Standardising Joint Assessment in Rheumatoid Arthritis


Summary
Evaluating joint involvement in rheumatoid arthritis in a key clinical assessment. We investigated the extent of variation in measurement of joint swelling and tenderness and evaluated the impact of training to standardise methods. Eight observers (medical and nursing staff) examined eight rheumatoid patients for joint swelling and tenderness before and after training in clinical methods. The EULAR handbook for joint evaluation was used for training and assessments were based on the 28 joint count. There was extensive variability in both numbers of swollen and tender joints. Coefficients of variation for articular indices recorded by the 8 observers in individual patients were often high (up to a maximum of 204%), indicating considerable differences between observers. Training had an impact on the assessment of the numbers of swollen joints which increased by a mean of 32% (P < 0.05) and the number of tender joints which increased by 41% (P < 0.01). Training had only a limited impact on the variation among observers in determining the number of swollen and number of tender joints. After training, the mean coefficients of variation were still 59% for swollen joints and 65% for tender joints. These results highlight the extent of variation in clinical assessment of rheumatoid arthritis and show the advantages of training. It leads to increased sensitivity of measurement. Standardisation appears essential for clinical studies.

Key words
Rheumatoid Arthritis, Articular Index, Joint Assessment

INTRODUCTION
In recent years there has been international agreement on the core data set for the clinical assessment of rheumatoid arthritis (RA) (1,2). Evaluating joint involvement is a central component of the core data set and includes determining both the number of swollen joints and the number of tender joints. Standardisation is needed. To help achieve such standardisation EULAR have developed a handbook of clinical methods to assess joint swelling and tenderness (3). How great is the natural variability between clinicians in the measurement of joint swelling and tenderness in RA using an internationally accepted method? Does simple training to standardise methods greatly reduce this variability? We examined these questions in a simple clinical study.

METHODS
Eight observers evaluated the same 8 RA patients in an outpatient clinical setting. The observers consisted of 5 medical staff (from consultant to registrar grade) and 3 nurses involved in clinical trial work. All observers were regularly involved in assessing RA patients for entry into clinical trials of slow-acting anti-rheumatic drugs or biological agents. The patients met the diagnostic criteria of RA of the American College of Rheumatology (4). They were selected as being current outpatient attenders who had RA which was partially controlled by slow-acting anti-rheumatic drugs. None had a very active disease or were in remission.

Joint swelling and tenderness were evaluated on two occasions before and after training in clinical methods. On both occasions the joints assessed were in the OMERACT core data set and analysis was based on the 28 joint count as recommended by EULAR. On neither occasion was any grading of counts at individual joints used. Patients were seen by different observers in random order.
Fig. 1: Changes in swollen joint counts and tender joint counts with training. Means and SDs are shown.

Training lasted 60 minutes and was based on the material in the EULAR handbook for joint evaluation. This explains how to detect joint swelling and tenderness and how to assess each joint. Care was taken to explain exactly how to examine the joints in the 28 joint count in a standardised manner (these consist of 10 proximal interphalangeal joints, 10 metacarpophalangeal joints, 2 wrists, 2 elbows, 2 shoulders and 2 knees).

The results were expressed as mean number of swollen joints (with standard deviations) given by the 8 observers for each patient. Coefficients of variation were also given for the relationships between these means and standard deviations. The results were compared between the first and second assessments by a two-way analysis of variance examining changes in the 8 patients using the results from the 8 individual observers as replicates.

**RESULTS**

At the initial assessments the mean number of patients' swollen joints was between 5-12 and that of tender joints 1-11. Taking each observer's result as a replicate reading, the mean number of swollen joints in the 8 patients found in the first assessment was 7.6 and the mean number of tender joints 6.6.

When these observations were repeated after a period of training, the assessment of both the mean numbers of swollen and tender joints increased (Figure). Training increased assessment of swollen joints by a mean of 32% and of tender joints by 41%. An analysis of variance showed highly significant differences between patients and period of observation but no effect from the different observers (Table).

Although there was no significant effect indicated by different observers on the mean joint counts, there was considerable variation between observers and this was not affected by training. The initial coefficients of variation between each observer in the different patients for both the number of tender and the number of swollen joints varied from 21-187%. The mean coefficient of variation for all patients was 82% for swollen joints and 66% for tender joints. After training, the coefficients of variation for both the number of tender and the numbers of swollen joints varied from 22-204%. The mean coefficients of variation after training were 59% for swollen joints and 65% for tender joints.

**DISCUSSION**

Our results highlight the extent of variation in clinical assessment of joint involvement in RA. There are two principal conclusions. First, training leads to increased sensitivity of measurement. Overall changes in the numbers of swollen and tender joints in the region of 30-40% are what would be expected with an active antirheumatic drug and are of clinical significance. The finding that training can have an effect of such magnitude