

# Rheumatoid arthritis: assessing disease activity and outcome

Deborah PM Symmons

**ABSTRACT** – The range of treatments for rheumatoid arthritis (RA) has increased significantly in recent years, with a parallel improvement in patient outcome. The development and assessment of new therapies and therapeutic strategies relies on the availability of valid and reliable outcome measures to assess the diverse impact of RA on the patient's life. This paper reviews the outcome measures in current use which assess disease activity, joint damage, physical function and health-related quality of life. Some measures have been combined into composite indices which are useful for summarising the patient's current condition and as primary outcome measures for clinical trials. There is still a need for better and more relevant tools especially for imaging multiple joints and for assessing fatigue.

**KEY WORDS:** outcome measures, rheumatoid arthritis

## Introduction

One hundred years ago aspirin was the mainstay of treatment for rheumatoid arthritis (RA). Today a range of non-steroidal anti-inflammatory drugs, nonbiologic disease-modifying antirheumatic drugs (DMARDs) and corticosteroids exists, as well as a rapidly increasing number of biologic agents. Patients rarely develop the severe disability and deformity so common in the past. The National Institute for Health and Clinical Excellence (NICE) *National clinical guideline for the management and treatment of adults with rheumatoid arthritis*, launched in 2009, emphasises:

- the need for early introduction of DMARDs
- the centrality of methotrexate in the management of early disease
- the value of combination therapy
- the early introduction of biologic agents
- the multidisciplinary approach
- the treatment of disease to a target of remission
- the importance of adopting a patient-centred approach.<sup>1</sup>

All these changes reflect a revolution in the treatment of RA and indeed in the outcome of the condition. However, none of this would have been possible if there had not also been a revolution

in the accurate assessment of disease activity and outcome in RA. As H James Harrington, the former principal of Ernst and Young, said 'If you can't measure something, you can't understand it. If you can't understand it, you can't control it. If you can't control it, you can't improve it'. If adequate tools to measure the efficacy of new treatments did not exist, it would not have been possible to design or publish the results of trials nor to persuade the NHS to devote the large sums of money now needed to pay for biologic agents.<sup>2</sup>

## Assessment of disease activity

Outcome in RA is multidimensional.<sup>3</sup> The key components are current disease activity and cumulative disease damage. Instruments used to assess disease activity in RA have become progressively less complex, less doctor centred and more patient centred over recent decades. Obviously, the inside of every joint cannot be inspected and so surrogate external signs and biomarkers are used to quantify the degree of synovial joint inflammation. The classic signs of inflammation are heat, redness, swelling, and pain. Heat and redness have not proved easy to assess or quantify. Assessment of joints, therefore, hinges predominantly on assessing pain and swelling by manual examination. A number of joint counts have been developed which differ according to whether they assess tenderness or swelling or both; the number of joints or joint areas that are assessed; whether or not the degree of tenderness or swelling is measured on a dichotomous (yes–no) or graded scale; and whether the joints are weighted according to their size (Table 1) (although weighting for joint size improves the correlation of a joint count with acute phase reactants, weighting also introduces interobserver error).<sup>4</sup> Some systems, such as the Lansbury Index and the Ritchie Index, only assess joints for tenderness. Joint counts which include a grading for the degree of tenderness also have higher interobserver error than those which do not.

The 28-joint count is the most widely used in Europe. It includes the shoulders, elbows, wrists, metacarpophalangeal (MCP) joints, proximal interphalangeal (PIP) joints and knees.<sup>5</sup> The most frequent criticism is that it does not include the ankle or toe joints. In America, the 68-joint count has been favoured. This is, of course, more time consuming to conduct. In an analysis of 735 patients included in RA clinical trials, the 28-joint and 68-joint counts were found closely correlated.<sup>5</sup> The most frequently involved joints were those that are included in the 28-joint count. Although the metatarsophalangeal (MTP) joints are often involved clinically, their exclusion in the context

This article is based on the Oliver Sharpey lecture given on 18 June 2009 by **Deborah PM Symmons**, professor of rheumatology and musculoskeletal epidemiology, Arthritis Research UK Epidemiology Unit, Manchester Academic Health Sciences Centre

**Table 1. Characteristics of joint counts.**

	Year	Number of joints	Tenderness alone	Swelling and/or tenderness	Weighting by joint area	Grades
Lansbury	1958	84	+	–	+	4
Cooperating clinics (ARA)	1965	66/68	–	+	–	–
Ritchie index	1968	26 joint areas	+	–	–	4
Thompson	1987	38	–	+	+	–
Fuchs	1989	28	–	+	–	–

\*Later modified to count tenderness and swelling separately. ARA = American Rheumatism Association.

of a clinical trial did not influence the conclusions. It was, however, noted that the ankle is involved more often than the shoulder and that it is easier to assess swelling in the ankle than the shoulder. Adapting the 28-joint count to substitute ankle for shoulder assessment may be justified.

The acute phase response (erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP)) is a laboratory measure of disease activity. However, this may be influenced by disease processes other than RA. Serum levels of cytokines such as tumour necrosis factor  $\alpha$  and interleukin 6 are likely to be more specific but measurement of these has not yet found its way into routine clinical practice.

Swollen joint counts and the acute phase response provide objective evidence of synovial inflammation. Tender joint counts rely on patient reporting and are really an assessment of pain. From the patient's perspective, the pain, swelling and stiffness from the joints constitute the most prominent symptoms of RA. Measurement of pain is very simplistic when the many physiological and cultural processes that feed into the perception of pain are considered. However, it seems that the simple horizontal or vertical visual analogue scale captures the essence of pain in RA and has good measurement properties. Patients also report that fatigue and poor quality sleep are very troublesome features of active RA. An instrument that measures fatigue in a reproducible and meaningful way should be incorporated into RA clinical trials.

### Core set of disease activity measures

Up until the early 1990s, it was difficult to compare the results of clinical randomised controlled trials in RA because each had different outcome measures. A series of meetings during the 1990s led to agreement on a core set of disease activity measures which is recommended for use in all RA clinical trials – regardless of the agent or type of intervention being assessed. Different initiatives were established in Europe, North America and worldwide but the final core sets of disease activity measures overlap except that the European (EULAR) set includes the 28-joint count, the North American (ACR) set includes the 68-joint count, and the worldwide (OMERACT) set does not specify the number of joints to be included.<sup>6–8</sup>

### Composite indices – strengths and weaknesses

There is something inherently appealing in using a composite measure which combines the various aspects of disease activity. As composite measures have a lower standard deviation than individual measures, clinical trials with a composite index as the primary endpoint can have smaller sample sizes than those with an individual outcome measure. The weakness of the composite measure is that the single numerical value may not always be easy to interpret or to compute. With time, composite indices of RA disease activity have tended to become less complex and to incorporate only those individual measures which have been shown to have good measurement properties (Table 2). The Disease Activity Score (DAS-28) is the most widely used index of disease activity.<sup>9</sup> The overall status in RA (OSRA) is a composite measure which summarises a patient's current clinical status in terms of:

- their age, sex and disease duration
- disease activity on a 10-point scale
- disease damage on a 10-point scale
- current treatment.<sup>10</sup>

It is quick to administer (requiring no blood tests) and can be carried out within the context of a routine clinical assessment.

The DAS was developed in Nijmegen in the early 1990s.<sup>11</sup> In total, 113 patients with RA were studied over a three-year period. High disease activity was reflected in the decision to start DMARD therapy and low disease activity in the decision either to reduce or leave DMARD therapy unchanged. Information on a wide range of clinical and laboratory parameters was collected and the clinical judgement of six rheumatologists was used to put patients into the high or low disease category state. Discriminate analysis and multiple regression was used to develop formulae that would predict accurately which state the patient should be in. The original DAS used the Ritchie articular index and a 44 swollen joint count. A simplified version was produced in 1995 which has four components: the number of swollen joints (out of 28), the number of tender joints (out of 28), either the ESR or CRP, and a patient global health (GH) assessment.<sup>9</sup> The DAS-28 ranges from around 1.6 to 9.1. Values below 3.2 are designated low disease activity and values below 2.6 are classified as remission.<sup>12</sup> Values above 5.1 are classified as

Table 2. Composite indices of disease activity.

	Steinbrocker <i>et al</i> <sup>16</sup>	Lansbury <sup>17</sup>	Mallya and Mace <sup>18</sup>	DAS/DAS-28 <sup>11,12</sup>
Number of swollen joints	+			+
Number of tender joints	+		+	+
Pain on motion	+	+		
Morning stiffness		+	+	
ESR/CRP	ESR	ESR	ESR	Either
Haemoglobin	+	+	+	
Weight	+			
Pain	+	+	+	
Patient global health				+
Muscle weakness		+		
Function	+			
Fatigue		+		

CRP = C-reactive protein; ESR = erythrocyte sedimentation rate.

high disease activity. This value was chosen because it corresponds with the usual entry criteria (in terms of number of tender and swollen joints and ESR) for most RA clinical trials. The EULAR response criteria are also based on the DAS-28<sup>13</sup> and are expressed in terms of both the change in disease activity ( $>1.2$  (ie two standard deviations) being viewed as a good improvement) and the achieved disease activity ( $-\leq 3.2$  being viewed as the desirable goal).

The ACR20 is used in the USA. In the context of clinical trials, patients in the placebo arm often improve by up to 20% in the seven items of the core set of outcome measures. Responses greater than 20% are, therefore, taken to indicate response attributable to an active DMARD. A patient is said to have achieved an ACR20 response if they have improved by at least 20% in five out of seven of the core disease activity measures which must include the number of swollen joints and the number of tender joints. ACR50 and ACR70 have equivalent definitions. The ACR response criteria assess only change and not absolute achieved disease activity.

## Assessing joint damage

Joint damage is seen, on examination, as joint deformity or loss in the range of movement. A number of clinical measures of joint damage have been developed over the years. None of these has found its way into routine clinical practice nor into clinical trials. Joint damage may also be quantified radiologically. Formal assessments are based predominantly on assessing erosions and joint space narrowing. Generally speaking, just the hands and feet are assessed. A number of studies have shown that radiographic changes in the hands and feet are representative of what is going on in the larger joints in the great majority of patients. Two systems are in common use. The first is the global grading system developed by Larsen and colleagues in 1977 and now modified.<sup>14</sup> The system in current use only includes the wrists, MCP, PIP and MTP joints. Each of these

joints is compared with a series of standard reference films and graded from zero to five. A score of  $\geq 2$  in any individual joint is erosive. The wrist is scored as a single joint and multiplied by five. Most recent trials have used the van der Heijde modification of the Sharp score method.<sup>15</sup> In this, erosions and joint space narrowing are scored separately at up to six sites per hand and six sites per foot. The total score ranges from 0 to 448. It is very time consuming but is sensitive to change in the context of clinical trials. More recently, both magnetic resonance imaging (MRI) and ultrasound have been used to image joints in RA. At the present time, it seems unlikely that MRI will find its way into routine clinical use. Although it can pick up very early and reversible changes in the bones and cartilage, cost and time constraints mean that only a small number of joints can be assessed in each individual patient. It has also been shown, in the context of clinical trials, that there is no increased sensitivity to change compared to conventional radiographs. However, this may change in the future. The use of ultrasound to detect synovitis in the joints has been a major development in the assessment of the rheumatoid patient. It is useful in the diagnosis of synovitis and may also be useful to establish remission. However, again, it is not practical to perform an ultrasound of every joint.

## Assessing physical function

The first attempt to assess physical disability was the Steinbrocker functional class, published in 1947, which categorised patients into one of four classes.<sup>16</sup> It is based purely on the physician's assessment. Although it is quick to apply and comprehensive, it is not sensitive to change. Even in long-term longitudinal studies, patients tend to remain in the same class for most of their illness. The Health Assessment Questionnaire (HAQ) was introduced in the 1980s and, unlike the radiological scoring systems, has undergone little subsequent modification apart from translation into other languages.<sup>2</sup> It is a self-completed questionnaire comprising 20 questions across eight

domains. The patient is asked to assess their difficulties in performing the task described across a four-point scale. The highest scoring question in each domain is used in the final score, which is averaged across the eight domains and ranges from zero (no disability) to three. The HAQ is probably the single most useful outcome measure in RA. It has immediate relevance to both the patient and the clinician. In a large number of studies, it has been shown to be influenced by current disease activity and by cumulative damage. Thus, in early disease, the HAQ is reversible whereas in later disease it tends to correlate with radiological scores and be less reversible. It has been shown to be a sensitive outcome measure for use in clinical trials and also to predict future mortality. Thus, it assesses not just the patient's current disability but also aspects of their overall health. The measurement properties of the HAQ have been intensively studied. It is a categorical and not a continuous measure. A patient with a HAQ score of two is not twice as disabled as one with a HAQ score of one. This means that, non-parametric and not parametric statistics should be used to describe the HAQ.

### Health-related quality of life and mortality

Some instruments which assess a patient's overall quality of life are specific for RA and these clearly have advantages certainly from the patient's perspective. However, generic measures enable us to benchmark patients with RA against those with a whole range of other conditions. This is particularly useful in health economic analysis and in making decisions about allocation of resources. The EuroQol (EQ5D) has been most widely used in RA. Recently, the SF-6D has been developed from the SF-36. It is a utility measure which can, therefore, be used in economic analyses. The EQ5D differs from the SF-6D in that it includes states rated as 'worse than death'. Rather surprisingly, a high proportion of patients with early RA rate themselves in these 'worse than death' states. This is mainly due to the high levels of pain reported by these patients.

Patients with RA seldom have this as their only medical problem as almost 70% who are on DMARDs have at least one other chronic disease. It is the interaction between these co-morbid conditions and RA that may lead to premature mortality. Mortality, of course, is the final and most robust outcome measure. Almost all mortality studies in RA have shown excess mortality – even those done in modern times. Patients with RA die from the same causes as the general population but at an earlier age. In other words, they are ageing faster than their contemporaries. It is not clear whether this premature ageing process starts with the development of RA or whether it precedes its onset. Mortality studies done exclusively on patients who have developed their RA since 2000, ie in the biologic era, are awaited with interest.

### References

- 1 National Institute for Health and Clinical Excellence. *Rheumatoid arthritis. National clinical guideline for management and treatment in adults*. London: Royal College of Physicians, 2009.

- 2 Kelly C, Mir F. Biological therapies: how can we afford them? *BMJ* 2009;339:b3276.
- 3 Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcomes in arthritis. *Arthritis Rheum* 1980;23:137–45.
- 4 Thompson P, Kirwan J. Joints count: a review of old and new articular indices of joint inflammation. *Brit J Rheum* 1995;34:1003–8.
- 5 Smolen JS, Breedveld FC, Eberl G *et al*. Validity and reliability of the 28 joint count for the assessment of rheumatoid arthritis activity. *Arthritis Rheum* 1995;38:38–43.
- 6 Smolen JS. The work of the EULAR standing committee on International Clinical Studies Including Therapeutic Trials. *Br J Rheum* 1992;31:219–20.
- 7 Felson D, Anderson JJ, Boers *et al*. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. *Arthritis Rheum* 1993;36:729–40.
- 8 Boers M, Tugwell P, Felson DT *et al*. World Health Organisation and International League of Associations for Rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. *J Rheumatol Suppl* 1994;41:86–9.
- 9 Prevoo ML, Van't Hof MA, Kuper HH *et al*. Modified disease activity scores that include twenty-eight joint counts: Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38:44–8.
- 10 Symmons DPM, Hassell AB, Gunatillaka KAN *et al*. Development and preliminary assessment of a simple measure of overall status in rheumatoid arthritis (OSRA) for routine clinical use. *Quart J Med* 1995;88:429–37.
- 11 van der Heijde DM, van't Hof MA, van Riel PL *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.
- 12 Aletaha D, Ward MM, Machold KP *et al*. Remission and active disease in rheumatoid arthritis: Defining criteria for disease activity states. *Arthritis Rheum* 2005;52:2625–36.
- 13 van Gestel AM, Prevoo ML, van't Hof MA *et al*. Development and validation of the European League against Rheumatism response criteria for rheumatoid arthritis. *Arthritis Rheum* 1996;39:34–40.
- 14 Larsen A, Dale K, Eek M. Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. *Radiol Diagn Stockh* 1977;18:481–91.
- 15 van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. *J Rheumatol* 2000;27:261–73.
- 16 Steinbrocker O, Traeger CH, Batterman RC. Therapeutic criteria in rheumatoid arthritis. *JAMA* 1994;271:659–62.
- 17 Lansbury J. Report of a 3 year study on the systemic and articular indexes in rheumatoid arthritis: theoretical and clinical considerations. *Arthritis Rheum* 1958;1:505–22.
- 18 Mallya RK, Mace BEW. The assessment of disease activity in rheumatoid arthritis using a multivariate analysis. *Rheumatol Rehabil* 1981;20:14–7.

**Address for correspondence: Professor DPM Symmons, Arthritis Research UK Epidemiology Unit, Manchester Academic Health Sciences Centre, University of Manchester, Stopford Building, Oxford Road, Manchester M13 9PT. Email: Deborah.Symmons@manchester.ac.uk**