Clinical judgment in rheumatoid arthritis. IV
Rheumatologists’ assessments of disease remain stable over long periods

J. R. KIRWAN AND H. L. F. CURREY
From the Bone and Joint Research Unit, The London Hospital Medical College, London E1 2AD

SUMMARY  Seven rheumatologists made judgments about the improvement or deterioration of identical sets of 50 ‘paper’ patients on two occasions one year apart. The stability of their judgments over one year ($r_s=0.70$) compared favourably with the reliability of duplicate judgments on each occasion ($r_s=0.76$). Multiple regression analysis of the patient data in relation to the disease assessments provided a model of each clinician’s underlying judgment policy. The stability of judgments predicted by these policy models was even higher ($r_s=0.83$).

Key words: judgment policy models, judgment analysis.

Rheumatologists’ assessments of disease activity in rheumatoid arthritis (RA) have been shown to differ widely,1–3 yet such judgments are fundamental to patient management decisions and clinical research. The possibility that this collective divergence of opinions4 and practice3 might conceal a more ordered approach by individual rheumatologists prompted this study, which investigates the stability of judgments made by individual clinicians one year apart.

Accepted for publication 13 April 1984.
Correspondence to Dr J. R. Kirwan, Bone and Joint Research Unit, ARC Building, 25–29 Ashfield Street, London E1 2AD.

Materials and methods
Clinical data taken largely from real patients were used to construct 30 paper patients4 5 as previously described.3 Before and after treatment values were included for 10 clinical variables (Table 1). Twenty randomly selected patients were duplicated, producing 50 in all. (There were no consistent differences between the 20 duplicate patients and the 10 included only once.) These were sent on two occasions one year apart to 10 randomly selected British rheumatologists. On each occasion they were invited to indicate their judgment, using a visual

Table 1  Clinical variables included in patient data

<table>
<thead>
<tr>
<th>Clinical variables</th>
<th>Range (× step size)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirin (or any other non-steroidal drug) usage</td>
<td>0–16 × 1 per day</td>
</tr>
<tr>
<td>ESR</td>
<td>10–100 × 1 mm/h</td>
</tr>
<tr>
<td>Haemoglobin</td>
<td>8–15 × 0.1 g/dl</td>
</tr>
<tr>
<td>Number of erosions on hand x-rays</td>
<td>0–24 × 1</td>
</tr>
<tr>
<td>Patient’s assessment of morning stiffness</td>
<td>0, 5, 15, 30–360 × 30 min</td>
</tr>
<tr>
<td>Patient’s global assessment of current disease activity</td>
<td>0–4 × 1 (0=lowest possible, 4=highest possible)</td>
</tr>
<tr>
<td>Patient’s pain score</td>
<td>0–4 × 1 (0=none, 4=worst possible)</td>
</tr>
<tr>
<td>Physician’s assessment of functional capacity</td>
<td>1–4 × 1 (1=independent, 4=totally dependent)</td>
</tr>
<tr>
<td>Physician’s measurement of grip strength</td>
<td>75–450 × 5 mmHg</td>
</tr>
<tr>
<td>Physician’s measurement of Ritchie articular index</td>
<td>0–78 × 1</td>
</tr>
</tbody>
</table>

Values were recorded before and after one year of treatment with an unspecified ‘second-line’ agent.
analogue scale of the improvement or deterioration shown by each patient.

All possible subsets multiple regression analysis was used to derive judgment policy models for each clinician on each occasion. These models are linear regression equations which relate the clinical data to the assessments made by each rheumatologist. Comparisons of judgments of duplicate patients on each occasion (reliability) and of repeated judgments one year apart (stability) were made using the Spearman rank correlation coefficient.

**Results**

Seven rheumatologists returned completed sets of judgments on both occasions. The reliability of judgments (the correlation between assessments of duplicate patients on each occasion) and the stability of judgments made on the same patients one year apart are shown in Table 2. Reliability and stability were highly correlated ($r_s=0.79$, $p<0.01$). How much of the variance in judgments was explained by the policy model equations ($R^2$) is shown for each rheumatologist on each occasion in Table 3, together with the correlations between the judgments predicted by each model for each occasion (stability of model predictions). (The policy model equations give the same judgment prediction for the same patient data, thus their reliability is $r_s=1.00$.) The stability of model predictions is significantly greater ($p<0.05$) than that of actual judgments for each rheumatologist (compare Tables 2 and 3).

**Discussion**

Judgments of change in RA by individual rheumatologists, though differing considerably from each other, are shown here to be reasonably stable over a long period. The correlation between judgments made one year apart (0.70), although less than the upper limit given by that for repeated judgments of the patients on the same occasion (0.76, $p<0.05$) nevertheless compares favourably with it. In addition, rheumatologists who showed higher within-occasion reliability also had higher stability of judgments, suggesting that those who follow a judgment policy more consistently also tend to maintain that policy.

Although a small sample, the 10 rheumatologists invited to take part in the study were selected randomly from those practising in Britain. Self-selection by the seven who completed both parts of the study may have introduced some bias, but it seems likely that the results have general applicability. Using paper patients has been shown to elicit judgments very similar to those made when the real patients they represent are assessed in a clinical setting, and the data used in this study were taken largely from real patients undergoing treatment for their arthritis. As a result naturally occurring intercorrelations between clinical variables have been preserved. These considerations suggest that the findings are a true reflection of normal rheumatological practice.

In previous studies of rheumatologists’ judgments, reliability has been reported only for within-occasion assessments. The temporal stability of judgments in general has only recently been studied. In an investigation of judgments made by non-medical undergraduates on two occasions two weeks apart the stability correlation coefficient was 0.72.
The derivation of multiple regression equations as judgment policy models provides a means of reducing the random inconsistencies of judgment and clarifies the way in which judgment policy relates to the patient data. As has been previously found, the variance explained by the rheumatologists' models was high (65–88%), indicating that the models had effectively identified the underlying judgment policies. On each occasion a model was calculated for each rheumatologist, and this was used to predict his 'consistent' judgment for each patient. These judgments proved to be very stable between occasions (pooled $r_s=0.83$), a finding also reported for the undergraduate students.7

There are many potential clinical and research applications for judgment modelling. Examples include maintaining stable assessment policies for the management of patients over long periods and improving agreement between physicians cooperating in clinical trials or other investigations.2 The feedback provided by the modelling process may also be helpful in education by improving agreement between student and teacher.9 In practice the analysis of clinical judgment in the circumstances of rheumatological practice will involve collecting data about real patients and will inevitably be a long-term process spread over weeks or months. The fact that rheumatologists' judgments are stable means that the accumulation and analysis of data in this way is justifiable. It also raises the possibility that changes in judgment policy may be worthwhile and result in prolonged, rather than transitory, effects.

Finally, there is reassurance in the discovery that rheumatologists are essentially orderly in their individual approach to patient assessment.

References

Clinical judgment in rheumatoid arthritis. IV. Rheumatologists' assessments of disease remain stable over long periods.

J R Kirwan and H L Currey

*Ann Rheum Dis* 1984 43: 695-697
doi: 10.1136/ard.43.5.695

Updated information and services can be found at:
[http://ard.bmj.com/content/43/5/695](http://ard.bmj.com/content/43/5/695)

**Email alerting service**

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Notes**

To request permissions go to:
[http://group.bmj.com/group/rights-licensing/permissions](http://group.bmj.com/group/rights-licensing/permissions)

To order reprints go to:
[http://journals.bmj.com/cgi/reprintform](http://journals.bmj.com/cgi/reprintform)

To subscribe to BMJ go to:
[http://group.bmj.com/subscribe/](http://group.bmj.com/subscribe/)