Annals of the Rheumatic Diseases, 1983, **42**, 644–647

Clinical judgment in rheumatoid arthritis. I. Rheumatologists' opinions and the development of 'paper patients'

J. R. KIRWAN,[1] D. M.CHAPUT DE SAINTONGE,[2] C. R. B. JOYCE,[3] AND H. L. F. CURREY[1]

From the [1]Bone and Joint Research Unit and [2]Department of Therapeutics, the London Hospital Medical College, London E1 2AD, and the [3]Medical Department, Ciba-Geigy AG, 4002 Basle, Switzerland

SUMMARY Opinions about the importance of various measures of disease activity in rheumatoid arthritis gathered from a survey of 20% of British rheumatologists showed a wide diversity for all clinical variables. 'Paper patients' have been developed as a method of investigating actual clinical decisions rather than expressed opinions. Assessments based on 'paper patients' correlate highly ($r=+0{\cdot}901$) with those made on the equivalent real patients when seen in person.

Little is known about which criteria of disease activity in rheumatoid arthritis (RA) are actually used by practising rheumatologists nor to what extent each contributes to the clinician's assessment of the patient's progress. The American Rheumatism Association introduced criteria for the diagnosis of RA in 1958,[1] but these are not concerned with disease severity. Most investigators collect large quantities of information about their patients,[2 3] but evidence has been provided by at least one study[4] that the individual rheumatologist uses only a few clinical variables when making judgments about patient status. A large group of British rheumatologists were therefore surveyed in order to gather opinions about the importance of various measures of disease activity. The diversity of replies prompted the development of 'paper patients' as a method of investigating actual clinical decisions rather than expressed opinions.

Materials and methods

SURVEY

Rheumatologists attending the combined British rheumatology societies' meeting in 1980 were invited to complete a simple form that listed clinical variables relevant to RA (Table 1). Respondents were requested to score each variable on a 0–100 scale according to their own estimate of its importance in assessing changes in rheumatoid disease activity. The variables listed were included because each had previously been mentioned by at least one of 15 clinicians asked to provide a list of important variables. A space was provided for the addition of any other variables considered important.

Table 1 *Clinical variables relevant to RA and included in the survey of rheumatologists*

Early morning stiffness	(EMS)*
Pain	(Pain)*
Patient's global assessment	(Global)*
Erythorcyte sedimentation rate	(ESR)*
Patient's disability	(FC)*
Haemoglobin	(Hb)*
Number of stiff joints	(Joints)*
Extra-articular signs	
Number of erosions	(*x*-ray)*
Articular index	(AI)*
Grip strength	(Grip)*
Side effects of therapy	
Latex test for rheumatoid factor	(Latex)
C-reactive protein	
Immunoglobulins	(Ig)

*These items were specifically mentioned on the forms used to develop the 'paper patients'.
FC=functional capacity.

DEVELOPMENT OF 'PAPER PATIENTS'

All patients suffering from RA attending a busy rheumatology outpatient department over a 2-week period were included in this part of the study. The clinician seeing each patient was asked to record on a simple form any information that was available to him about the patient that might affect his or her assessment of disease activity. Some clinical variables

Accepted for publication 22 November 1982.
Correspondence to Dr J. R. Kirwan.

(those obtaining the 10 highest average scores in the survey) were specifically mentioned on the form (Table 1), but it was stressed that no extra effort should be made to obtain any of this information if it would not have otherwise been available and used. There was space to record any other observations which the clinician felt might be influencing his assessment. A horizontal 10 cm visual analogue scale (VAS) labelled 'no disease activity' and 'maximum possible disease activity' was provided. On this the clinician recorded his assessment of 'current disease activity' in the patient before him.

Some weeks later the same clinicians were each presented with copies of all the forms from the patients he or she had seen containing only the basic information shown in Table 1. Some forms were duplicated. All these forms were interleaved with forms containing information from patients seen by other clinicians. A new VAS was provided, and the clinician recorded his assessment of 'current disease activity' in the patient represented by the data on paper before him (a 'paper patient'). The cases were dealt with in a preset order, and there were sufficient numbers to ensure that the duplicate sets were not recognised. Correlations were calculated on scores following angular transformation.[5]

Results

SURVEY

Seventy-four rheumatologists returned a completed questionnaire, about 40% of those attending the meeting. Analysis of the results by means of absolute scores, relative scores (reported here), or ranking of scores gave the same general result. Many respondents replied ambiguously to 3 of the clinical variables (C-reactive protein, drug side effects, and extra-articular signs), and these are therefore difficult to interpret. The distribution of scores for the remaining 12 clinical variables is shown in Fig 1. There is a wide range in scores allocated to all the variables, and a substantial group of doctors awarded few or zero marks to many items—for example, immunoglobulins, articular index—to which others gave a very high score.

DEVELOPMENT OF CLINICAL JUDGMENT INVESTIGATION

The 9 rheumatologists conducting outpatient clinics saw and assessed 'current disease activity' in a total of 70 patients with RA. Each subsequently assessed the information recorded at the time they had seen these patients personally. The correlation between the judgment of severity in the 'real' (interview) and equivalent 'paper' patients was r=+0·901 (Fig. 2). The correlation coefficients were equally high (range

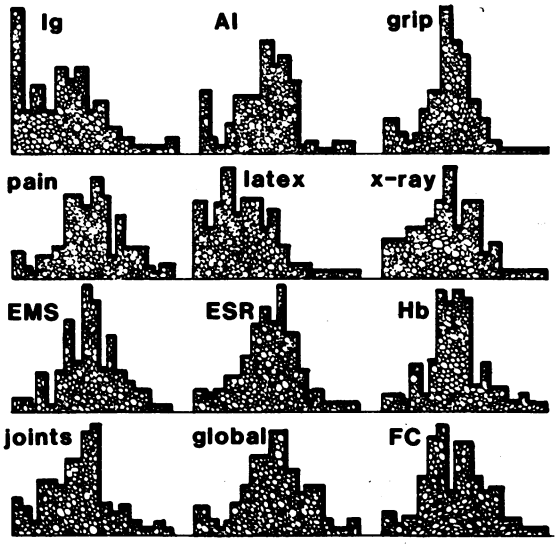

Fig. 1 *Distribution of scores for 12 clinical variables included in the survey. For each histogram the horizontal axis represents the relative scores awarded by respondents and tthe vertical axis represents the number of respondents awarding a particular score. For key see Table 1.*

r=0·770 to 0·962) for all but one doctor (who contributed judgments on only 6 patients). Scores given to duplicates of 'paper' patients were even more highly correlated: r=+0·971 (Fig. 3). Many 'paper patients' were also scored by several different clinicians, and the variation between their judgments was similar throughout the range of disease activity.

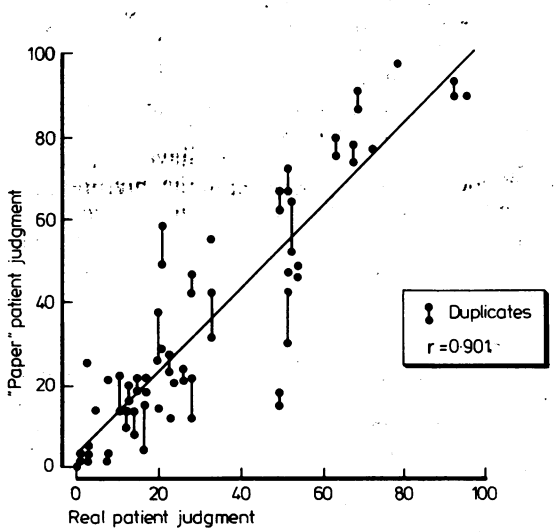

Fig. 2 *Correlation between VAS scores (0—100 mm) for real and equivalent 'paper' patients when judging 'current disease activity'.*

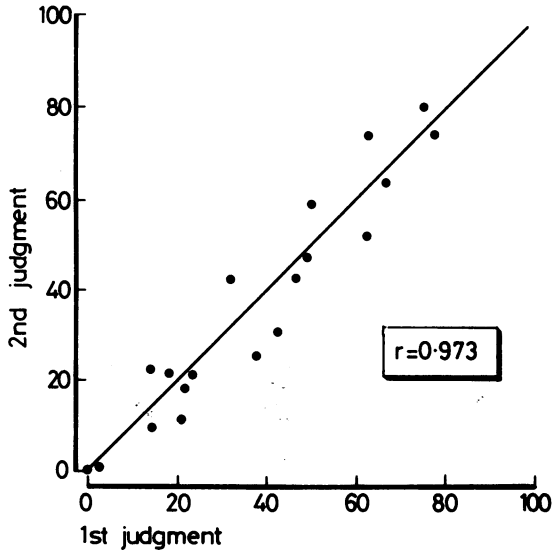

Fig. 3 *Correlation between VAS scores (0—100 mm) for duplicate 'paper patient' judgments.*

Discussion

Although the sample of rheumatologists completing the survey was self-selected, it was nevertheless a high proportion (about 20%) of all British rheumatologists. An enormous diversity of opinion is evident, even about measures which have traditionally been accepted as measures of disease activity, such as the erythrocyte sedimentation rate, as well as those of more recent introduction, such as immunoglobulins. It is possible (perhaps even probable) that opinions expressed in a survey do not accurately reflect actual practice. An equivalent diversity in practice would be important in the interpretation of clinical trials as well as their design and execution. It might also be important in explaining conflicting views and differences in practice between rheumatologists.

In order to explore the existence and possible importance of such differences in practice a suitable testing system is needed that would isolate judgments from the process of gathering information, should allow comparisons between clinicians, should allow repeated judgments, and should reflect well the judgments clinicians would make when seeing real patients. We believe that the development of 'paper patients' meets these requirements.

During the 'paper patients' study rheumatologists recorded only information readily available to them in their routine outpatient clinics. Although an opportunity was always provided to record additional information considered by the physician to be contributing to the judgment of severity in a given case, few notes of this kind were made. Such information was omitted from the 'paper patient' records to prevent them being easily identified. For most patients (65%) only 5 clinical variables were recorded (early morning stiffness, pain score, patient's overall opinion, doctor's assessment of functional capacity, and doctor's articular index score), but there was nevertheless a high correlation between judgments made on paper and those based on the real patient. Judgments on 'paper patients' were also highly reproducible, and their variability between clinicians was independent of the apparent disease activity.

More complicated methods of representing patients, developed for use in general problem solving, include patient management problems,[6] computer simulations,[7] role playing oral examinations,[8] and simulated patients.[9] Although these methods have been designed to resemble the clinical setting, their use has not been directly compared with performance with real patients. 'Paper patients', while simple in design and apparently unlike clinical circumstances, are in fact a valid representation of real patients and provide a useful tool for the further investigation of actual clinical judgment.

Analysis of clinical judgment may elucidate the nature of the differences in judgment patterns which lead clinicians to disagree about the assessment of a particular patient and will pinpoint those items of patient information which contribute to judgment of disease activity. Analyses such as these require judgments to be made by different clinicians on the same set of patient observations. Only by using 'paper patients' does this become a feasible proposition.

We thank the members of the Department of Rheumatology, The London Hospital, for their help and co-operation.

J. R. Kirwan holds an Arthritis and Rheumatism Council Research Fellowship. D. M. Chaput de Saintonge gratefully acknowledges the support of the Nuffield Foundation.

References

1 Ropes M W, Bennett G A, Cobb S, Jacox R, Jessar R A. Revision of diagnostic criteria for rheumatoid arthritis. *Bull Rheum Dis* 1958; **9**: 175–6.

2 Multicentre Study Group. Levamisole in rheumatoid arthritis. *Ann Rheum Dis* 1982; **41**: 159–63.

3 Dixon J S, Pickup M E, Lowe J R, Hallett C, Lee M R, Wright V. Discriminatory indices of response of patients with rheumatoid arthritis treated with D-penicillamine. *Ann Rheum Dis* 1980; **39**: 301–11.

4 Joyce C R B, Berry H, Chaput de Saintonge D M, Domenet J, Fowler P, Mason R M. Judgement analysis of investigator's assessments: a way to reduce one important source of error in multicentre trials. In: Fehr K, Huskisson E C,Wilhelmi E, eds. *International Co-ordination of Drug Trials.* EULAR Bulletin, Monograph 1/1977.

5 Armitage P. *Statistical method in medical research.* Oxford: Blackwell, 1971: 356–7.

6 McGuire C H, Babbott D. Simulation technique in the measurement of problem-solving skills. *J Educ Measurement* 1967; **4:** 1–10.

7 Hubbard J P. *Measuring medical education.* Philadelphia: Lea and Febiger, 1971.

8 Lamont C T, Hennen B K E. The use of simulated patients in a certification examination in family medicine. *J Med Educ* 1972; **47:** 789–95.

9 Barrows H S. *Simulated patients.* Illinois: Thomas, 1968.