Systematic Review and Metaanalysis of Patient Self-Report versus Trained Assessor Joint Counts in Rheumatoid Arthritis

JENNIFER L. BARTON, **LINDSEY A. CRISWELL**, **RACHEL KAISER**, **YEA-HUNG CHEN**, and **DEAN SCHILLINGER**

Department of Medicine, University of California, San Francisco, and the San Francisco Department of Public Health, San Francisco, California, USA

Abstract

Objective—Patient self-report outcomes and physician-performed joint counts are important measures of disease activity and treatment response. This metaanalysis examines the degree of concordance in joint counts between trained assessors and patients with rheumatoid arthritis (RA).

Methods—Studies eligible for inclusion met the following criteria: English language; compared patient with trained assessor joint counts; peer-reviewed; and RA diagnosis determined by board-certified or board-eligible specialist or met 1987 American College of Rheumatology criteria. We searched PubMed and Embase to identify articles between 1966 and January 1, 2008. We compared measures of correlation between patients and assessors for either tender/painful or swollen joint counts. We used metaanalysis methods to calculate summary correlation estimates.

Results—We retrieved 462 articles and 18 were included. Self-report joint counts were obtained by a text and/or mannequin (picture) format. The summary estimates for the Pearson correlation coefficients for tender joint counts were 0.61 (0.47 lower, 0.75 upper) and for swollen joint counts 0.44 (0.15, 0.73). Summary results for the Spearman correlation coefficients were 0.60 (0.30, 0.90) for tender joint counts and 0.54 (0.35, 0.73) for swollen joint counts.

Conclusion—A self-report tender joint count has moderate to marked correlation with those performed by a trained assessor. In contrast, swollen joint counts demonstrate lower levels of correlation. Future research should explore whether integrating self-report tender joint counts into routine care can improve efficiency and quality of care, while directly involving patients in assessment of RA disease activity.

Key Indexing Terms

RHEUMATOID ARTHRITIS; JOINTS; METAANALYSIS

Self-reported outcomes are important measures of disease activity and treatment response in clinical trials and practice in rheumatoid arthritis (RA). Systematic assessment of swelling and tenderness in joints, or joint counts, by physicians has been cited as the most specific measure of disease activity in RA and has been shown to be predictive of mortality[1]. Joint counts are included in standard composite measures of disease activity routinely used in clinical trials

Address correspondence to Dr. J. Barton, Division of Rheumatology, University of California, 374 Parnassus Avenue, Box 0500, San Francisco, CA 94143. Jennifer.Barton@ucsfmedctr.org.

J.L. Barton, MD, Assistant Professor; L.A. Criswell, MD, MPH, Professor; R. Kaiser, MD, MPH, Assistant Professor; D. Schillinger, MD, Professor, Department of Medicine, University of California, San Francisco; Y. Chen, MS, San Francisco Department of Public Health.

[Disease Activity Score-28 (DAS28) and the American College of Rheumatology (ACR) Core Data Set]. Formal joint counts, however, are inconsistently performed in clinical practice[2].

Since the early 1990s, multiple studies have assessed the reliability, validity, and sensitivity to change of self-report joint counts. There are many potential advantages of a self-report joint count. Involving patients in disease activity assessment may enhance self-management behavior, and ultimately improve health outcomes. Self-management programs in arthritis have been shown to improve health status, reduce pain and fatigue, and increase self-efficacy[3,4]. A self-report joint count in RA, if shown to be reliable, could serve as an important marker of disease activity over the course of RA, since, in the collaborative model of chronic care, it is ultimately the patient, rather than any one clinician, who is responsible for continuing self-care and decision-making. Active engagement with one's chronic disease has been shown to be associated with health improvements[5–7]. Last, as the opportunities for Web-based self-management support and clinical decision-making across chronic diseases increase, there is an urgent need to explore reliable means of eliciting symptoms from patients with RA so that advances in disease management that are being applied in diseases such as congestive heart failure, diabetes, and asthma can also be adapted to RA.

Self-report in RA

Self-reported outcomes in RA have become central to the measurement of response to treatment both in clinical trials and in routine practice. Over 10 countries have published guidelines or consensus statements on the use of biologic therapies in RA that include a composite score to measure disease activity, such as the DAS28[8]. New ACR recommendations for use of biologics in the US also promote the use of a disease activity score, many of which include a formal joint count[9]. Choy, *et al* have developed and validated a patient-based disease activity score (PDAS) for RA that incorporates a self-report tender and swollen joint count[10]. However, this tool has yet to be widely adopted or tested in clinical practice.

Can patients reliably report on joint tenderness and inflammation? While some studies suggest that self-report tender joint counts are reliable, others have not, and no systematic review of the research has been performed. Because of the discrepancy between physician and patient joint counts observed in some studies, and the potential positive influence of direct patient involvement through the use of self-report of disease activity, we undertook a systematic review of the literature and performed a metaanalysis to examine the degree of concordance between patients and trained assessors around this important clinical measure of disease activity in RA.

MATERIALS AND METHODS

Search strategy

Three search strategies were used to identify appropriate articles between 1966 and January 1, 2008. We searched PubMed, Embase, and the Cochrane databases by combining the following 3 searches with AND to identify all definitions of joint counts and self-report, as well as terms for RA: (1) rheumatoid arthritis; (2) joint OR joints OR articular OR disease activity; and (3) self-report* OR self-assess* OR self-monitor* OR self-administ* OR self-evalua* OR self-perce* OR self-examin* OR self-rate OR self-rating.

Bibliographies of all included studies were searched as well. A single investigator (JLB) performed the PubMed, Embase, and Cochrane Library searches and reviewed the titles and abstracts identified by the search.

Literature search limits and article selection criteria. Inclusion and exclusion criteria

Studies eligible for inclusion met the following criteria: (1) objective was to compare patient with trained assessor counts of joint swelling and/or tenderness and pain; (2) appeared in peer-reviewed publication; (3) included subjects with RA as determined by a board-certified or board-eligible specialist or as defined by 1987 ACR criteria[11]; and (4) were English language. Studies were excluded if they were letters to the editor, case reports, or review articles. Studies were also excluded based on abstract review if they did not meet inclusion criteria.

Abstraction of articles

Data abstraction was performed by 2 authors (JLB, RK). In situations where there was disagreement, an adjudicator (LAC) intervened. The following information was abstracted systematically from each article: author, year, country, setting (type of outpatient clinic or inpatient), study design, duration of study, ACR criteria for RA met, number of subjects, covariates included in study (age, sex, disease duration, inflammatory markers, medications, ethnicity, etc.), assessor type (physician, nurse, research assistant), format of joint count (mannequin or text), main outcome (correlation coefficient), association between correlation coefficient and other covariates, and predictors of concordance.

Quality assessment

Quality of included articles was assessed by ensuring that studies met the inclusion criteria listed above. These qualitative measures were chosen instead of a quality score because such scales have been shown to be incomplete and unreliable due to heterogeneity between studies[12].

Statistical methods

The primary outcome measure was a summary correlation coefficient for both tender/painful joint counts and swollen joint counts. Given that studies used varying measures of correlation, we chose to calculate a summary Pearson correlation coefficient for tender/painful joint counts as well as one for swollen joint counts for all studies that reported a Pearson correlation coefficient. For those studies that reported Spearman correlation coefficients, we separately calculated the summary estimate for both types of joint counts as well. We were unable to include 3 studies in the metaanalyses that reported an intraclass correlation coefficient (ICC) as the measure of correlation[13–15]. This precluded calculation of a summary ICC since, to our knowledge, there is no established method for performing metaanalysis using the ICC. Metaanalysis of correlation coefficients was conducted via methods described by Hunter and Schmidt[16]. Summary estimates were computed by weighting the sample size. We did not correct for any "design artifacts," which are defined by Hunter and Schmidt as factors that may affect the size of the correlation coefficient, such as sampling or measurement error[16]. We produced, for each analysis, a summary correlation and confidence interval for the estimate. All analyses were conducted using R[17]. Interpretation of correlation coefficients was based on published thresholds[18] as follows: r zero to 0.2, no correlation; 0.2 to 0.4, low correlation; 0.4 to 0.6, moderate correlation; 0.6 to 0.8, marked correlation; and 0.8 to 1.0, high correlation.

Because reliability of self-report joint counts could be influenced by mode of elicitation of joint symptoms, we performed a subgroup analysis by separately exploring correlations when the self-report was obtained using a mannequin format (Figure 1) from those correlations when self-report was obtained using a text format. We assessed for the presence of publication bias using the Begg's test, which assesses the significance of the correlation between the ranks of effect estimates and the ranks of their variances; p value < 0.1 was considered significant[19, 20].

RESULTS

The PubMed search identified 364 articles and the Embase search identified an additional 98 for a total of 462 articles. The Cochrane Library search yielded no additional articles. Four hundred forty articles were excluded based on a review of the title and abstract, leaving 22 articles for full review of the text. After full review, 18 articles remained in the final analysis (Table 1). The 4 studies were excluded based on the following: (1) no direct comparison of patient with assessor joint count (n = 2); (2) duplicate patient population (n = 1); (3) lack of assessor joint count (n = 1).

Methodologic qualities of the studies

Methods across the 18 studies were heterogeneous. We abstracted data on blinding of assessors, type of joint counts performed (tender or swollen), format of self-report (mannequin or text), and measure of correlation between assessor and patient (Pearson, Spearman, or ICC). Of the 18 studies, 9 reported blinding of assessors, 15 reported tender joint counts, and 10 reported swollen joint counts. Three studies reported a combined tender/swollen joint count[21−23]. Ten elicited self-report using text format and 8 using the mannequin. Three studies directly compared the 2 self-report methods[13,24,25].

A minority of the studies described some features of validity of the self-report joint counts that included assessments of reliability, sensitivity to change, and correlation with acute-phase reactants. Seven studies reported on intrarater reliability (test and retest self-report)[14,22,23, 25−28] and all demonstrated moderate to high correlation. In terms of sensitivity to change, only 5 reports provided results, with the majority of these noting sensitivity to change for the self-report tender joint count, but not for the swollen joint count over time[14,15,22,26,29]. The range of time elapsed between joint assessments was 0.5 to 10 months, with 3–6 months being the average. Ten studies evaluated the correlation between self-report and assessor joint counts with an acute-phase reactant[14,21−23,27−32]. In the majority of cases, these studies concluded that self-report joint counts do not correlate with acute-phase reactants and that self-report swollen joint count is not an adequate substitution for a physician examination to detect swollen or inflamed joints.

Summary correlation coefficients

The correlation coefficients between patient and assessor tender and swollen joint counts of the included articles are shown in Tables 2 (tender) and 3 (swollen). The 3 studies that reported on a combined tender and swollen joint count were not included in the summary results[21−23]. The summary estimate for Pearson correlation coefficients for tender joint counts was 0.61 (0.47 lower, 0.75 upper) and for swollen joint counts, 0.44 (0.15 lower, 0.73 upper). The Spearman summary estimate for tender joint counts was 0.60 (0.30 lower, 0.90 upper) and for swollen joint counts, 0.54 (0.35 lower, 0.73 upper). Figure 2 displays Forrest plots for these summary estimates. These plots indicate moderate to marked correlation for tender joint counts, but only moderate correlation for swollen joint counts.

There was no evidence of publication bias for any of the 4 summary estimates using Begg's test.

Overall, the range of correlation coefficients for tender joint counts fell into the moderate to high range, 0.45 to 0.92. The mannequin format yielded higher scores (0.54 to 0.92) compared to the text format for reporting tender joint counts (0.45 to 0.89). Overall, the swollen joint count correlation coefficients were lower, with a range of 0.16 to 0.67 (no to marked correlation). The mannequin format (range of 0.31 to 0.67) again fared better than text (0.16

to 0.63) for swollen joint counts. Overall, the mannequin format had higher correlation coefficients than the text format when the 2 were compared within and across studies.

Patient characteristics and predictors of patient-assessor differences

Only one of the 18 studies reported on predictors of patient-physician differences in joint counts[25]. Using regression analysis, the authors found that greater age, worse patient global disease rating, and poorer function (but not educational attainment) were significantly associated with discordance in the mannequin self-report format of tender joint counts. There were no important variables associated with differences in the mannequin swollen joint counts. With respect to the text format, longer disease duration predicted greater physician-patient differences in tender joint counts, and education level for swollen joint counts. For example, for patients with less than complete high school education the average percentage difference in swollen joint count scores was 21% and for those with more than a high school education, there was 0% difference[25].

Six studies collected data on patient education level[13,14,24,25,28,33]. Three of these reported less than high school education in at least half or more of the subjects[13,14,24]. Two of these studies that included only tender joint counts using both mannequin and text formats reported moderate to high correlation (0.55 to 0.77) between self-report and assessors for both methods, and the third reported a marked degree of correlation for tender joint counts using a text method (ICC = 0.78), but low correlation for swollen joint counts (ICC = 0.31)[14]. Race/ethnicity data were collected in only 5 studies, with 4 of the 5 including 70% or more Caucasian subjects[13,14,28,29,33]. Nearly all the studies reported a majority (> 70%) of female subjects. Only one report highlighted the importance of literacy[14]. In that study, the mannequin format (ICC = 0.64) performed better than text (ICC = 0.55) for self-report of joint tenderness or pain.

DISCUSSION

Our systematic review and metaanalysis demonstrates that patient self-report tender joint counts have moderate correlation with those of a trained assessor or physician. Swollen joint counts demonstrate lower levels of correlation.

Since the 1950s, joint counts have served as an essential component of disease activity measurement in clinical trials, and less consistently in clinical practice, despite the fact that a joint count is part of the ACR Quality Indicator Set[34]. A self-report articular index could provide an efficient means to directly involve patients in the assessment of disease activity in RA and could benefit both patient and practitioner. Benefits include savings of time and expense, simplicity of use by the patient, and a reliable self-report measure to assess disease activity and response to treatment in clinical trials[15,26].

While physician interpretation of tender or swollen joints may differ from that of the patient, one could argue that valuable information is derived from both observers and that one perspective is not "better" than the other. Reliable data on swelling and inflammation may best be gathered from the physician, but a more meaningful and accurate description of pain or tenderness can arguably be derived from the patient. As evidence that patient-reported outcomes are just as important, if not more important indicators of response to therapy, Ward[35] assessed 14 different clinical measures (both physician and/or patient-reported, but no joint counts) in a group of 24 subjects with RA to determine which were most accurate and responsive to change. Of the 14 measures tested, patient and physician global assessments, patient pain score, and a patient self-report disability index were more sensitive to change over time than other measures (including physician-reported and laboratory measures). Three of these are obtained via patient self-report[35].

While the effects of greater patient involvement in symptom reporting with respect to health outcomes have not been fully investigated in RA, self-management programs that target self-efficacy, informed decision-making, and communication in arthritis have been shown to reduce pain, fatigue, and health distress[3]. Another potential benefit of self-report joint counts in RA is the elimination of interobserver variation by removal of different physicians or nurses in the assessment of joints over time. And while our study demonstrated lower correlations on swollen joint counts, the study by Levy, *et al* demonstrated that a brief training session (5 minutes) helped patients discriminate between a truly swollen joint and a chronically enlarged joint and greatly enhanced the accuracy of a self-report swollen joint count[36]. It is also possible that a Web-based, touch-screen questionnaire could lead to the elimination of data entry, provide immediate access to results, and be an acceptable or even preferable option for patients to report joint counts as well as other measures[32].

We found some evidence that education level, a potent marker for poor health communication, may not pose a barrier to accurately reporting tender joint counts, especially if the format used is pictorial (mannequin)[13,24]. When low literate and non-English-speaking subjects have been asked to recall medication regimens, such as warfarin doses, reporting was more accurate using a visual representation of their dosages, providing support for the notion that a mannequin or pictorial representation of joints may be more accurate across different language and literacy levels[37]. Athale, *et al* studied a Web-compatible instrument to collect self-report measures in RA (including painful and swollen joint counts) and found moderate to high correlation between paper and computer-reported measures using a mannequin format[38].

The studies included in this review presented a number of weaknesses. As a whole, there was a lack of uniformity in several areas. The selection of a measurement tool to formally assess the joint count varied significantly among studies, including variability of number of joints counted, format of report, different counts between assessor and patient, and the combination of pain and swelling on some indices. The measure of correlation between assessor and self-report counts (Pearson, Spearman, or ICC) was not consistent across studies. The level of disease activity and the number of joints involved was for the most part low, but not uniform. And last, patient characteristics such as education level, disease duration, and level of knowledge about RA were rarely reported. The lack of consistency among studies was a limiting factor for the metaanalysis.

Our study has several limitations. First, the lack of conformity among studies with regard to patient characteristics did not allow a comprehensive determination of how they correlate with differences between patient self-report and assessor tender or swollen joint counts. It is possible that self-report joint counts of patients with more severe disease, longer disease duration, or greater disability may be systematically different than assessor or physician joint counts.

Second, because of a paucity of information from studies and the relative lack of diversity of enrolled subjects, we were not able to establish whether the reliability estimates are robust across education, age, race, language, or literacy levels. Future studies should include a diverse population to allow a broader and more generalizable application of a self-report joint count.

This systematic review and metaanalysis demonstrates that self-report tender joint counts in RA have moderate to marked correlation with those of a trained assessor or physician. Self-report joint counts have been incorporated into a patient-based disease activity score (the PDAS) that has been proven to be valid and sensitive, but has yet to be widely adopted[10]. Use of a mannequin self-report form (as in the PDAS) appears to enhance correlation. Swollen joint counts have at most moderate correlation between patients and assessors. Based on these results, a self-report tender joint count using visual aids, as proposed in the PDAS, may provide an efficient means to directly involve patients in assessment of RA activity. Studies are needed

to assess the feasibility and effectiveness of incorporating a self-report disease activity score into an interactive, self-management program in clinical practice.

Acknowledgments

References

1. Pincus T, Brooks RH, Callahan LF. Prediction of long-term mortality in patients with rheumatoid arthritis according to simple questionnaire and joint count measures. Ann Intern Med 1994;120:26–34. [PubMed: 8250453]

2. Pincus T, Segurado OG. Most visits of most patients with rheumatoid arthritis to most rheumatologists do not include a formal quantitative joint count. Ann Rheum Dis 2006;65:820–2. [PubMed: 16291813]

3. Osborne RH, Wilson T, Lorig KR, McColl GJ. Does self-management lead to sustainable health benefits in people with arthritis? A 2-year transition study of 452 Australians. J Rheumatol 2007;34:1112–7. [PubMed: 17343319]

4. Lorig K, Gonzalez VM, Ritter P. Community-based Spanish language arthritis education program: a randomized trial. Med Care 1999;37:957–63. [PubMed: 10493473]

5. McPherson ML, Smith SW, Powers A, Zuckerman IH. Association between diabetes patients' knowledge about medications and their blood glucose control. Res Social Adm Pharm 2008;4:37–45. [PubMed: 18342821]

6. Guerci B, Drouin P, Grange V, Bougneres P, Fontaine P, Kerlan V, et al. Self-monitoring of blood glucose significantly improves metabolic control in patients with type 2 diabetes mellitus: the Auto-Surveillance Intervention Active (ASIA) study. Diabetes Metab 2003;29:587–94. [PubMed: 14707887]

7. Greenfield S, Kaplan SH, Ware JE Jr, Yano EM, Frank HJ. Patients' participation in medical care: effects on blood sugar control and quality of life in diabetes. J Gen Intern Med 1988;3:448–57. [PubMed: 3049968]

8. Lopez-Olivo MA, Kallen MA, Ortiz Z, Skidmore B, Suarez-Almazor ME. Quality appraisal of clinical practice guidelines and consensus statements on the use of biologic agents in rheumatoid arthritis: a systematic review. Arthritis Care Res 2008;59:1625–38.

9. Saag KG, Teng GG, Patkar NM, Anuntiyo J, Finney C, Curtis JR, et al. American College of Rheumatology 2008 recommendations for the use of nonbiologic and biologic disease-modifying antirheumatic drugs in rheumatoid arthritis. Arthritis Rheum 2008;59:762–84. [PubMed: 18512708]

10. Choy EH, Khoshaba B, Cooper D, MacGregor A, Scott DL. Development and validation of a patient-based disease activity score in rheumatoid arthritis that can be used in clinical trials and routine practice. Arthritis Rheum 2008;59:192–9. [PubMed: 18240256]

11. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24. [PubMed: 3358796]

12. Egger, M. Systematic reviews in health care: meta-analysis in context. London: BMJ Publishing Group; 2001.

13. Alarcon GS, Tilley BC, Li S, Fowler SE, Pillemer SR. Self-administered joint counts and standard joint counts in the assessment of rheumatoid arthritis. J Rheumatol 1999;26:1065–7. [PubMed: 10332969]

14. Escalante A. What do self-administered joint counts tell us about patients with rheumatoid arthritis? Arthritis Care Res 1998;11:280–90. [PubMed: 9791327]

15. 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. [PubMed: 1734905]

16. Hunter, J.; Schmidt, FL. Methods of meta-analysis. Newbury Park, CA: Sage; 1990.

17. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2007.

18. Franzblau, AN. A primer of statistics for non-statisticians. New York: Harcourt, Brace & World; 1958.

19. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994;50:1088–101. [PubMed: 7786990]

20. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol 2005;58:882– 93. [PubMed: 16085191]

21. 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. [PubMed: 7633798]

22. 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. [PubMed: 2213775]

23. Taal E, Abdel-Nasser AM, Rasker JJ, Wiegman O. A self-report Thompson articular index: What does it measure? Clin Rheumatol 1998;17:125–9. [PubMed: 9641509]

24. Calvo FA, Calvo A, Berrocal A, Pevez C, Romero F, Vega E, et al. Self-administered joint counts in rheumatoid arthritis: Comparison with standard joint counts. J Rheumatol 1999;26:536–9. [PubMed: 10090158]

25. Wong AL, Wong WK, Harker J, Sterz M, Bulpitt K, Park G, et al. Patient self-report tender and swollen joint counts in early rheumatoid arthritis. J Rheumatol 1999;26:2551–61. [PubMed: 10606362]

26. 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. [PubMed: 17216678]

27. 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. [PubMed: 8724295]

28. Stucki G, Stucki S, Bruhlmann P, Maus S, Michel BA. Comparison of the validity and reliability of self-reported articular indices. Br J Rheumatol 1995;34:760–6. [PubMed: 7551662]

29. 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. [PubMed: 8877920]

30. Fransen J, Langenegger T, Michel BA, Stucki G. Feasibility and validity of the RADAI, a self-administered rheumatoid arthritis disease activity index. Rheumatology 2000;39:321–7. [PubMed: 10788543]

31. 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. [PubMed: 10334682]

32. Greenwood MC, Hakim AJ, Carson E, Doyle DV. Touch-screen computer systems in the rheumatology clinic offer a reliable and user-friendly means of collecting quality-of-life and outcome data from patients with rheumatoid arthritis. Rheumatology 2006;45:66–71. [PubMed: 16263782]

33. 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. [PubMed: 8399430]

34. MacLean CH, Saag KG, Solomon DH, Morton SC, Sampsel S, Klippel JH. Measuring quality in arthritis care: methods for developing the Arthritis Foundation's quality indicator set. Arthritis Rheum 2004;51:193–202. [PubMed: 15077259]

35. Ward MM. Clinical measures in rheumatoid arthritis: which are most useful in assessing patients? J Rheumatol 1994;21:17–27. [PubMed: 7864914]

36. 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. [PubMed: 17444590]

37. Schillinger D, Machtinger EL, Wang F, Palacios J, Rodriguez M, Bindman A. Language, literacy, and communication regarding medication in an anticoagulation clinic: a comparison of verbal vs. visual assessment. J Health Commun 2006;11:651–64. [PubMed: 17074733]

38. Athale N, Sturley A, Skoczen S, Kavanaugh A, Lenert L. A web-compatible instrument for measuring self-reported disease activity in arthritis. J Rheumatol 2004;31:223–8. [PubMed: 14760788]

39. Dwyer KA, Coty MB, Smith CA, Dulemba S, Wallston KA. A comparison of two methods of assessing disease activity in the joints. Nursing Res 2001;50:214–21.

1) Please indicate with a mark, on the picture below all joints which are painful at present

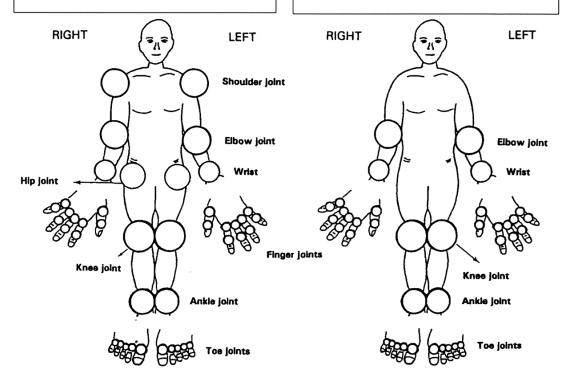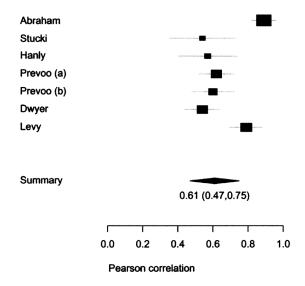2) Please indicate with a mark, on the picture below all joints which are swollen or tender at present

RIGHT   LEFT

Shoulder joint
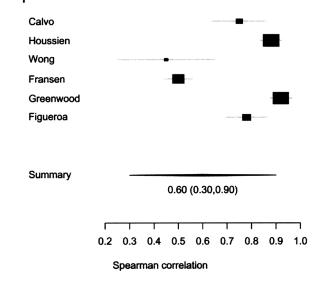
Elbow joint

Wrist

Hip joint

Finger joints

Knee joint

Ankle joint

Toe joints

RIGHT   LEFT

Elbow joint

Wrist

Knee joint

Ankle joint

Toe joints

Figure 1.
Mannequin figure. Reproduced from Stucki, *et al*[28] with permission.
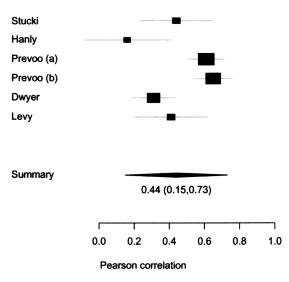
a. Pearson Tender Joint Count

b. Pearson Swollen Joint Count

c. Spearman Tender Joint Count

d. Spearman Swollen Joint Count
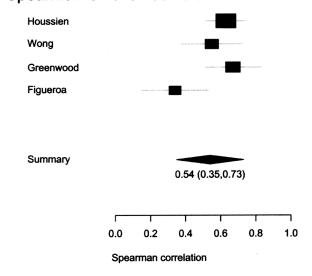
Figure 2.
Forrest plots of summary correlation estimates.

Table 1

Characteristics of 18 studies comparing patient with trained assessor joint counts in rheumatoid arthritis.

Source	No. of Joints	No. of Patients	No. of Assessors	Same Joint Count?	Blinding?
Text format					
Mason[15]†	20*	45	3	Y	Y
Abraham[33]†	20*	32	1	N	Y
Hewlett[2]1F	64F	50F	1	Y	NA
Hanly[29]	20*	61	4	Y	Y
Fransen[30]†	16	584	NA	N	NA
Dwyer[39]†	20*	185	NA	N	NA
Greenwood[32]††	28	45	1	Y	Y
Figueroa[26]	42	79	3	N	NA
Mannequin format					
Stewart[22]	38	30	NA	Y	Y
Prevoo[27]	28	141; 101	NA	Y	?
Escalante[14]	40	103	1	Y	?
Taal[23]	38	43	1	Y	Y
Houssien[31]	28	100	1	Y	Y
Levy[36]††	28	60	1	Y	Y
Both					
Stucki[28]	16 text; 38 mannequin	55	NA	N	Y
Wong[25]	Mannequin: 50 tender joints, 48 swollen joints; Text: 20 tender joints, 18 swollen joints	60	NA	Y	Y
Alarcon[13]†	36	67	1	Y	Y
Calvo[24]	36	60	1	Y	?

† Tender joint count only.

†† Computer report.

* Rapid Assessment of Disease Activity in Rheumatology (RADAR) groups hand knuckles, finger knuckles, toe knuckles.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NA: not available.

Table 2

Tender joint count correlation coefficients.

Source	Correlation	Correlation Type
Text		
Mason[15]	0.81	ICC
Alarcon[13]	0.55	ICC
Abraham[33]	0.89	Pearson
Hanly[29]	0.57	Pearson
Dwyer[39]	0.54	Pearson
Calvo[24]	0.75	Spearman
Houssien[31]	0.88	Spearman
Wong[25]	0.45	Spearman
Fransen[30]	0.50	Spearman
Figueroa[26]	0.78	Spearman
Mannequin		
Escalante[14]	0.78	ICC
Alarcon[13]	0.64	ICC
Stucki[28]	0.54	Pearson
Prevoo[27]	0.62	Pearson
Prevoo[27]	0.60	Pearson
Levy[36]*	0.79	Pearson
Calvo[24]	0.77	Spearman
Wong[25]	0.69	Spearman
Greenwood[32]*	0.92	Spearman

*
Computer touch-screen mannequin.

ICC: intraclass correlation coefficient.

Table 3

Swollen joint count correlation coefficients.

Source	Correlation	Correlation Type
Text		
Hanly[29]	0.16	Pearson
Dwyer[39]	0.31	Pearson
Houssien[31]	0.63	Spearman
Wong[25]	0.55	Spearman
Figueroa[26]	0.34	Spearman
Mannequin		
Escalante[14]	0.31	ICC
Stucki[28]	0.44	Pearson
Prevoo[27]	0.61	Pearson
Prevoo[27]	0.65	Pearson
Levy[36]*	0.41	Pearson
Wong[25]	0.61	Spearman
Greenwood[32]*	0.67	Spearman

*
Computer touch-screen mannequin.

ICC: intraclass correlation coefficient.