week 12, and Week 24, respectively. At Week 24, the intraclass correlation coefficients (r) between physician and patient TJC (r = 0.7771) were greater than those for SJC (r = 0.4341). There were no clinically relevant differences in correlations between tender or swollen joints when the cohort was stratified by ETN or ETN + MTX treatment (data not shown).

The correlation between CDAI scores calculated using either physician or patient joint counts was 0.73 or higher at baseline, Week 12, and Week 24 (Table 3). Similarly, the correlation between SDAI scores calculated using physician or patient joint counts was 0.75 or higher at baseline, Week 12, and Week 24. Thus, the DAS28, CDAI, and SDAI all show a high degree of correlation when calculated using either physician or patient joint counts over the 24-week study period.

Table 1. Criteria for European League Against Rheumatism response.

DAS28 at Current Visit	DAS28 Improvement from Baseline		
	> 1.2	> 0.6 to ≤ 1.2	≤ 0.6
≤ 3.2	Good response	Moderate response	No response
> 3.2 to ≤ 5.1	Moderate response	Moderate response	No response
> 5.1	Moderate response	No response	No response

DAS28: disease activity score.

Table 2. Patient demographics.

Characteristic	ETN, n = 218	ETN/MTX, n = 229	All, n = 447
Mean age, yrs (range)	54.3 (21–88)	54.7 (19–82)	54.5 (19–88)
Sex, n (%)			
Men	40 (18.3)	53 (23.1)	93 (20.8)
Women	178 (81.7)	176 (76.9)	354 (79.2)
Race, n (%)			
White	176 (80.7)	188 (82.1)	364 (81.4)
Black	17 (7.8)	11 (4.8)	28 (6.3)
Hispanic	21 (9.6)	22 (9.6)	43 (9.6)
Asian	3 (1.4)	5 (2.2)	8 (1.8)
American Indian or Alaska Native	1 (0.5)	1 (0.4)	2 (0.4)
Other	0 (0.0)	2 (0.9)	2 (0.4)
Age, yrs, n (%)			
< 65	169 (77.5)	174 (76.0)	343 (76.7)
≥ 65	49 (22.5)	55 (24.0)	104 (23.3)
Mean weight, kg (SD)	78.9 (20.2)*	82.3 (25.4)	80.7 (23.1)
Mean height, cm (SD)	164.8 (8.5)*	166.5 (9.6)	165.7 (9.1)
Mean BMI, kg/m^2 (range)	29.0 (18.0–50.3)*	29.5 (15.7–67.2)	29.3 (15.7–67.2)
Mean C-reactive protein, mg/l (SD)	16.4 (20.9)	14.4 (29.9)	ND
Mean tender joint count (0–68), (SD)	23.7 (15.7)	21.5 (16.8)	22.6 (16.3)
Median tender joint count (0–68), (range)	21.5 (0–68)	19.0 (0–68)	21.0 (0–68)
Mean swollen joint count (0–66), (SD)	14.5 (9.6)	13.3 (9.1)	13.9 (9.4)
Median swollen joint count (0–66), (range)	13.0 (0–50)	11.0 (0–50)	21.0 (0–50)
Mean HAQ (SD)	1.4 (0.7)	1.3 (0.7)	1.3 (0.7)
Mean DAS28 (SD)	5.4 (1.3)	5.0 (1.3)	5.2 (1.3)
Mean Pt-DAS28 (SD)	5.5 (1.3)	5.1 (1.4)	5.3 (1.3)
Mean patient global assessment of disease activity (SD); scale 0–100	58.7 (23.4)*	53.7 (24.0)	56.1 (23.8)
Mean physician global assessment of disease activity (SD); scale 0–100	55.1 (22.6)	51.9 (24.9)	53.5 (23.9)

* n = 217. BMI: body mass index; DAS28: disease activity score; ETN: etanercept; HAQ: Health Assessment Questionnaire; MTX: methotrexate; ND: not determined; Pt-DAS28: patient-derived DAS28.

Table 3. Correlations between physician and patient tender and swollen joints and HAQ-DI, CDAI, and SDAI scores calculated using physician/patient joint counts.

	Baseline	Week 12	Week 24
Correlation between patient and physician joint counts*			
Tender joint counts	0.78		0.78
Swollen joint counts	0.55	0.53	0.43
Correlation between HAQ-DI and patient/physician joint counts†			
HAQ-DI vs physician tender joint counts	0.42	0.54	0.51
HAQ-DI vs physician swollen joint counts	0.25	0.38	0.31
HAQ-DI vs patient tender joint counts	0.47	0.58	0.54
HAQ-DI vs patient swollen joint counts	0.37	0.39	0.48
Correlation between DAS28 and Pt-DAS28†	0.73	0.75	0.69
Correlation between physician CDAI and patient CDAI†	0.73	0.78	0.75
Correlation between physician SDAI and patient SDAI†	0.75	0.79	0.76

* Intraclass correlation coefficients for physician- and patient-derived tender and swollen joint counts. † Spearman correlation coefficients for physician- and patient-derived tender and swollen joint counts. CDAI: Clinical Disease Activity Index; DAS28: Disease Activity Score; Pt-DAS28: Patient-derived Disease Activity Score; SDAI: Simplified Disease Activity Score; HAQ-DI: Health Assessment Questionnaire-Disability Index.

To assess whether baseline differences resulted in differences at 12 and 24 weeks, logistic regression models were used to adjust for covariates. When physician and patient baseline assessments differed by ≥ 5 tender/swollen joints, patients were 4.7 (95% CI 2.87–7.60) and 2.7 (95% CI 1.71–4.21) times more likely to have such differences at 12 and 24 weeks, respectively, compared with patients without such differences at baseline.

To assess agreement between the physician and patient assessments across varying levels of joint involvement, Bland-Altman plots were used to illustrate differences between physician and patient joint counts versus their average at baseline and at Week 24 (Figure 1). At baseline, the mean absolute difference between physician and patient tender and swollen joints was 4.81 and 5.66, respectively (Figure 1A, 1C). By Week 24, the mean absolute difference between physician and patient tender and swollen joint counts decreased compared with baseline and was 3.86 and 4.37, respectively. In addition, the average number of joint counts decreased, indicating disease improvement due to treatment (Figure 1B, 1D). Generally, the largest variation between physician and patient assessment occurred at intermediate levels of joint involvement, or between 10 to 20 joints. More patients tended to view their disease condition to be worse than their physician, as evidenced by more data points clustering below the mean.

When individual joints were scored by physicians and patients for tenderness and swelling, the greatest agreement was observed for larger joints, particularly shoulders, knees, and elbows (Table 4). Agreement was less frequent when assessing swollen joints, with the exception of the knee (kappa = 0.40–0.55). Agreement was similar across time-

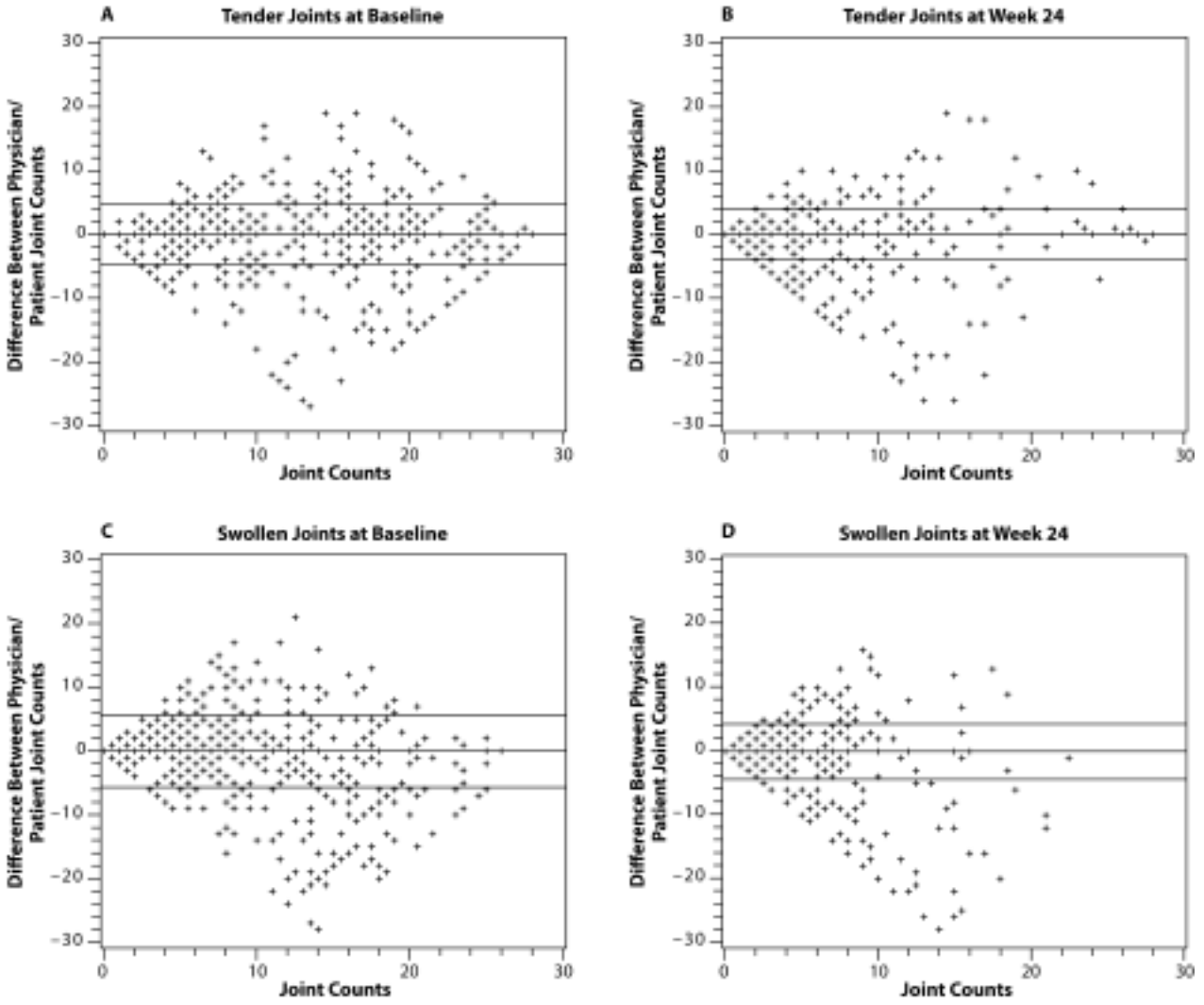

Figure 1. Bland-Altman plots of agreement between physician and patient tender and swollen joint counts over time. For each plot (+) represents the difference in joint counts assessed by the physician and the patient. Reference lines above and below zero represent the mean of the absolute difference between physician and patient joint counts for the analysis. Joint differences plotted at zero represent perfect agreement; positive joint count differences represent a higher physician joint count compared with the patient joint count; negative joint count differences indicate that patients counted more joints than the physician. When assessing a plot in total, a diamond-shaped geometry represents good correlation at low and high joint numbers with less agreement at an intermediate number of joints. A. Tender joint counts at baseline. The mean of the absolute difference between physician and patient joint counts was ± 4.81. B. Tender joint counts at Week 24. The mean of the absolute difference between physician and patient joint counts was ± 3.86. C. Swollen joint counts at baseline. The mean of the absolute difference between physician and patient joint counts was ± 5.66. D. Swollen joint counts at Week 24. The mean of the absolute difference between physician and patient joint counts was ± 4.37.

Table 4. Agreement between physician and patient individual joint assessment by kappa statistics.

Joint	Baseline		Week 12		Week 24	
	Tender Joints	Swollen Joints	Tender Joints	Swollen Joints	Tender Joints	Swollen Joints
Shoulder	0.49	0.15	0.48	0.17	0.55	0.22
Knee	0.53	0.40	0.50	0.53	0.59	0.55
Elbow	0.52	0.34	0.61	0.42	0.50	0.25
Wrist	0.45	0.29	0.48	0.29	0.54	0.22
MCP 1	0.39	0.20	0.45	0.28	0.47	0.17
MCP 2	0.46	0.27	0.51	0.25	0.49	0.28
MCP 3	0.43	0.26	0.47	0.27	0.45	0.23
MCP 4	0.44	0.24	0.43	0.26	0.46	0.20
MCP 5	0.46	0.26	0.39	0.19	0.41	0.20
PIP 1	0.36	0.33	0.39	0.19	0.39	0.25
PIP 2	0.35	0.23	0.38	0.30	0.39	0.22
PIP 3	0.37	0.26	0.51	0.34	0.40	0.22
PIP 4	0.39	0.25	0.44	0.23	0.38	0.20
PIP 5	0.36	0.24	0.43	0.12	0.29	0.08

MCP: metacarpophalangeal; PIP: proximal interphalangeal; 1: thumb, 2: index, 3: middle, 4: ring, 5: small. Kappa statistics measure agreement between the 2 counts. A value of 1.0 indicates perfect agreement, 0 indicates the same agreement as expected by chance, and a negative value indicates less agreement than would be expected by chance.

points and treatment groups. Similarly, HAQ-DI moderately correlated with physician- and patient-derived TJC (0.54, 0.58 at Week 12 and 0.51, 0.54 at Week 24, respectively) and less with SJC (0.38, 0.39 at Week 12 and 0.31, 0.48 at Week 24, respectively; Table 3).

Agreement of DAS28 and Pt-DAS28. Agreement between DAS28 and Pt-DAS28 was high throughout the study (Figure 2). For DAS28, 8%, 35%, and 57% of patients could be classified as having low, moderate, or high disease activity at baseline, respectively. For Pt-DAS28, 7%, 36%, and 57% had low, moderate, and high disease activity at baseline. Of the 34 patients who had low disease activity at baseline by physician assessment, 14 (41%) were also categorized as having low disease activity by patient assessment. Of the 256 patients with high disease activity by physician assessment, 218 (85%) also had high disease activity by patient assessment. At Week 24, 48%, 39%, and 12% of patients were classified as having low, moderate, or high disease activity by DAS28 and 43%, 39%, and 18% by Pt-DAS28. Overall agreement across disease severity was 75% at baseline, 73% at Week 12, and 72% at Week 24. Kappa statistics for agreement between physician and patient assessments of disease severity were 0.54 (95% CI 0.48–0.61) at baseline, 0.57 (95% CI 0.50–0.63) at Week 12, and 0.55 (95% CI 0.48–0.62) at Week 24. Similar results were found when patients were assessed according to treatment.

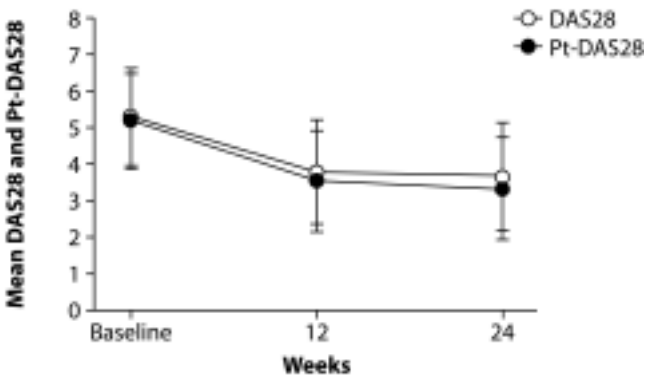

Figure 2. Agreement between DAS28 and Pt-DAS28 at baseline (n = 447), Week 12 (n = 416 and n = 414) and Week 24 (n = 378 and n = 379), respectively. DAS and Pt-DAS expressed as mean ± SD. DAS28: 28-joint count Disease Activity Score; Pt-DAS28: patient DAS28.

Agreement between EULAR responses using DAS28 and Pt-DAS28. When DAS28 and Pt-DAS28 were used to classify patients according to EULAR criteria, 78% and 72% of patients, respectively, were classified as EULAR responders (good or moderate) at Week 24. Sensitivity (0.83) and specificity (0.67) were high, suggesting good agreement between EULAR response based on DAS28 and Pt-DAS28. At 24 weeks, 79% of patients showed agreement in EULAR response using DAS28 or Pt-DAS28 (kappa = 0.45).

DISCUSSION

This study demonstrated good agreement between physician- and patient-derived DAS28 scores over time in patients treated with ETN. Physician- and patient-derived TJC and SJC were well correlated, with better agreement noted for tender versus swollen joints. The HAQ-DI was similarly correlated with physician- and patient-derived ten-

der and swollen joint counts. EULAR response rates for patients were similar when DAS28 and Pt-DAS28 were used to assess disease activity. Together, these data support the possible use of patient-derived tender and swollen joint counts to help determine disease activity in patients with RA.

Other studies have compared the level of agreement between physician and patient assessment of TJC and SJC, and, consistent with the data presented here, stronger correlations have been observed between physicians and patients for tender joints compared with swollen joints[9,17-19]. One explanation for this finding is that physicians may rely more heavily on patient input when assessing tenderness, whereas physical measurements of synovitis may play a greater role when assessing swelling. In addition, swelling may be difficult for patients to assess compared with pain. Typically, there is a higher correlation between physician and patient assessments with pain compared with swelling[18]. The weaker correlation seen with swollen joints may be interpreted as a limitation of the DAS.

When assessments of tender and swollen joints were compared individually, better agreement was generally observed for large joints than small joints and for tender joints than swollen joints. Although this trend was consistent among most joints, the knee joint assessment showed good agreement for both tenderness and swelling. The reason for this difference is unclear, but agreement was consistent over the study duration. Potentially, swelling may be easier to determine in the knee compared with other large joints, such as the shoulder.

Agreement between physician and patient assessments of tender and swollen joints remained stable over the course of the study. Although there can be considerable variation in determining joint involvement in patients with RA, training for both medical staff and patients has been shown to significantly increase the accurate detection of both tender and swollen joints[18,20]. However, even among trained medical professionals with RA experience, substantial variation has been shown to occur when different assessors examine the same patients[20]. A potential means of reducing variation in joint counts would be to use data collected from the patient, which might be more consistent because joint assessment would be performed routinely by only 1 person. In addition, appropriate patient training would be expected to increase the accuracy of joint assessments over time and with experience[18]. Although patient self-assessment may be beneficial for many patients, the limitations of some patients in performing this assessment should be considered. In general, patients tended to have higher joint counts than physicians, and despite repeated assessments over time, some patients in this study with large differences between physician assessment and patient self-assessment at baseline had similarly large differences in subsequent measurements. Perhaps some patient self-assessments may be affected by musculoskeletal conditions unrelated to RA disease activity, while physicians may tend to focus more on joint pain and swelling secondary to RA.

Another possible limitation to our results was the considerable average difference between physician and patient SJC at baseline (5.7) and Week 24 (4.4). This is potentially important given that SJC is a stronger predictor of structural damage than TJC[21]. The good correlation of physician and patient DAS may be due in part to the stronger agreement with TJC between physicians and patients and heavier weighting of tender versus swollen joint counts in the DAS28. However, patient and physician correlations in scales that equally weigh tender and swollen joint counts such as the CDAI and SDAI appear to have correlations similar to the DAS28. Nonetheless, the overall results from this analysis support the use of patient self-assessment of tender and swollen joint counts during RA treatment. Patient-derived joint counts may be beneficial when physician time and resources are limited, but they are not meant to replace the physician assessment. Patient self-assessment is likely to result in more frequent monitoring, which may allow more rapid medication adjustments by the physician, thereby improving patient outcomes[2]. Additionally, self-monitoring may increase patient involvement and prompt patients to seek changes in treatment earlier. Further studies are warranted to determine whether the Pt-DAS has predictive value for radiographic outcomes and functional status over time.

ACKNOWLEDGMENT

The authors thank Susan DeRocco, PhD, and Ali Hassan, PhD (supported by Amgen Inc.); and Marc Kubasak, PhD, and Edward Mancini, DPM, of Amgen Inc. for writing assistance. We also thank Nan Zhang, PhD, for statistical support in the preparation of this report.

REFERENCES

1. Wolfe F, Cush JJ, O'Dell JR, Kavanaugh A, Kremer JM, Lane NE, et al. Consensus recommendations for the assessment and treatment of rheumatoid arthritis. J Rheumatol 2001;28:1423-30.
2. Fransen J, van Riel PL. The Disease Activity Score and the EULAR response criteria. Clin Exp Rheumatol 2005;23:S93-9.
3. van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. Ann Rheum Dis 1990;49:916-20.
4. 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.
5. van der Heijde DM, van 't Hof MA, van Riel PL, van Leeuwen MA, van Rijswijk MH, van de Putte LB. Validity of single variables and composite indices for measuring disease activity in rheumatoid arthritis. Ann Rheum Dis 1992;51:177-81.
6. Welsing PM, van Gestel AM, Swinkels HL, Kiemeney LA, van Riel PL. The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. Arthritis Rheum 2001;44:2009-17.

7. Dore RK, Mathews S, Schechtman J, Surbeck W, Mandel D, Patel A, et al. The immunogenicity, safety, and efficacy of etanercept liquid administered once weekly in patients with rheumatoid arthritis. Clin Exp Rheumatol 2007;25:40-6.
8. Scott DL, van Riel PL, van der Heijde D, Benke AS. Assessing disease activity in rheumatoid arthritis. The EULAR handbook of standard methods. Vienna: European League Against Rheumatism; 1993.
9. 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 1999;38:48-52.
10. van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. Arthritis Rheum 1998;41:1845-50.
11. DAS-Score.NL. Nijmegen, The Netherlands: Department of Rheumatology, University Medical Center; 2009. [Internet. Accessed January 7, 2010.] Available from: http://www.das-score.nl/www.das-score.nl
12. Cole JC, Motivala SJ, Khanna D, Lee JY, Paulus HE, Irwin MR. Validation of single-factor structure and scoring protocol for the Health Assessment Questionnaire-Disability Index. Arthritis Rheum 2005;53:536-42.
13. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137-45.
14. Aletaha D, Nell VP, Stamm T, Uffmann M, Pflugbeil S, Machold K, et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. Arthritis Res Ther 2005;7:R796-806.
15. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. Rheumatology 2003;42:244-57.
16. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics 1977;33:363-74.
17. 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.
18. Levy G, Cheetham C, Cheatwood A, Burchette R. Validation of patient-reported joint counts in rheumatoid arthritis and the role of training. J Rheumatol 2007;34:1261-5.
19. Figueroa F, Braun-Moscovici Y, Khanna D, Voon E, Gallardo L, Luinstra D, et al. Patient self-administered joint tenderness counts in rheumatoid arthritis are reliable and responsive to changes in disease activity. J Rheumatol 2007;34:54-6.
20. Scott DL, Choy EH, Greeves A, Isenberg D, Kassinor D, Rankin E, et al. Standardising joint assessment in rheumatoid arthritis. Clin Rheumatol 1996;15:579-82.
21. Rau R, Herborn G, Menninger H, Sangha O. Radiographic outcome after three years of patients with early erosive rheumatoid arthritis treated with intramuscular methotrexate or parenteral gold. Extension of a one-year double-blind study in 174 patients. Rheumatology 2002;41:196-204.