# Clinical Composite Measures of Disease Activity for Diagnosis and Followup of Undifferentiated Peripheral Inflammatory Arthritis: A Systematic Review

ISABEL CASTREJÓN, LUCÍA SILVA-FERNÁNDEZ, CLAIRE BOMBARDIER, and LORETO CARMONA

***ABSTRACT.*** ***Objective.*** To critically appraise the validity of activity indices used in the followup of patients with undifferentiated peripheral inflammatory arthritis (UPIA).

***Methods.*** A systematic review was performed in Medline, Embase, the Cochrane Library, and abstracts presented at the 2007 and 2008 meetings of the American College of Rheumatology and European League Against Rheumatism. Selection criteria were: patients with UPIA, the assessment of instruments to evaluate disease activity, and assessment of validity of the instruments. Two reviewers screened titles and abstracts independently and collected data using ad hoc standard forms.

***Results.*** The search yielded 179 articles and 834 abstracts, of which 4 articles and 1 abstract were included. We found no study that validated Disease Activity Score (DAS), Clinical Disease Activity Index (CDAI), or Simplified Disease Activity Index (SDAI). Included studies addressed validation of 4 questionnaires: WHO Disability Assessment Schedule (WHODAS), London Handicap Scale (LHS), Disease Repercussion Profile (DRP), and the Health Assessment Questionnaire (HAQ); and 3 indexes: RA Disease Activity Index (RADAI), McGill Range of Motion Index (McROMI), and NOAR Damaged Joint Count (NOAR-DJC). Questionnaires were self-administered and feasible; RADAI was the most feasible index. Internal consistency was studied in the questionnaires (Cronbach's $\alpha > 0.83$). Responsiveness was tested in the DRP, LHS, and HAQ, but the approach to study sensitivity to change was poorly explained, with no clear intervention. Construct validity, examined by means of convergence with other instruments, was generally moderate, and slightly higher for the RADAI.

***Conclusion.*** No instrument of disease activity has been fully validated for use in UPIA. We found no direct evidence of what is the most useful index to follow up patients with UPIA. (J Rheumatol 2011;38 Suppl 87:48–53; doi:10.3899/jrheum.101075)

*Key Indexing Terms:*
UNDIFFERENTIATED ARTHRITIS ASSESSMENT TOOLS DISEASE ACTIVITY
OUTCOME MEASURES SYSTEMATIC REVIEW

*From the Rheumatology Department, Hospital Universitario la Princesa; Rheumatology Department, Hospital Universitario Puerta de Hierro Majadahonda, Madrid, Spain; Division of Rheumatology and Department of Health Policy, Management, and Evaluation, University of Toronto; Division of Clinical Decision Making and Health Care, Toronto General Research Institute, University Health Network, Mount Sinai Hospital, Toronto, Canada; and Research Unit, Fundación Española de Reumatología, Madrid, Spain.*

*Supported by an unrestricted educational grant from Abbott. Abbott had no role in the design, literature search, data collection, data analysis, data interpretation, or writing of this report. Dr. Bombardier holds a Pfizer Chair and Canada Research Chair in Knowledge Transfer for Musculoskeletal Care.*

*I. Castrejón, MD, Rheumatology Department, Hospital Universitario la Princesa; L. Silva-Fernández, MD, Rheumatology Department, Hospital Universitario Puerta de Hierro Majadahonda; C. Bombardier, MD, FRCP, Division of Rheumatology and Department of Health Policy, Management, and Evaluation, University of Toronto; Division of Clinical Decision Making and Health Care, Toronto General Research Institute, University Health Network, Mount Sinai Hospital; L. Carmona, MD, PhD, Research Unit, Fundación Española de Reumatología.*

*Address correspondence to Dr. I. Castrejón, S. Reumatología, Hospital La Princesa, C/ Diego de León 62, 28006 Madrid, Spain. E-mail: isabelcastrejon@ser.es*

Many instruments for disease activity assessment have been developed in recent years. These indices are frequently used in clinical trials as well as in daily practice as they are useful to evaluate response to treatment or to make a decision to start or change treatment. Use of such indices has become an important aspect of the care for patients with rheumatoid arthritis (RA)[1,2]. However, we are unaware if they are equally useful for patients with undifferentiated peripheral inflammatory arthritis (UPIA).

UPIA is a form of arthritis that does not fulfill classification criteria for a more definitive diagnosis. Patients with UPIA are hard to follow in clinical practice, as they comprise a very heterogeneous group, sharing characteristics of different diagnoses. Due to the lack of a more precise clinical picture and outcome, it is important to have comprehensive tools that help the clinician anticipate outcomes, including more precise diagnosis, and thus make therapeutic decisions. Studies focused on UPIA have used many different indices to evaluate outcome. However, the sole fact of using

an index in a study does not confer validity for evaluating outcome in that particular population, in this case, disease activity in UPIA. An instrument should demonstrate that it measures what is intended, discriminates between different disease states, and shows change in the numerical result when the patient improves.

Our objective was to analyze the validity of any available activity index, instrument, or scale, used to evaluate disease activity of patients with UPIA. The clinical correlate to our objective was to answer the question, "Which clinical assessments of disease activity [e.g., Disease Activity Score (DAS), Simplified Disease Activity Index (SDAI), and Clinical Disease Activity Index (CDAI)] should be done (at baseline and repeat at what interval) in patients with undifferentiated arthritis?"

## MATERIALS AND METHODS

This systematic review is part of the 3e (evidence, expertise, exchange) Initiative in Rheumatology. The 3e Initiative is a multinational effort, aimed at promoting evidence-based medicine, by formulating detailed recommendations addressing clinical problems[3,4]. In contrast to guidelines developed by a limited panel of experts, the 3e Initiative involves a broad international panel of rheumatologists. Further, the initiative promotes epidemiology, by teaching and conducting systematic literature research following a strict methodology[5]. The objective of the 3e Initiative of 2008-2009 was to develop practical recommendations for the investigation and followup of UPIA, by integrating systematically generated evidence and expert opinion of a broad panel of international rheumatologists.

*Rephrasing the research question*. The clinical question as formulated by experts from 18 countries was translated into an epidemiological research question according to the PICO (Patient, Intervention, Comparator, and Outcome) approach[6]. Definitions: Patients were defined as adults with UPIA; Intervention as any index, instrument, or scale used to evaluate disease activity; Comparators were the above indices compared to themselves or to another index; and Outcome was any aspect of validity: construct validity, feasibility, reliability, or responsiveness.

The final search question was rephrased as, "What are the most suitable clinical measures to evaluate the diagnosis of UPIA? and, What are the most useful indexes to evaluate the followup?". From a clinical point of view, we wanted to know whether any index could help differentiate patients with UPIA from others that would develop a specific diagnosis.

*Scenarios*. Whether a disease activity measure is useful is a difficult question to address. We identified at least 3 possible approaches to the answer, each determining a different search strategy. The first approach would be to retrieve all studies that included common indices of disease activity (DAS, SDAI, CDAI, etc.) in which at least one group had UPIA. Next, we would evaluate whether these indices were discriminating groups of high and low activity in the UPIA arm. The second approach would be to search for all UPIA studies, and check which indices were used. This approach would tell us only if the indices discriminate between patient states. These 2 approaches would require in-depth knowledge of the studies through contact with the authors, as the information on discrimination of variables is not usually available in studies that do not specifically address validity.

The third and most objective approach, which was selected by the team, involved searching for studies in which the population studied was UPIA, in which indices were used, and that specifically measured any aspect of validity.

*Systematic literature search*. We performed a systematic literature search for articles published between 1950 and January 2009 in Medline, Embase, and the Cochrane Library, using a comprehensive search strategy (see online appendix, available from: www.3eupia.com). Abstracts presented at the 2007 and 2008 meetings of the European League Against Rheumatism (EULAR) and American College of Rheumatology (ACR) were also searched. Two reviewers (LS and IC) screened the titles and abstracts for selection criteria independently, using a third reviewer for consensus, and collected the data using ad hoc standard forms. Review articles were also retrieved for identifying additional references by hand search.

*Selection of articles*. Relevant articles were selected in a systematic procedure. First, titles and abstracts of all identified references were screened, excluding articles that clearly did not address the topic of interest. Second, selected articles were reviewed in the full report, applying the following inclusion criteria: validation studies, UPIA patients, adults (> 18 years), and disease activity measures. Articles that did not fulfill all the inclusion criteria were excluded (Table 1). We did not include other standard clinical monitoring measures such as pain, global assessment, joint count, erythrocyte sedimentation rate, etc., because they were being investigated in other reviews in the 3e Initiative.

*Data extraction and analysis*. Publication details, patient characteristics, instruments to measure disease activity, and aspects of the assessment of validity were extracted using standard forms. To evaluate validity of each instrument, feasibility, reliability, responsiveness, and construct validity were analyzed. To measure feasibility, we deduced from study information the time required to complete an instrument and its ease of use. We created an ad hoc measure that went from 0 (unfeasible) to 3 (completely feasible).

Reliability embraces the concept that repeat administration of a measurement tool in stable subjects will yield the same result, thus measuring an instrument's stability. Reliability also includes an instrument's internal consistency or "good construction," as expressed by the Cronbach's alpha statistic (< 0.70 = individual items provide an inadequate contribution to the overall scale; > 0.90 suggests redundancy). The stability of the instrument should be tested twice: by the same operator at different times (intraobserver test-retest) and by a different operator (interobserver reliability) at the same time. Intra- and interobserver reliability are measured either with the kappa statistic or with the intraclass correlation coefficient.

Responsiveness, also called sensitivity to change, is defined as "the ability of an instrument to accurately detect change when it has occurred"[7]. It measures whether the instrument detects that the patient has improved or worsened. Responsiveness implies that an intervention with an effect of known direction is given to the studied patients. It is quantified by the effect size (ES) or the standardized response mean (SRM). In accord with the literature, ES were considered as follows: ~0.2 = small, ~0.5 = moderate, and > 0.8 = large.

Construct validity is a measure of how close to what the instrument says it measures it really measures. One way of conceptualizing construct validity is to consider it as testing hypotheses of what a valid instrument would and would not correlate with; thus the instrument is compared to other instruments measuring the same construct (high correlation) and different construct (low correlation). Establishing the validity of an instrument to measure disease activity is difficult, as no established "gold standard" is available. In most of the studies retrieved the construct validity was examined in terms of convergence with variables that should have a converging relationship (correlation > 0.60 is considered a good correlation).

## RESULTS

A total of 179 references and 834 meeting abstracts were identified. After title/abstract screening, 19 articles were retrieved for full article review, of which 4 fulfilled inclusion criteria. One meeting abstract was also included. Thus, 5 records were included addressing some aspects of the validation of 4 questionnaires and 3 physical measures. We found no study on the validation of the most common activity measures such as DAS[8] or SDAI[9] in patients with UPIA.

*Table 1.* Excluded studies and reason for exclusion.

| Study | Reason for Exclusion |
|---|---|
| Tully[21] | Arthritis and osteoarthritis population, age > 60 yrs |
| Cohen[22] | RA > 1 yr duration and prognostic factors of quality of life after 5 yrs of followup |
| El Miedany[23], Kievit[24], Suurmeijer[25], Salaffi[26] | Population was RA, not UPIA |
| Lerner[27] | RA patients as controls to validate disability in other diseases |
| Smolen[28] | RA population. Review |
| Hamilton[29] | RA population (per ACR criteria) > 1 yr disease duration. Not an index: Gait analysis on contact-sensitive walk mat system |
| Cole[30] | Scleroderma and RA population. Study goal was to examine structural validity of HAQ in patients with SSc |
| Saraux[31], van der Helm[32, 33], El Miedany[34] | Prediction rules (multivariate) for RA |

A summary of the results of the validity of the different instruments can be found in Table 2.

*Description and feasibility of the questionnaires*. The 4 questionnaires for which a validation in a UPIA population was published were the World Health Organization Disability Assessment Schedule (WHODAS), London Handicap Scale (LHS), Disease Repercussion Profile (DRP), and the Health Assessment Questionnaire (HAQ), all self-administered.

The WHODAS is a short-form questionnaire comprising 36 Likert-formatted questions divided into 6 domains (understanding/communicating, getting around, self-care, getting along with people, life activities, and participation in society)[10]; final score ranges from 0 (best) to 100 (worst). The LHS has 6 domains covering handicap dimensions: mobility, physical independence, occupation, social interaction, orientation, and economic self-sufficiency; score ranges from 100 (no disadvantage) to 0 (extreme disadvantage)[11]. The DRP consists of 6 visual analog scales on the importance to the patients of 6 domains: functional and social activity, employment/money, relationships, emotions, and body image; score ranges from 0 (none) to 10 (extremely important)[12]. The HAQ includes 20 items on ability for daily activities: dressing and grooming, rising, eating, walking, hygiene, functional reach and grip and activities; score ranges from 0 (none) to 3 (complete disability)[13].

Since all questionnaires are self-administered, they are feasible, although the WHODAS and the LHS seem to take longer.

*Description and feasibility of the indexes*. We identified 3 indices that had been validated in UPIA, or at least some aspect of the index had been validated: the Rheumatoid Arthritis Disease Activity Index (RADAI), McGill Range of Motion Index (McROMI), and NOAR Damaged Joint Count (NOAR-DJC).

The RADAI is a self-administered questionnaire that yields an index of activity. It comprises 5 individual items that have a high association with clinically assessed joint synovitis and acute-phase response, providing a global score[14]. A validity study in UPIA was available only in abstract form[15].

The McROMI index is based on a visual estimate of range of motion (ROM) of 19 movements in 9 joint areas (neck, shoulder, elbow, wrist, forearm, hand, hip, knee, and ankle), all bilaterally, except for the neck. The authors propose that a limited ROM from inflammation and pain may occur early in the disease process. To obtain a score, each movement is graded from 0 to 3, 3 being the most abnormal,

*Table 2.* Summary of results of validity of instruments retrieved in the search strategy. Feasibility (0 = not feasible to 3 = completely feasible) is based on the time it takes, and how easy it is to answer the questions.

| | Feasibility | Internal Consistency | Test-Retest | Responsiveness | Construct Validity |
|---|---|---|---|---|---|
| WHODAS | 2 | +++ | +++ | | ++ |
| DRP | 3 | ++ | | – | ++ |
| LHS | 2 | ++ | | – | ++ |
| HAQ | 3 | ++ | | – | |
| McROMI | 1 | | | | ++ |
| NOAR-DJC | 1 | | ++ | | ++ |
| RADAI | 3 | | | | +++ |

WHODAS: WHO Disability Assessment Schedule; DRP: Disease Repercussion Profile; LHS: London Handicap Scale, HAQ: Health Assessment Questionnaire; McROMI: McGill Range of Motion Index; NOAR-DJC: NOAR Damaged Joint Count, RADAI: Rheumatoid Arthritis Disease Activity Index.

and the maximum score 111. The McROMI requires assistance to complete, and the movements and scores assigned to the different degrees of mobility are not easy to remember[16].

Last, the NOAR-DJC index assesses the presence or absence of deformity in 51 joints. Distal interphalangeal joints are included because the NOAR-DJC was intended for use in patients with early inflammatory polyarthritis. Deformity is defined as the inability to adopt an anatomical position and a reduction in range of movement and/or surgical alteration of the joint. The authors consider deformity to be a reversible feature of early arthritis. NOAR-DJC requires a guideline and a manikin to perform[17].

Concerning feasibility of the above indexes, the RADAI seems to be the most feasible because the others are time-consuming and difficult to perform.

*Reliability and responsiveness*. The internal consistency of the 4 questionnaires was very good, all showing Cronbach's $\alpha > 0.8$. Regarding the indices, internal consistency is not absolutely necessary because indices are composed of very similar items. The RADAI is a mixture of a questionnaire and an index, as the way questions are asked (internal consistency) is not all that relevant. That is why we cannot say it is totally unnecessary.

On the other hand, test-retest reliability is very important for measures that imply an operator, such as mobility and deformity indices. In testing the NOAR-DJC, the interclass correlation coefficient (ICC) was slightly higher for the intraobserver study (0.88, 95% CI 0.79–0.94, $p < 0.001$) than for the interobserver study (0.74, 95% CI 0.53–0.86, $p < 0.001$). Both results showed fairly good reliability. Surprisingly, this was not tested in the McROMI as might have been expected in this type of index. In questionnaires, test-retest reliability was evaluated only in the WHODAS, where intraobserver reliability was assessed in 20 subjects of the population: the ICC between time 1 and time 2 was 0.94 (95% CI 0.86–0.98).

Regarding responsiveness, it was tested only in the DRP, LHS, and HAQ, all in the same study[18]. The approach to measure responsiveness in this study was rather poor, with no clear intervention or anticipated change.

*Construct validity*. To evaluate construct validity of the questionnaires we had to consider that the authors wanted to measure disability. The WHODAS has a moderate correlation with measures of disability such as HAQ (Kendall's tau-b = 0.55, $p < 0.001$). Unexpectedly, the correlation with the Medical Outcome Study Short-Form Survey 36 was moderate (Kendall's tau-b = –0.43, $p = 0.001$, for the mental component and Kendall's tau-b = –0.51, $p = 0.001$, for the physical component). Evidently the correlation with measures of activity (DAS28, tender and swollen joint count, self-report of pain, and global assessments) was poor (0.17 to 0.41).

The DRP showed a moderate correlation with measures of disability such as the HAQ (Spearman's correlation = 0.59, $p < 0.001$) and measures of handicap such as the LHS (Spearman's correlation = –0.51, $p < 0.001$). The LHS showed good correlation with disability as measured by the HAQ (–0.71, $p < 0.001$; the correlation is negative because the questionnaires are ordered in different directions). The correlation with disease activity measures was low (–0.36 to 0.01).

The HAQ was tested only against the LHS (Spearman's correlation = 0.71, $p < 0.001$) and the DRP (Spearman's correlation = 0.59, $p < 0.001$) as a similar construct, showing moderate to good correlations. The HAQ in UPIA does not have a good correlation with disease activity measures (0.17–0.41), although they are larger than for the DRP or the LHS. In summary, construct validity of these questionnaires is not bad for what they intend to measure, but they clearly do not measure disease activity.

With regard to the McROMI and NOAR-DJC, they are presented as instruments that should measure a construct close to disease activity. The problem with UPIA is that the number of swollen joints varies much more than in a specific disease (oligo to polyarthritis versus polyarthritis in RA, for instance), and relying on the number of swollen joints for disease activity may not be completely adequate. When the McROMI was tested against measures of disease activity, the correlation was poor, being the higher one when it was compared with DAS28-C-reactive protein (tau-b = 0.42, $p < 0.001$). Correlation was not better when compared to measures of function, which best correlated with the HAQ (tau-b = 0.44, $p < 0.001$). The NOAR-DJC was tested against different activity measures after 1 and 5 years of followup. Results after 1 year of followup showed low correlations: tender joint count, r = 0.18 (95% CI 0.12, 0.24); swollen joint count, r = 0.21 (95% CI 0.16, 0.27); HAQ, r = 0.39 (95% CI 0.34, 0.44 ); and eroded joint count, r = 0.19 (95% CI 0.10, 0.27). After 5-year followup, correlations were slightly better: tender joint count, r = 0.28 (95% CI 0.20, 0.35); swollen joint count, r = 0.33 (95% CI 0.25, 0.39); HAQ, r = 0.45 (95% CI 0.40, 0.50); and eroded joint count, r = 0.42 (95% CI 0.35, 0.49). This article actually shows the correlation in a subgroup that developed RA after 1 year: they were the same or worse.

The RADAI was tested only in UPIA versus the DAS28 (Pearson's correlation 0.596, $p < 0.0001$). This was the best correlation with a disease activity measure that we found.

## DISCUSSION

Despite their well established use to evaluate disease activity in RA, it remains unclear whether composite indices may also be useful in patients with UPIA.

Our systematic review summarized available evidence from the literature on the most suitable clinical instruments to evaluate the diagnosis and followup of UPIA. Our results showed that no disease activity instrument has been fully

validated for use in UPIA; lack of validation is particulary apparent for the most commonly used indexes such as DAS28, SDAI, and CDAI, which should be included in a research agenda. We observed that since the questionnaires retrieved are designed mainly to measure disability, although they were in part validated in a UPIA population, they cannot be recommended to evaluate disease activity, not to mention diagnostic evolution and followup in these patients. The indices may be useful to evaluate disability in patients with UPIA and indirectly to evaluate disease progression. Physical disability is the most powerful determinant of all severe longterm outcomes in RA[19], and possibly in UPIA as well.

Concerning the indices retrieved, only the RADAI, a mixed questionnaire-index, seems to be useful. However, it was not completely validated. Construct validity was examined versus the DAS28 only, showing a good correlation. This finding suggests that the RADAI may be a valid and feasible instrument for the assessment of disease activity in patients with UPIA, although it would be necessary to establish whether it can detect clinically important changes.

The 2 other indices are clearly not suitable, as they are time-consuming and difficult to perform, and validation is clearly incomplete. We are unaware of the effect size, and the construct validity is not very promising.

Validation of an instrument is a continuing process, and testing validity is established not from a single approach but from a series of converging studies. Future validation of these indices in different populations is necessary, especially because validity of an instrument is population-specific[20].

There are important limitations to our analyses; the study of validity of any clinical measure in UPIA is very challenging, in particular since "disease activity" implies that a defined disease should be diagnosed, which is not the case in UPIA. Further, many of the included studies do not evaluate all aspects of the validation for an instrument.

In this systematic review we have not found direct evidence on what the most useful index is to follow up patients with UPIA, or at which intervals these should be repeated. The experts decided, in light of very limited evidence and mainly based on their experience, that disease activity should be monitored; however, no specific instrument can be recommended.

Future research is needed to evaluate the validity of the most common indices in populations with UPIA. Also, it would be important to adjust these indices to the different characteristics of these patients, possibly including larger joint counts or other extraarticular features.

## ACKNOWLEDGMENT

We thank all members of the 3e scientific committees, all participants of the national meetings, and acknowledge the support from Margaux Orange.

## REFERENCES

1. Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, van Zeben D, Kerstens PJ, Hazes JM, et al. Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): a randomized, controlled trial. Arthritis Rheum 2005;52:3381-90.
2. Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. Lancet 2004;364:263-9.
3. Machado P, Castrejon I, Katchamart W, Koevoets R, Kuriya B, Schoels M, et al. Multinational evidence-based recommendations on how to investigate and follow-up undifferentiated peripheral inflammatory arthritis: integrating sytematic literature research and expert opinion of a broad international panel of rheumatologists in the 3e Initiative. Ann Rheum Dis 2010 August 19. [Epub ahead of print]
4. Visser K, Katchamart W, Loza E, Martinez-Lopez JA, Salliot C, Trudeau J, et al. Multinational evidence-based recommendations for the use of methotrexate in rheumatic disorders with a focus on rheumatoid arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3e Initiative. Ann Rheum Dis 2009;68:1086-93.
5. van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the Cochrane collaboration back review group. Spine (Phila Pa 1976) 2003;28:1290-9.
6. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996;312:71-2.
7. De Bruin AF, Diederiks JP, de Witte LP, Stevens FC, Philipsen H. Assessing the responsiveness of a functional status measure: the Sickness Impact Profile versus the SIP68. J Clin Epidemiol 1997;50:529-40.
8. 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.
9. 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.
10. Baron M, Schieir O, Hudson M, Steele R, Kolahi S, Berkson L, et al. The clinimetric properties of the World Health Organization Disability Assessment Schedule II in early inflammatory arthritis. Arthritis Rheum 2008;59:382-90.
11. Harwood RH, Rogers A, Dickinson E, Ebrahim S. Measuring handicap: the London Handicap Scale, a new outcome measure for chronic disease. Qual Health Care 1994;3:11-6.
12. Carr AJ, Thompson PW. Towards a measure of patient-perceived handicap in rheumatoid arthritis. Br J Rheumatol 1994;33:378-82.
13. Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the Health Assessment Questionnaire, disability and pain scales. J Rheumatol 1982;9:789-93.
14. 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.
15. Bykerk VP, Mironyuk L, Chen H, Maguire L, Jamal S, Bombardier C. Validity of the RADAI in early rheumatoid arthritis [abstract]. Ann Rheum Dis 2007;66 Suppl 2:344.
16. Baron M, Steele R. Development of the McGill Range of Motion Index. Clin Orthop Relat Res 2007;456:42-50.
17. Bunn DK, Shepstone L, Galpin LM, Wiles NJ, Symmons DP. The NOAR Damaged Joint Count (NOAR-DJC): a clinical measure for assessing articular damage in patients with early inflammatory polyarthritis including rheumatoid arthritis. Rheumatology

2004;43:1519-25.
18. Harwood RH, Carr AJ, Thompson PW, Ebrahim S. Handicap in inflammatory arthritis. Br J Rheumatol 1996;35:891-7.
19. 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.
20. Pedhauzer E, Schmelkin L. Measurement, design and analysis: an integrated approach. Hillsdale, NJ: Lawrence Erlbaum Associates; 1991:37-9.
21. Tully M, Cantrill J. The test-retest reliability of the modified Patient Generated Index. J Health Serv Res Policy 2002;7:81-9.
22. Cohen JD, Dougados M, Goupille P, Cantagrel A, Meyer O, Sibilia J, et al. Health Assessment Questionnaire score is the best predictor of 5-year quality of life in early rheumatoid arthritis. J Rheumatol 2006;33:1936-41.
23. El Miedany Y, Palmer D. Can standard rheumatology clinical practice be patient-based? Br J Nurs 2008;17:673-5.
24. Kievit W, Welsing PM, Adang EM, Eijsbouts AM, Krabbe PF, van Riel PL. Comment on the use of self-reporting instruments to assess patients with rheumatoid arthritis: the longitudinal association between the DAS28 and the VAS general health. Arthritis Rheum 2006;55:745-50.
25. Suurmeijer TP, Doeglas DM, Moum T, Briancon S, Krol B, Sanderman R, et al. The Groningen Activity Restriction Scale for measuring disability: its utility in international comparisons. Am J Public Health 1994;84:1270-3.
26. Salaffi F, Stancati A, Neri R, Grassi W, Bombardieri S. Measuring functional disability in early rheumatoid arthritis: the validity, reliability and responsiveness of the Recent-Onset Arthritis Disability (ROAD) index. Clin Exp Rheumatol 2005;23:S31-42.
27. Lerner D, Amick BC 3rd, Rogers WH, Malspeis S, Bungay K, Cynn D. The Work Limitations Questionnaire. Med Care 2001;39:72-85.
28. Smolen JS, Aletaha D. Activity assessments in rheumatoid arthritis. Curr Opin Rheumatol 2008;20:306-13.
29. Hamilton J, Brydson G, Fraser S, Grant M. Walking ability as a measure of treatment effect in early rheumatoid arthritis. Clin Rehabil 2001;15:142-7.
30. Cole JC, Khanna D, Clements PJ, Seibold JR, Tashkin DP, Paulus HE, et al. Single-factor scoring validation for the Health Assessment Questionnaire-Disability Index (HAQ-DI) in patients with systemic sclerosis and comparison with early rheumatoid arthritis patients. Qual Life Res 2006;15:1383-94.
31. Saraux A, Berthelot JM, Chales G, Le Henaff C, Thorel JB, Hoang S, et al. Ability of the American College of Rheumatology 1987 criteria to predict rheumatoid arthritis in patients with early arthritis and classification of these patients two years later. Arthritis Rheum 2001;44:2485-91.
32. Van der Helm-van Mil AH, Detert J, le Cessie S, Filer A, Bastian H, Burmester GR, et al. Validation of a prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: moving toward individualized treatment decision-making. Arthritis Rheum 2008;58:2241-7.
33. Van der Helm-van Mil AH, le Cessie S, van Dongen H, Breedveld FC, Toes RE, Huizinga TW. A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: how to guide individual treatment decisions. Arthritis Rheum 2007;56:433-40.
34. El Miedany Y, Youssef S, Mehanna AN, El Gaafary M. Development of a scoring system for assessment of outcome of early undifferentiated inflammatory synovitis. Joint Bone Spine 2008;75:155-62.