View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Radboud Repository

Radboud Repository

Radboud University Nijmegen

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/23314

Please be advised that this information was generated on 2017-12-05 and may be subject to change.

Validity and Reproducibility of Self-Administered Joint Counts. A Prospective Longitudinal Followup Study in Patients with Rheumatoid Arthritis

MARLOU L.L. PREVOO, INA H. KUPER, MARTIN A. van't HOF, MIEK A. van LEEUWEN, LEVINUS B.A. van de PUTTE, and PIET L.C.M. van RIEL

ABSTRACT. ***Objective.*** To investigate the reproducibility and validity of self-administered joint counts (JC), measuring tenderness, swelling and the combination of both, in a longitudinal study.
Methods. At the outpatient department a form self-administered by patients (SAI-form), was used to measure joint involvement. Concurrent joint examinations were performed by an assessor. The JC and scores for groups of joints by assessors were correlated with those by patients. As a retest the form was completed again by the patients within 10 days. Correlations between the JC measured by the test and retest were computed to investigate reproducibility.
Results. Correlations between test and retest were high (> 0.7). Correlations between JC and groups of joints measured by the assessors and by the patients were moderate (0.6). Correlations with other disease activity variables did not differ between assessors' and patients' joint examination scores.
Conclusion. The patient-administered joint examination was reproducible; however, correlation with the assessors' joint examination was moderate. The value of the self-administered joint count needs further examination and cannot yet replace the assessor's joint examination. (*J Rheumatol 1996;***23:841–5**)

Key Indexing Terms:
ARTICULAR SCORES PROCESS VARIABLES RHEUMATOID ARTHRITIS

Rheumatoid arthritis (RA) is a systemic illness with chronic, symmetric polyarthritis as its most important clinical feature. Clinical evaluation of joints is considered important for the evaluation of an individual patient's disease in daily clinical practice as well as in clinical trials[1]. However, since there is no consensus on how to best measure joint involvement, many different types of joint counts (JC)[2–5] are in use. A recent study to evaluate several published JC for validity and reliability to measure disease activity[6] concluded that 28 JC that measure number of tender joints, swollen joints, or both tenderness and swelling were as valid and reliable as more comprehensive JC (i.e., JC with more joints and/or grading for tenderness, or weighting for surface area). Another study indicates that despite their simplified form, these JC were sensitive to changes in clinical trials in which 2 groups were compared[7]. Because of the validity, reliability, sensitivity, and simplicity of these 28 JC (measuring tenderness and/or swelling), the 28 JC were chosen to be part of a validated standard set of criteria for the assessment of activity and outcome of RA (EULAR core set)[8], and they were accepted as outcome measures in clinical trials by the American College of Rheumatology (ACR)[9]. Simpler and less time consuming would be a JC performed by the patient comparable to the self-administered visual analog scale (VAS) pain score, and functionality score by questionnaire. The self-assessment of joint involvement has been investigated in 5 studies[10–14], each evaluating a different questionnaire and a different JC. Our goal was to compare the validity and reproducibility of self-administered JC to assess the disease activity in joints.

MATERIALS AND METHODS

Patients already included in a followup study (begun in 1985) at the universities of Nijmegen and Groningen[6,15] were asked to participate in this followup study. To be eligible patients with recent onset RA visiting the outpatient department of the University of Nijmegen (Clinic 1) or Groningen (Clinic 2) had to fulfil the following criteria: RA according to the 1987 ACR criteria, disease duration less than one year, and no previous disease modifying antirheumatic drugs (DMARD). The participants, 173 patients from Clinic 1 and 101 patients from Clinic 2, were seen by research nurses and rheumatologists at least every 3 months.

In 1992, a self-administered joint evaluation form (SAJ) was developed. The SAJ form consisted of 2 mannequins to evaluate (1) swollen joints, and (2) tender joints (Figure 1). The following joints, indicated by a circle, were evaluated bilaterally for tenderness and swelling: the shoulder, elbow, wrist, each metacarpophalangeal (MCP), each proximal interphalangeal (PIP) of the fingers, interphalangeal of the thumb, hip, knee, ankle, and each metatarsophalangeal (MTP). On all SAJ forms instructions were included: patients were asked to indicate which joints were tender (graded

From the Department of Rheumatology, University Hospital Nijmegen, Department of Statistics, Nijmegen University, and the Department of Rheumatology, University Hospital Groningen.

Supported by a grant from Het Nationaal Reumafonds.

M.L.L. Prevoo, MSc, University Hospital Nijmegen; I.H. Kuper, MD, University Hospital Groningen; M.A. van Leeuwen, MD, Nijmegen University; M.A. van't Hof, PhD, Associate Professor, University Hospital Groningen; L.B.A. van de Putte, MD, Professor; P.L.C.M. van Riel, MD, Associate Professor, University Hospital Nijmegen.

Address reprint requests to Dr. M.L.L. Prevoo, Department of Rheumatology, University Hospital Nijmegen, PO Box 9101, 6500 HB Nijmegen, The Netherlands.

Submitted December 30, 1994 revision accepted November 22, 1995.

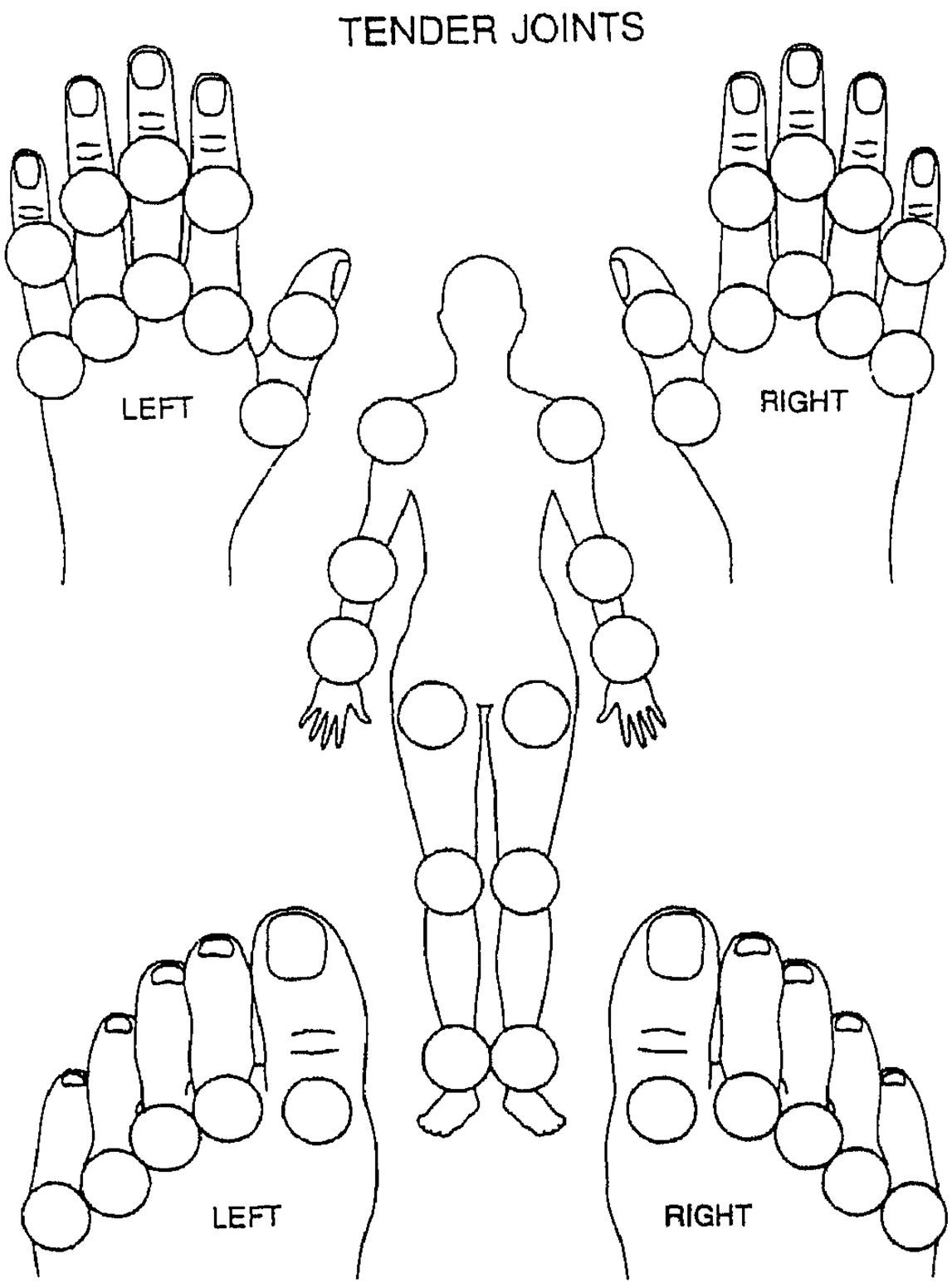

Figure 1. The self-administered form to indicate tender joints.

0 = no tenderness, 1 = slightly tender, 2 = moderately tender, 3 = very tender) or swollen (not graded, 0 = not swollen, 1 = swollen). Help with instructions was provided the first time patients completed the SAJ form.

In Clinic 1, patients completed the first SAJ form (TEST) during a visit, under supervision of one observer (MP), who had no knowledge of disease activity or joint involvement. Patients were asked to complete the same SAJ form at home one day after their visit, and to return the form by prepaid mail (RETEST). Subsequent SAJ were mailed to patients about 1 week before their 3 month visit. Forms were completed the morning of the visit, and collected by research nurses. In Clinic 2 patients completed the first SAJ form under supervision of research nurses or rheumatologists prior to joint examination. Patients completed subsequent SAJ during visits.

In both clinics, 3 month joint examination was given by research nurses or by rheumatologists (assessors). Joint scores were as follows: tenderness: (graded) 0 = not tender, 1 = tender, 2 = tender with wince response, 3 = tender with wince and withdraw response; swollen joints: (binary) 0 = not swollen, 1 = swollen. Other assessments include pain on a 10 cm VAS: 0 = no pain, 10 = worst pain possible; general health on a VAS: 0 = best possible, 10 = worst possible; grip strength by vigorimeter (mm Hg), erythrocyte sedimentation rate (ESR) according to Westergren. Every 6 months the patients completed a Health Assessment Questionnaire (HAQ)[16].

JC were calculated from assessor scores (assessors' JC) and from patient SAJ forms (SAJ-JC). Joints excluded because patients found them difficult to identify include temporomandibular, sternoclavicular, acromioclavicular, subtalar, midtarsal, and the cervical spine. The following JC were calculated: modified Ritchie Articular Index [42 joints examined for tenderness (graded)], the total number of swollen joints (40 joints), total number of tender joints (42 joints), total number of joints that are tender as well as swollen (40 joints), as well as the Thompson count (THOM, 38 joints examined for the combination of tenderness and swelling, in which the joints were weighted for surface area), 28 joint count measuring tender joints, 28 swollen joint count, and the combination of tenderness and swelling (28 T&S). Definitions of the JC are shown in Table 1.

Statistics. The reproducibility of the SAJ forms was determined by comparing results of TEST versus RETEST. After square root transformation of JC to obtain "normality," Pearson correlation coefficients between TEST and RETEST JC were computed.

To investigate the validity of the SAJ-JC to measure joint involvement, assessor JC were considered the gold standard. Systematic differences between JC measured by patients (SAJ) and those measured by assessors were tested by Wilcoxon paired signed rank tests. Regression analyses were performed to control for systematic differences. Pearson correlation coefficients between JC of patients and assessors were computed.

RESULTS

Of 144 patients in Clinic 1 asked to participate 3 patients were unable to complete the SAJ form. A total of 141 patients (98%) completed the first SAJ form. After completing the first SAJ form, these 141 patients were sent a total of 327 SAJ forms of which 284 (87%) were returned. Some SAJ forms were not filled in correctly (for instance, grading for tenderness was missing) or the concurrent scores of the assessors were not available. A total of 386 SAJ forms

Table 1. Definitions for assessor and patient JC.

Joints	No. of Joints	RAI	THOM (*weights*)	28 T	28 T&S	28 S	TOTT	TOTT&S	TOTS
Shoulder	2	+	–	+	+	+	+	+	+
Elbow	2	+	+ (48)	+	+	+	+	+	+
Wrist	2	+	+ (32)	+	+	+	+	+	+
MCP (1–5)	10	+	+ (5)	+	+	+	+	+	+
PIP (1–5)	10	+	+ (3)	+	+	+	+	+	+
Hip	2	+	–	–	–	–	+	–	–
Knee	2	+	+ (95)	+	+	+	+	+	+
Ankle	2	+	+ (32)	–	–	–	+	+	+
MTP (1–5)	10	+	+ (8)	–	–	–	+	+	+

RAI: Ritchie Articular Index; THOM: Thompson articular score; 28 T: 28 JC measuring tender joints; 28 T&S: 28 JC measuring tenderness and swelling; 28 S: 28 JC measuring swollen joints; TOTT: total number of tender joints; TOTT&S: total number of joints measuring tenderness and swelling; TOTS: total number of swollen joints.

(91%) of all SAJ forms (n = 425) could be used for further analyses. Patients who completed the first SAJ form (n = 141), 111 (83%) completed retests within 10 days of the first test. At study start (at first SAJ) the median disease duration of these 141 patients was 3.8 years (25–75% interval: 2.8–5.3 years), their median age was 59 years (25–75% interval: 51–67 years), 59% were women, and 81% were IgM rheumatoid factor positive (IgM RF > 10 IU). In Clinic 2, 95 patients participated and the total number of SAJ forms was 231 (no retest). Their median disease duration was 5.6 years (25–75% interval: 4.4–6.3 years); median age was 56 years (25–75% interval: 41–63 years); 69% were women, and 83% IgM RF positive.

Correlations between TEST and RETEST JC ranged from 0.77 to 0.87. Correlations measuring a combination of tenderness and swelling [THOM (0.77), 28 T&S (0.78)] were lowest.

Median values of assessor and JC patient for both clinics are shown in Table 2. All JC except 28 measuring swollen joints are significantly different (systematic difference) between the assessors and the patients (p value Wilcoxon's signed rank tests < 0.05). In both clinics an overestimation of patient versus assessor JC was observed. Pearson correlation coefficients between assessor JC (gold standard) and patient JC are presented in Table 3. The range of correlations of all JC was 0.48–0.65 for Clinic 1 and 0.47–0.65 for Clinic 2. Also, correlations between assessor JC and SAJ-JC were computed separately for 3 observers (data not shown). These did not differ from the overall correlations.

To identify whether the low correlations between assessors and patients could be explained by specific joints, we determined assessor–patient agreement for groups of joints. To group the large number of joints, principal component analyses were performed separately for scores measuring swelling and tenderness (nongraded, binary). For both characteristics, similar groups for the patients and assessors (eigenvalue > 1) were found. The following 7 groups could be formed: MCP, PIP, MTP, and all large joints (wrists, ankles, knees, shoulders). In further analyses these large joints are grouped together. The reliability of the 4 groups

Table 3. Pearson correlations between the assessor and patient JC.

	Nijmegen Clinic 1 (n = 386)	Groningen Clinic 2 (n = 229)
RAI	0.62	0.65
THOM	0.48	0.47
28 T	0.62	0.60
28 T&S	0.56	0.61
28 S	0.61	0.65
TOTT	0.65	0.60
TOTT&S	0.53	0.61
TOTS	0.51	0.64

RAI: Ritchie Articular Index; THOM: Thompson articular score; 28 T: 28 JC measuring tender joints; 28 T&S: 28 JC measuring tenderness and swelling; 28 S: 28 JC measuring swollen joints; TOTT: total number of tender joints; TOTT&S: total number of joints measuring tenderness and swelling; TOT: total number of swollen joints.

measured by the assessors or patients both for tender and swollen joints was determined by Cronbach's alphas ranging as follows: MCP 0.84–0.92, PIP 0.87–0.93, MTP 0.91–0.94, and large joints 0.58–0.86. The correlations (nonparametric, Spearman rank correlations) between the mean scores of the groups of joints scored by the assessor and by the patients are presented in Table 4. The correlation for swelling of the MTP is low, and differs significantly ($p < 0.05$) from the correlations of all other groups of joints.

As low correlations were found between assessor and patient JC and groups of joints, we computed correlations of patient and assessor scores with other disease activity vari-

Table 4. Spearman correlations between assessor and patient (n = 615) scores for groups of joints.

	Spearman Correlations between Assessors and Patients	
	Tenderness	Swelling
MTP	0.48	0.25*
PIP	0.51	0.56
MCP	0.50	0.57
LJ	0.52	0.54

* Significantly ($p < 0.05$) different from the other correlations.
LJ: group of joints including wrists, ankles, knees, shoulders, elbows.

Table 2. Median values, differences and their (p10, p90) of assessor and patient JC.

	Nijmegen Clinic 1 (n = 386)				Groningen Clinic 2 (n = 229)			
JC	Assessor JC	Patient JC	Median Difference Assessor-Patient	p*	Assessor JC	Patient JC	Median Difference Assessor-Patient	p*
RAI	3 (0, 10)	10 (0, 26)	–5 (–18, 0)	< 0.01	5 (0, 18)	5 (0, 21)	0 (–9, 5)	0.02
THOM	17 (0, 193)	57 (0, 288)	–16 (–212, 38)	< 0.01	5 (0, 98)	12 (0, 175)	0 (–136, 32)	< 0.01
28 T	2 (0, 10)	7 (0, 24)	–4 (–16, 0)	< 0.01	2 (0, 13)	3 (0, 18)	0 (–7, 4)	< 0.01
28 T&S	1 (0, 7)	3 (0, 14)	–1 (–8, 2)	< 0.01	1 (0, 6)	1 (0, 10)	0 (–5, 1)	< 0.01
28 S	5 (1, 14)	4 (0, 16)	0 (–7, 7)	0.22	2 (0, 11)	2 (0, 13)	0 (–6, 4)	0.07

* Wilcoxon signed rank test; Due to ranking of values (patients have more extreme values) the signed rank test leads to significant results despite zero median values. RAI: Ritchie Articular Index; THOM: Thompson articular score; 28 T: 28 JC measuring tender joints; 28 T&S: 28 JC measuring the combination of tenderness and swelling; 28 S: 28 JC measuring swollen joints.

ables. For this correlation between patient and assessor scores for groups of joints with ESR, Pain on a VAS pain score, General Health on a VAS general health, grip strength, and HAQ were computed (Table 5). All Spearman rank correlations of scores for groups of joints with these variables do not significantly differ between assessors and patients. Similar results were found if tenderness was graded (not binary). There were no differences in correlations between the 2 clinics.

DISCUSSION

In the assessment of disease activity in patients with RA, there is an increase in the use of self-administration forms, such as questionnaires to assess function, VAS to measure pain, general health, and global disease activity. In recent years, self-administered forms to measure joint involvement including JC[10–14] have been evaluated. Although an attempt was made to test the validity of these different SAJ-JC, a comparison of SAJ-JC with assessor JC has never been made. We evaluated available SAJ-JC, including 28 JC recently shown to be reliable and valid measurements of disease activity[6], in a large cohort of patients with RA.

In view of the level of difficulty of the SAJ forms the response rate to the SAJ form was high. Since most patients were capable of filling in the SAJ form correctly, the number of forms used in the analyses was high.

Reproducibility of the self-administered JC appeared satisfactory, and comparable to test–retest correlations of other assessments[16–18].

However, the validity of SAJ-JC with respect to the assessor JC is low. This low correlation could not be explained by differences in the number of joints, in weighting of joints, and/or grading of tenderness (comprehensive JC showed as low a correlation as 28 JC). No difference resulted from the way the SAJ form was introduced and explained to patients in Clinic 1 versus 2. Finally, no group of joints could explain these low correlations. If the purpose of the self-administered score is to replace the assessor score, correlation between these scores should be higher than 0.9, and, after correction for systematic differences, scores should be almost identical. On the basis of our results we conclude therefore that joint involvement measured by assessors cannot be replaced by patient self-administered scores. In 4 of 5 studies evaluating self-administered JC it was concluded that validity of self-administered JC to measure assessor JC was good[10–13]; however, correlations in these studies range from 0.40 to 0.89. Only Hewlett, *et al* concluded correctly that the validity to assess assessor JC was low[14], in agreement with our results.

We found the correlation between assessor and patient scores for swelling of the MTP was low, and differs significantly ($p < 0.05$) from the correlations of all other groups of joints. Smolen, *et al*[19] found a clear cut dissociation between the measurement of joint pain and swelling of the MTP joints by assessors, which could indicate that the assessment of swelling in MTP joints is more difficult than in other joints; or MTP joint tenderness may reflect processes different from those of the underlying disease. The low correlation between the assessor and patient joint scores for swelling of MTP confirms that the measurement of these scores differs from all other joints.

Due to the low validity versus assessor scores, the relation of the self-administered scores with some other disease variables was computed. The relation of the SAJ scores with these disease variables did not differ from the assessor scores.

In our study the assessor score was taken as the gold standard. Agreement between patient and assessor scores was low compared to correlation coefficients for the Ritchie score (range 0.91 to 98) between 3 assessors[3]. However, correlation of patient scores with other disease activity variables was comparable with that of assessor scores with these variables. Thus, there should be further investigation to determine whether patient scores would be useful in clinical trials and in daily clinical practice.

We conclude that the reproducibility of the self-adminis-

Table 5. Spearman rank correlations between patient and assessor scores for groups of joints (patient–assessor) versus other disease activity variables.

	ESR (n = ± 615)		Pain (n = ± 615)		GH (n = ± 386)		Grip Strength (n = ± 615)		HAQ (n = ± 190)	
	Assessor	Patient	Assessor	Patient	Assessor	Patient	Assessor	Patient	Assessor	Patient
Tenderness										
MTP	0.01	0.12	0.34	0.23	0.36	0.25	0.20	0.20	0.32	0.37
PIP	0.11	0.14	0.39	0.35	0.38	0.35	0.27	0.27	0.37	0.37
MCP	0.08	0.16	0.43	0.36	0.41	0.36	0.33	0.26	0.42	0.29
LJ	0.06	0.17	0.50	0.44	0.45	0.36	0.32	0.34	0.50	0.39
Swelling										
MTP	0.07	0.03	0.28	0.23	0.23	0.14	0.10	0.22	0.16	0.18
PIP	0.02	0.09	0.22	0.29	0.18	0.15	0.23	0.25	0.19	0.15
MCP	0.06	0.14	0.33	0.35	0.24	0.13	0.26	0.24	0.21	0.14
LJ	0.07	0.21	0.38	0.32	0.28	0.19	0.31	0.33	0.39	0.24

Pain: VAS measuring pain; GH: VAS measuring general health; LJ: group of joints including wrists, ankles, knees, shoulders, elbows.

tered forms to determine joint involvement was good in this population. Agreement between health professionals and patients was low for JC as well as for groups of joints. The relation of self-administered scores for groups of joints with other disease variables, however, did not differ from the assessor scores. Although self-administered joint involvement scores cannot yet replace the joint examination by health professionals, further studies should investigate their place in the assessment of RA.

REFERENCES

1. van der Heide A, Jacobs JWG, Dinant HJ, Bijlsma JWJ: The impact of endpoint measures in rheumatoid arthritis clinical trials. *Semin Arthritis Rheum 1992;*21:287–94.
2. Lansbury J, Haut DD: Quantitation of the manifestations of rheumatoid arthritis. 4. Area of joint surfaces as an index to total joint inflammation and deformity. *Am J Med Sci 1956;*232:150–5.
3. Ritchie DM, Boyle JA, McInnes JM, *et al:* Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. *Q J Med 1968;*37:393–406.
4. Fuchs HA, Brooks RH, Callahan LF, Pincus T: A simplified twenty-eight-joint quantitative articular index in rheumatoid arthritis. *Arthritis Rheum 1989;*32:531–7.
5. Thompson PW, Silman AJ, Kirwan JR, Currey HLF: Articular indices of joint inflammation in rheumatoid arthritis. *Arthritis Rheum 1987;*27:618–23.
6. Prevoo MLL, van Riel PLCM, van't Hof MA, *et al:* Validity and reliability of joint indices. A longitudinal study in patients with recent onset rheumatoid arthritis. *Br J Rheumatol 1993;*32:589–94.
7. Fuchs HA, Pincus T: Reduced joint counts in controlled clinical trials in rheumatoid arthritis. *Arthritis Rheum 1993;*37:470–5.
8. van Riel PLCM: Provisional guidelines for measuring disease activity in RA clinical trials (editorial). *Br J Rheumatol 1992;*31:793–4.
9. Felson D, Anderson J, Boers M, *et al:* Reduced joint counts in rheumatoid arthritis clinical trials. *Arthritis Rheum 1994;*37:463–4.
10. Mason JH, Anderson JJ, Meenan RF, Haralson KM, Lewis-Stevens D, Kaine JL: The rapid assessment of disease activity in rheumatology (RADAR) questionnaire. *Arthritis Rheum 1992;*35:156–62.
11. 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.
12. Pincus T, Callahan LF, Brooks RH, Fuchs HA, Olsen NJ, Kaye JJ: Self-report questionnaire scores in rheumatoid arthritis compared with traditional physical, radiographic, and laboratory measures. *Ann Intern Med 1989;*110:259–66.
13. Abraham N, Blacmon D, Jackson JR, Bradley LA, Lorish CD, Alarcón GS: Use of self-administered joint counts in the evaluation of rheumatoid arthritis. *Arthritis Care Res 1993;*6:78–81.
14. Hewlett SE, Haynes J, Shepstone L, Kirwan JR: Patients' ability to report signs of disease activity in rheumatoid arthritis (abstr). *Br J Rheumatol 1994;*(suppl 2)33:21.
15. van der Heijde DMFM, van't Hof MA, van Riel PLCM, *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.
16. van der Heijde DMFM, van Riel PLCM, van der Putte LBA: Sensitivity of a Dutch health assessment questionnaire in a trial comparing hydroxychloroquine vs. sulphasalazine. *Scand J Rheumatol 1990;*19:407–12.
17. Meenan RF, Mason JH, Anderson JJ, Guccione AA, Kazis LE: AIMS2. The content and properties of a revised and expanded arthritis impact measurement scales health status questionnaire. *Arthritis Rheum 1992;*35:1–10.
18. 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 analogue scale. *Arthritis Rheum 1987;*30:630–6.
19. Smolen JS, Breedveld FC, Eberl G, Jones I, Leeming M, Wylie GL, Kirkpatrick J: Validity and reliability of the twenty-eight-joint count for the assessment of rheumatoid arthritis activity. *Arthritis Rheum 1995;*1:38–43.