Inter-observer agreement of standard joint counts in early rheumatoid arthritis: a comparison with grey scale ultrasonography—a preliminary study

F. Salaffi1, E. Filippucci1, M. Carotti2, E. Naredo3, G. Meenagh4, A. Ciapetti1, V. Savic5 and W. Grassi1

Objectives. The aims of the present study were to assess the inter-observer agreement of standard joint count and to compare clinical examination with grey scale ultrasonography (US) findings in patients with early rheumatoid arthritis (RA).

Methods. The study was conducted on 44 RA patients with a disease duration of <2 yrs. Clinical evaluation was performed independently by two rheumatologists for detection of tenderness in 44 joints and swelling in 42 joints. All patients underwent US assessment by a rheumatologist experienced in this method and blinded to the clinical findings. Joint inflammation was detected by US when synovial fluid and/or synovial hypertrophy was identified using OMERACT preliminary definitions. The inter-observer reliability was calculated by overall agreement (percentage of observed exact agreement) and kappa (κ)-statistics. The reliability of US was calculated in 12 RA patients.

Results. There was fair to moderate inter-observer agreement on individual joint counts for either tenderness or joint swelling apart from the glenohumeral joint. US detected a higher number of inflamed joints than did clinical examination. The mean (± s.d.) US joint count for joint inflammation was 19.1 (±4.1), while the mean (±s.d.) number of swollen joints was 12.6 (±3.6), with a significant difference of \( P = 0.01 \).

Conclusions. Our results provide evidence in favour of the hypothesis that clinical examination is far from optimal for assessing joint inflammation in patients with early RA. Furthermore, this study suggests that US can considerably improve the detection of signs of joint inflammation both in terms of sensitivity and reliability.

KEY WORDS: Rheumatoid arthritis, Synovitis, Ultrasonography, Joint count.

Introduction

Accurate assessment of disease activity in rheumatoid arthritis (RA) is essential in the clinical management of RA patients and in RA clinical trials. Counting the number of swollen joints is a clinical method of quantifying the amount of inflamed synovial tissue [1]. Joint counts are included in historical indices of disease activity, such as the Lansbury Index [2], and are a major component of the disease activity score (DAS) [3, 4] and similar indices [5–7], such as the American College of Rheumatology (ACR) Core Data Set for clinical trials in RA [8], the ACR criteria for improvement [9] and the ACR remission criteria [10]. Clinically detectable synovitis antedates joint damage [11], and rheumatologists should include a joint count at each visit for each RA patient [12]. Several studies reported considerable variation in joint counts between both observers and centres in clinical trials and in daily clinical practice [13–15]. Although regarded as an ‘objective’ measure, the joint count is only an indirect assessment of inflammation in the joint. In this respect, imaging modalities such as magnetic resonance imaging (MRI) and high-resolution ultrasonography (US) offer further possibilities to evaluate synovitis and hence disease activity [16, 17]. Recent reports addressing the use of US in the evaluation of RA indicated that clinical joint examination may be inadequate in clinical trials for assessing the reduction in signs and symptoms of RA [18–22]. US has no contraindications, poses no problems regarding patient compliance and allows the examination of more than one anatomic area in a single study. US is a non-invasive, inexpensive and free-of-radiation-hazards imaging technique allowing a quick and sensitive assessment of soft tissue inflammation.

The main aims of the present study were to assess the inter-observer agreement regarding standard joint count and to compare clinical examination with grey scale US findings in patients with early RA.

Methods

Patients

A total of 44 patients with recent-onset RA (disease duration <2 yrs), attending the care facilities of the Department of Rheumatology of the Universita` Politecnica delle Marche, were recruited. Demographic and clinical characteristics of the patients are illustrated in Table 1. The patient selection criteria were as follows: fulfilment of the ACR, (formerly the American Rheumatism Association) 1987 revised criteria for RA [23], age >18 yrs, duration of symptoms <2 yrs and active disease that was defined by the presence of not less than three swollen joints and at least three of the following four features: either an erythrocyte sedimentation rate (ESR) ≥28 mm/h or a C-reactive protein (CRP) level >19 mg/l, morning stiffness ≥29 min, >5 swollen joints and >10 tender joints [24]. Patients who had had traumatic, septic or microcrystalline arthritis, previous joint surgery or isotopic synovectomy in the previous 12 months were excluded. The study was performed according to the principles of the Declaration of Helsinki. The protocols were approved by the ethics committees. Informed consent was obtained from all the patients.

Clinical assessment

Clinical examinations were performed independently and sequentially by two rheumatologists who carried out a consensus on joint assessment before the study: the first (F.S.) with extensive experience in quantitative joint evaluation and the second (A.C.) with experience from 300 supervised joint count examinations of RA patients. The data obtained by the former rheumatologist...
Table 1. Demographic and clinical characteristics of the patients

|                          | Value |
|--------------------------|-------|
| Age, mean ± s.d. yrs     | 53 ± 9.8 |
| Sex, % women             | 72.7  |
| Disease duration, mean ± s.d. months | 17 ± 3.8 |
| No. of swollen joints, mean ± s.d. | 12.6 (±3.6) |
| No. of tender joints, mean ± s.d. | 18.8 (±6.7) |
| Duration of morning stiffness, mean ± s.d. min | 42.6 (±12.6) |
| RF positivity (by nephelometry, with titres of >20 IU/ml), % of patients | 26 |
| ESR, mean ± s.d. mm/h    | 47.1 (±21.7) |
| CRP, mean ± s.d. mg/dl   | 26 (±17.8) |
| DMARD use, no. of patients (%) | 44 (100) |
| Methotrexate             | 33 |
| Leflunomide              | 3     |
| Combination therapy, no. of patients (%) | 21 (47.7) |
| Etanercept plus methotrexate | 10   |
| Adalimumab plus methotrexate | 7     |
| Leflunomide plus methotrexate | 4     |
| NSAID use on demand, no. of patients (%) | 40 (90.9) |
| Steroid use (4–16 mg prednisolone equivalent daily), no. of patients% | 21 (47.7) |

RF, rheumatoid factor; ANA, anti-nuclear antibody; ESR, erythrocyte sedimentation rate; CRP, C-reactive protein; DMARD, disease-modifying anti-rheumatic drug; NSAID, non-steroidal anti-inflammatory drug.

were used for comparison with US findings. The following 42 joints were assessed bilaterally for tenderness and swelling: acromioclavicular, glenohumeral, elbow, wrist (radiocarpal), metacarpophalangeal (MCP), proximal interphalangeal (PIP) of hands, knee, ankle (tibiotalar) and metatarsophalangeal (MTP) joints. Moreover, hip joints were assessed for tenderness on passive motion. The sum of both tender joint count (TJC) and swollen joint count (SJC) was recorded for each patient. Clinical inter-observer agreement for both tenderness and swelling was calculated. The presence or absence of joint swelling was compared with US detection of joint inflammation.

**US assessment**

On the same day of the clinical examination, all patients underwent a US assessment by a rheumatologist experienced in US (E.F.) and blinded to the results of the joint count assessment. US examinations were performed using an AU5 ‘Harmonic’ (Esaote Biomedica, Genoa, Italy) equipped with two broadband linear probes (7.5–10 and 10–14 MHz). The patients were evaluated using the US scanning protocol introduced in a recent paper by Naredo et al. [21]. Joint inflammation was detected by US when synovial fluid and/or synovial hypertrophy was identified using OMERACT preliminary definitions [25]. Figure 1 shows a representative US image illustrating both synovial fluid and synovial hypertrophy at MCP joint level in an RA patient.

Each US examination took <60 min, and representative images were archived. Inter-observer reliability was determined by comparing the findings obtained by the experienced ultrasonographer rheumatologist (E.F.) and those of an experienced radiologist (M.C.) who examined 484 joints in a random subset of 12 patients. Each investigator performed the US examinations independently and sequentially while blinded to all other study data. Intra-observer reliability was assessed by blinded rescorings of the archived US images in the same subset 2 months after the baseline US examination.

**Statistical analysis**

Data evaluation and statistical analysis were performed using SPSS version 11.0 software (SPSS Inc., Chicago, IL, USA) and MedCalc (Belgium, release 9.0) for Windows XP. Normally distributed continuous data were summarized with mean and s.d.
lowest $k$-values, especially in the evaluation of the joint swelling ($k = 0.20$), showing a non-consistent agreement. In the assessment of joint swelling, the highest values of $k$ were found at PIP and MCP joints level (especially at the second finger), with a moderate to good inter-observer agreement. Moderate levels of agreement were found at wrist, knee and ankle joints, with $k$-values of 0.49, 0.56 and 0.50, respectively. Table 3 reports exact inter-observer agreement on the detection of joint swelling at PIP, MCP and MTP joints.

**Inter- and intra-observer agreement between US investigators**

The reliability of US was calculated in 12 RA patients. The inter-observer agreement showed an exact agreement of 91 and 89% for the presence/absence of synovial fluid and synovial hypertrophy, with $k = 0.72$ and $k = 0.66$, respectively. The US assessment of synovial fluid and synovial hypertrophy showed a good intra-observer agreement with $k$-values of 0.70 and 0.61, respectively. An exact agreement of 90 and 85% was found for the presence/absence of synovial fluid and synovial hypertrophy, respectively.

### Table 2. Clinical inter-observer agreement for tenderness and swelling (mean $k$-values for each joint)

| Joint                        | Tenderness | Swelling |
|------------------------------|------------|----------|
| Acromioclavicular            | 0.35       | 0.39     |
| Glenohumeral                 | 0.34       | 0.20     |
| Elbow                        | 0.36       | 0.37     |
| Wrist                        | 0.42       | 0.49     |
| Metacarpophalangeal          |            |          |
| First                        | 0.39       | 0.43     |
| Second                       | 0.44       | 0.65     |
| Third                        | 0.62       | 0.51     |
| Fourth                       | 0.50       | 0.52     |
| Fifth                        | 0.42       | 0.50     |
| Proximal interphalangeal     |            |          |
| First                        | 0.40       | 0.45     |
| Second                       | 0.44       | 0.69     |
| Third                        | 0.46       | 0.50     |
| Fourth                       | 0.59       | 0.76     |
| Fifth                        | 0.59       | 0.52     |
| Hip                          | 0.46       | 0.56     |
| Knee                         | 0.42       | 0.50     |
| Ankle                        | 0.42       | 0.50     |
| Metatarsophalangeal          |            |          |
| First                        | 0.36       | 0.41     |
| Second                       | 0.36       | 0.40     |
| Third                        | 0.36       | 0.42     |
| Fourth                       | 0.31       | 0.40     |
| Fifth                        | 0.35       | 0.44     |

*Assessed for tenderness only.

**Table 3. Clinical inter-observer agreement on the detection of joint swelling**

| Second investigator | Joint swelling (–) | Joint swelling (+) | Total |
|---------------------|--------------------|--------------------|-------|
| PIP joints: percentage of exact agreement (86.8%). | | | |
| Joint swelling (–) | 289                | 11                 | 300 (68.2%) |
| Joint swelling (+) | 47                 | 93                 | 140 (31.8%) |
| Total              | 336 (76.4%)        | 104 (23.6%)        | 440   |
| MCP joints: percentage of exact agreement (80.2%). | | | |
| Joint swelling (–) | 236                | 35                 | 271 (61.6%) |
| Joint swelling (+) | 52                 | 117                | 169 (38.4%) |
| Total              | 288 (65.5%)        | 152 (34.5%)        | 440   |
| MTP joints: percentage of exact agreement (68.8%). | | | |
| Joint swelling (–) | 245                | 60                 | 305 (69.3%) |
| Joint swelling (+) | 77                 | 58                 | 135 (30.7%) |
| Total              | 322 (73.2%)        | 118 (26.8%)        | 440   |

**Table 4. Levels of agreement between mean left/right joint count for swelling and US joint count for joint inflammation ($k$-values)**

| Joint                          | Mean left/right |
|--------------------------------|----------------|
| Acromioclavicular              | 0.35           |
| Glenohumeral                   | 0.20           |
| Elbow                          | 0.57           |
| Wrist                          | 0.46           |
| Metacarpophalangeal            |                |
| First                          | 0.45           |
| Second                         | 0.50           |
| Third                          | 0.53           |
| Fourth                         | 0.52           |
| Fifth                          | 0.53           |
| Proximal interphalangeal       |                |
| First                          | 0.50           |
| Second                         | 0.60           |
| Third                          | 0.51           |
| Fourth                         | 0.49           |
| Fifth                          | 0.57           |
| Knee                           | 0.61           |
| Ankle                          | 0.54           |
| Metatarsophalangeal            |                |
| First                          | 0.47           |
| Second                         | 0.43           |
| Third                          | 0.44           |
| Fourth                         | 0.44           |
| Fifth                          | 0.47           |

**Table 5. Overall agreement between clinical and US assessment of joint inflammation**

| Joint                          | Clinical examination |
|--------------------------------|----------------------|
|                                | US assessment        |
|                                | Joint swelling (–)   | Joint swelling (+) | Total |
| PIP joints: percentage of exact agreement (83.9%). | | | |
| Joint swelling (–)             | 270                  | 66                 | 336 (76.4%) |
| Joint swelling (+)             | 5                    | 99                 | 104 (23.6%) |
| Total                          | 275 (62.5%)          | 165 (37.5%)        | 440   |
| MCP joints: percentage of exact agreement (76.1%). | | | |
| Joint swelling (–)             | 191                  | 97                 | 288 (65.5%) |
| Joint swelling (+)             | 8                    | 144                | 152 (34.5%) |
| Total                          | 199 (45.2%)          | 241 (54.8%)        | 440   |
| MTP joints: percentage of exact agreement (70.2%). | | | |
| Joint swelling (–)             | 220                  | 100                | 320 (72.7%) |
| Joint swelling (+)             | 31                   | 89                 | 120 (27.3%) |
| Total                          | 251 (57%)            | 189 (43%)          | 440   |

### Agreement between clinical and US findings

US detected a higher number of inflamed joints than clinical examination. Hip joints were not included in the analysis because joint swelling was not assessed. US detected signs of joint inflammation in 936 of 1848 joints (50.6%), while clinical examination found 594 swollen joints (32.1%) ($P = 0.005$). The mean (± S.D.) US joint count for joint inflammation was 19.1 (± 4.1), while the mean number of swollen joints was 12.6 (± 3.6), with a significant difference of $P = 0.01$. Table 4 shows the levels of agreement between individual joint count for swelling and US joint count for joint inflammation. The highest $k$-values were found at knee and PIP joints. The $k$-values showing a level of agreement from moderate to good were obtained in all the other joints with the exception of the shoulder. In particular, the $k$-value for the glenohumeral joints was 0.20, showing poor agreement.

Table 5 reports the percentage of observed exact agreements between clinical examination (joint swelling) and US assessment (detection of synovial fluid and/or synovial hypertrophy) of PIP, MCP and MTP joints. At PIP joint level, US signs of joint inflammation were detected in 99 joints in which clinical examination found joint swelling (22.5%), whereas no signs of inflammation were seen by US in 270 clinically non-swollen
joints (61.4%) with an overall agreement of 83.9%. At MCP joint level, US found inflammation in 144 swollen joints (32.7%) and did not find inflammation in 191 non-swollen joints (43.4%), with an overall agreement of 76.1%. At MTP joint level, the US and clinical findings were in agreement on the presence of signs of inflammation in 89 joints (20.2%) and on the absence of signs of inflammation in 220 joints (50%), with an overall agreement of 70.2%.

Discussion

In the present study, we investigated the extent of agreement in detecting the presence of joint swelling and tenderness in patients with early RA. There was extensive variability in the number of both swollen and tender joints. Difference of joints between observers on the scoring range recorded by the two observers in individual patients were often high, indicating considerable differences between observers. In particular, we were able to show that shoulders were far more often involved in discordant observations when compared with concordant observations. The shoulder is a challenging anatomic area to be assessed by the rheumatologist. The deep location of the glenohumeral joint makes joint effusion difficult to be detected, especially in obese patients [27]. A further possible explanation for the low accuracy of the clinical assessment in this cohort and in other series [27, 28] could be found in the poor correlation between clinical findings and anatomical abnormalities in the shoulder. US examination found in most of the swollen shoulders the presence of a sub-deltoid bursitis, while the US detection of inflamed glenohumeral joint was frequently underestimated by the clinical examination. In a recent paper, focusing on the correlation between glenohumeral joint swelling detected on physical examination and effusion revealed by US in patients with RA, Luukkainen et al. [27] found a $\kappa$-coefficient of 0.202, which is exactly what we found.

Our results are consistent with those reported by Szkudlarek et al. [29] for the MCP joints. In their study, the overall agreement on the presence or absence of signs of inflammation between US and clinical assessment was 63%. Similarly, Luukkainen et al. [18] evaluated the relationship between clinically detected joint swelling and joint effusion detected by US in MTP joints and talocrural joints in patients with RA and showed poor agreement. Further, in the original work by Koski [30], there was also some overlap between the normal and synovitic values in MTP joints. MTP joints represent challenging joints to be assessed by clinical examination. US have recently been proved to be a reliable tool for evaluating joint inflammation at MTP joints [29]. Despite lower values of $\kappa$-inter-observer agreement, our results are similar to those found by Szkudlarek et al. [29]. Thus, we believe that MTP joints should be included in joint counting both clinically and by US. Moreover, US provides for a sensitive detection of bone erosions, especially in patients with early RA, especially at the fifth MTP joint level [29].

The increasing use of high-cost biological treatments focuses attention on clinical assessments. Reports from both the Health and European League Against Rheumatism (EULAR) [31] and the National Institute for Health and Clinical Excellence [32] recommend basing the decision to treat patients with anti-tumour necrosis factor therapy on DAS28 scores. However, missing the information from the MTP joints that are frequently and may be primarily affected in early RA [16, 33], could jeopardize biometric reliability of the composite index. Fransen et al. [34], in fact, demonstrated in their clinical follow-up cohort that the DAS performed better than the DAS28 in detecting remission. Similarly, Makinen et al. [35] and Landewe et al. [36] suggested that DAS28 has insufficient construct validity and should be used with consideration in clinical practice and in clinical trials.

US was more sensitive than clinical examination in detecting joint inflammation. However, its higher sensitivity may vary considerably according to the selected joint. At small joints of the hand and feet, the relatively high number of clinical swollen joints not inflamed as found by US (Table 5) was mainly due to the presence of other pathology detectable by US, such as tenosynovitis, periarticular soft tissue oedema or osteophytes. Longitudinal researches aiming at investigating the value of US findings of joint inflammation in patients who satisfy the remission criteria with normal findings on clinical and laboratory studies are required. Imaging assessment, such as US, may be necessary for the accurate evaluation of disease status and, in particular, for the definition of true remission [37–39].

The present study has the following limitations. First, because it is time-consuming, the US examination of the 44 joints did not include power Doppler assessment. Second, during US examinations, no distinction was made between normal and pathological synovial fluid. The presence of an even minimal amount of fluid within the joint cavity, fulfilling the OMERACT preliminary definitions, was considered abnormal. This may have led to an overestimation of the joint inflammation as some intra-articular non-inflammatory fluid collection could have been interpreted as pathological.

Further limitations to this study, which must be emphasized, are that the data derive from a single clinical trial, and further analyses of additional clinical trials are required to determine whether the results are generalizable.

In conclusion, our data showed that joint US examination was more sensitive than clinical examination in the detection of joint inflammation in patients with early RA. We have also shown that US is a reproducible method of assessing joint inflammation, with good levels of agreement between readers. This is consistent with published data [37] from studies of patients with increased levels of synovitis. The present study strongly encourages the use of US to improve joint assessment in patients with RA in daily management and clinical trials. The enhanced sensitivity of US will probably lead to a re-adjustment of the ‘synovitis thermostat’, with more patients classified as having polyarthritis and fewer as pathological.

Rheumatology key messages

- This study showed fair to moderate inter-observer agreement on individual joint counts.
- US detected a higher number of inflamed joints than clinical examination.
- The use of US may improve joint assessment in patients with RA.

Disclosure statement: The authors have declared no conflicts of interest.

References

1 Sokka T, Pincus T. Quantitative joint assessment in rheumatoid arthritis. Clin Exp Rheumatol 2005;23:S58–62.
2 Lansbury J. A method for summation of the systemic indices of rheumatoid activity. Am J Med Sci 1956;232:300–10.
3 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;
4 van der Heijde DMFM, van't Hof MA, van Riel PLCM, van der Putte LBA. Development of a disease activity score based on judgment in clinical practice by rheumatologists. J Rheumatol 1993;20:579–81.
5 Aletaha D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. Clin Exp Rheumatol 2005;23:S100–8.
6 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.
7 Salaffi F, Peroni M, Ferraccioli GF. Discriminating ability of composite indices for measuring disease activity in rheumatoid arthritis: a comparison of the Chronic Arthritis Systemic Index, Disease Activity Score and Thompson’s articular index. Rheumatology 2000;39:90–6.
8 Felson DT, Anderson JJ, Boers M et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis
Kane D, Grassi W, Sturrock R, Balint PV. Musculoskeletal ultrasound—a state of the art review in rheumatology. Arthritis Rheum 1993;36:729–40.
9 Felson DT, Anderson JJ, Boers M et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1981;24:1209–15.
10 Sokka T, Kautiainen H, Moltonen T, Hannonen P. Erosions develop rarely in joints without clinically detectable inflammation in patients with early rheumatoid arthritis. J Rheumatol 2003;30:2580–4.
11 Scott DL, Antoni C, Choy EH, van Riel PLCM. Joint counts in routine practice. Rheumatology 2003;42:919–23.
12 Hart LE, Tugwell P, Buchanan VW, Norman GR, Grace EM, Southwell D. Grading of tenderness as a source of interrater error in the Ritchie articular index. J Rheumatol 1985;12:716–7.
13 Lewis PA, O'Sullivan MM, Rumfeld WR, Coles EC, Jessop JD. Significant changes in inflammatory activity in rheumatoid arthritis: a comparative study of clinical evaluation with ultrasonography and conventional radiography, and clinical examination. Arthritis Rheum 2004;50:2103–12.
14 Thompson PW, Hart LE, Goldsmith CH, Spector TD, Bell MJ, Ramsden MF. Comparison of four articular indices for use in clinical trials in rheumatoid arthritis: patient, order and observer variation. J Rheumatol 1991;18:661–5.
15 Rutland AL, Szkudlarek M, Boers M, Dinant HJ, Bezemer PD, van Riel PLCM. Retardation of joint damage in patients with early rheumatoid arthritis by initial aggressive treatment. Ann Rheum Dis 2005;64:375–81.
16 Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
17 Luukkanen R, Sanila MT, Luukkanen P. Poor relationship between joint swelling detected on physical examination and effusion diagnosed by ultrasonography in glenohumeral joints in patients with rheumatoid arthritis. Clin Rheumatol 2007;26:865–7.
18 Naredo E, Aguado P, De Miguel E et al. Painful shoulder: comparison of physical examination and ultrasonographic findings. Ann Rheum Dis 2002;61:362–6.
19 Thompson PW, Hart LE, Goldsmith CH, Spector TD, Bell MJ, Ramsden MF. Comparison of four articular indices for use in clinical trials in rheumatoid arthritis: patient, order and observer variation. J Rheumatol 1991;18:661–5.
20 Boutry N, Larde A, Lapegue F, Solau-Gervais E, Flipo RM, Cotten A. Magnetic resonance imaging of the metacarpophalangeal joints of the hands remains normal. Arthritis Rheum 1988;31:315–24.
21 Landewe R, van der Heijde D, van Riel PLCM. Retardation of joint damage in patients with early rheumatoid arthritis by initial aggressive treatment with disease-modifying antirheumatic drugs: five-year experience from the FIN-RACo study. Arthritis Rheum 2004;50:2072–81.
22 Wakefield RJ, Balint PV, Szkudlarek M et al. OMERACT 7 Special Interest Group. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. J Rheumatol 2005;32:2458–7.
23 Arnett FC, Edworthy SM, Bloch DA et al. The American Rheumatism Association FIN-RACo Trial Group. Retardation of joint damage in patients with early rheumatoid arthritis by initial aggressive treatment. Arthritis Rheum 1993;36:729–40.