

Validity of single variables and composite indices for measuring disease activity in rheumatoid arthritis

Désirée M F M van der Heijde, Martin A van't Hof, Piet L C M van Riel, Miek A van Leeuwen, Martin H van Rijswijk, Levinus B A van de Putte

Abstract

There is no agreement as to which variable best mirrors disease activity in rheumatoid arthritis (RA) and no studies have been performed on the validity of disease activity variables. In this study the validity of 10 commonly used single variables and three composite indices was tested.

All patients participated in a large follow up study in two clinics. The patients (n=233) had classical or definite RA and a disease duration of less than one year at entry. The mean follow up time was 30 months; the follow up frequency was once every four weeks; 6011 records were used in the analysis. The validation criteria included correlations with the other variables (correlational validity), with the physical disability (criterion validity I), and with the radiographically determined damage of hands and feet (construct validity). The judgment of a group of rheumatologists in clinical practice was also used as a model of criterion validity (II).

In this comparison the disease activity score and Mallya index showed the best validity. The best single variable was the number of swollen joints. The validity of most single variables was poor and these variables were not suitable as single endpoint measures in clinical trials.

Rheumatoid arthritis (RA) is a heterogeneous disease with a greatly varying clinical course and outcome. Two aspects are important in evaluating the disease: the disease activity and the ultimate outcome. The assessments to evaluate these features are often called 'process' and 'outcome' measurements, respectively.¹ Radiographic abnormalities and physical disability are two tools used to evaluate outcome and much research has been carried out on these methods. Methods to score radiographic damage have been widely tested—for example, for inter- and intraobserver variations.^{2,3} Questionnaires to measure physical disability have been developed and extensively tested and these are valid assessments for measuring outcome.^{4,5}

Process measurements are widely used in the evaluation of short term clinical trials. Although rheumatologists have discussed in workshops which endpoint measures should be used in clinical trials, no agreement exists,^{6,7} and in most clinical trials many variables are used for evaluation. This may cause conflicting results and, in addition, a correction for multiple testing is needed. This requires a larger group of patients to ensure enough power for the study. Single clinical and laboratory variables

are used most in measuring the disease process. Some composite indices have also been developed, such as the Mallya and Riel indices and the disease activity score.⁸⁻¹⁰ In contrast to the questionnaires, little research has been carried out on the validity of the various variables and indices. The validity of a variable means that the variable is measuring what it is supposed to measure. For a process variable in rheumatoid arthritis, this variable must measure 'disease activity'. This disease activity must be related to the ultimate result of the disease process.

A problem in the study of the validity of process variables in rheumatoid arthritis is that a standard for disease activity is not available. Analogous to the validation of questionnaires on physical disability,^{4,5} we have chosen an explorative method, resulting in a combination of several aspects of validity. This method for validation is widely used in the social and behavioural sciences and has been recommended by Tugwell and Bombardier for the selection of endpoint measures in clinical trials.¹¹ The overall results of all these facets of validity give the ultimate conclusion on the validity of a particular process variable.

Much research has been carried out on disease activity variables and on judgment analysis,^{12,13} although usually only one aspect of disease activity has been evaluated. For example, the possible effect of a new assessment was checked by the correlation with another disease activity variable such as the Ritchie score or the acute phase response.^{14,15} However, the reliability of these 'reference' disease activity variables is not well enough known; little research has been carried out on the comparison of different laboratory and clinical disease activity variables.^{16,17} As far as we know, no study has compared many laboratory and clinical disease activity variables and tested their ability to evaluate the ultimate outcome, measured by radiographically detected damage and physical disability in a large database of patient records.

The aim of this study was to compare the validity of 10 single variables and three composite indices. The disease activity variables were chosen arbitrarily from the large number of assessments available. We selected especially those measures which are frequently used to measure disease activity in clinical trials and in clinical practice. The most valid variable to measure disease activity in clinical trials and clinical practice was determined.

Patients and methods

PATIENTS

All patients took part in an ongoing prospective

Department of
Rheumatology,
University Hospital
Nijmegen,
The Netherlands
D M F M van der Heijde
P L C M van Riel
L B A van de Putte

Department of Medical
Statistics,
Nijmegen University,
The Netherlands
M A van't Hof

Department of
Rheumatology,
University Hospital
Groningen,
The Netherlands
M A van Leeuwen
M H van Rijswijk

Correspondence to:
Dr D M F M van der Heijde,
Department of Rheumatology,
University Hospital Nijmegen,
Geert Grooteplein Zuid 8,
6525 GA Nijmegen,
The Netherlands.

Accepted for publication
18 March 1991

study of RA. Consecutive patients who met the following criteria were included in the study: classical or definite RA according to the American Rheumatism Association criteria, disease duration shorter than one year at entry to the study, and not previously treated with slow acting antirheumatic drugs (SAARDs). The project is a cooperation between the departments of rheumatology of the University Hospital Nijmegen and the University Hospital Groningen in The Netherlands. At time of analysis 133 patients participated in the study in Nijmegen (clinic I). The follow up ranged from 8 to 57 months (mean 28; number of check ups 2937). In Groningen (clinic II), 100 patients participated, with a follow up from 8 to 58 months (mean 31; number of check ups 3074).

ASSESSMENTS

In the two clinics three specially trained research nurses assessed all the patients in the outpatient department. The following assessments were carried out every four weeks: number of tender joints and number of swollen joints. All the following joints were examined bilaterally for the presence or absence of swelling and tenderness: temporomandibular, sternoclavicular, acromioclavicular, shoulder, elbow, wrist, each metacarpophalangeal, each proximal interphalangeal and interphalangeal of the thumb, hip, knee, ankle, subtalar, midtarsal, metatarsophalangeal, and interphalangeal of the big toe. The Ritchie articular index,¹⁸ time of morning stiffness (minutes), pain (on a visual analogue scale of 10 cm, 0=no pain, 10=worst pain possible), general health (visual analogue scale of 10 cm, 0=best possible, 10=worst possible), grip strength with a vigorimeter (kPa), erythrocyte sedimentation rate (ESR) according to Westergren (millimetres in one hour), haemoglobin (mmol/l), and C reactive protein (mg/l). The normal range for haemoglobin in our clinics is 1 mmol/l lower for women than for men. Therefore we decreased the haemoglobin value of the men by 1 mmol/l, ensuring comparability regardless of sex.

We calculated the Mallya index of disease activity for each patient at each visit.⁸ This index consists of morning stiffness, pain (visual analogue scale), grip strength, articular index, haemoglobin (uncorrected for sex) and ESR. Each variable is divided into four classes and the mean score of the six variables gives the full score (range one to four). The Riel index (a modified Mallya index) is made up of the variables morning stiffness, number of tender joints, haemoglobin (corrected for sex) and ESR and is calculated in a analogous manner to the Mallya index.¹⁰ The disease activity score is

composed of the Ritchie index, number of swollen joints, ESR, and general health.⁹ A total of 13 variables were analysed in this study.

Every six months the patients of clinic I completed a questionnaire on physical disability comparable with the health assessment questionnaire.⁵⁻¹⁹ Plain anterior radiographs of the hands and feet were also made every six months. The number of erosions and the joint space narrowing in the hands were scored by a modified version of Sharp's method.²⁰ The foot radiographs were scored in a similar way. These scoring methods have been extensively described before.²¹ Summation of the erosions and joint space narrowing gives the total score. The scores of all 77 patients of clinic I who were followed for at least two years were used in the analysis of radiographic features.

VALIDATION

As there is no standard available for the description of the quality of the variables considered for disease activity, a more explorative validation procedure was chosen. The following aspects of validation were considered. Correlational validity: does the measure correlate with other measures which are supposed to measure disease activity? For this analysis mutual correlations of the variables under investigation have been calculated. Criterion validity: does the assessment fit with the theory about the disease? Does the method measure the true clinical status? Therefore, correlations with physical disability have been determined and the clinical judgment of rheumatologists on disease activity has also been analysed. Construct validity: does the process variable (disease activity) lead to the ultimate result, the outcome? Correlations between the disease activity variables and radiographically determined damage have been calculated.

STATISTICAL PROCEDURES

Initially, all variables were assessed on their suitability for standard statistical analysis. If necessary, a transformation to 'normality' was performed. The transformed values were used in further analysis. The study of validity was mainly based on correlations.

Results

Table 1 presents the initial characteristics of the patients; table 2 summarises the variables used in the analysis. Owing to a high skewness, some of the variables are transformed to approximate a normal distribution. The skewness for clinic II is comparable with that of clinic I, and therefore the same transformations were performed.

CORRELATIONAL VALIDITY

The mutual correlations of the different disease activity variables give the correlational validity. The first calculations are performed on the data from each patient: the correlations between the variables within a patient and thereafter the

Table 1 Characteristics of the patients at the start of the follow up study

	Clinic I (n=133)	Clinic II (n=100)
No (%) women	88 (66)	65 (65)
Mean (SD) age (years)	54.3 (14.5)	47.7 (15.6)
Mean (SD) disease duration (months)	6.1 (4.0)	6.9 (4.6)
IgM rheumatoid factor >5 IU (No (%))	103 (80)	88 (88)
Physical disability	0.64 (0.57)	
Erosion of hands and feet (number)	2.9 (4.7)	
Total score, hands and feet	5.3 (4.7)	

Table 2 Mean, standard deviation (SD), minimum, maximum, and skewness before and after transformation of the variables of clinic I patients (n=133) and clinic II patients (n=100)

Variable	Clinic I (2937 check ups)					Clinic II (3074 check ups)				
	Mean	SD	Minimum	Maximum	Skewness before transformation	Skewness after transformation	Mean	SD	Minimum	Maximum
Disease activity score	3.35	1.23	0.32	7.26		—	3.28	1.28	0.35	8.35
Mallya index	2.23	0.54	1.00	3.83		—	2.18	0.58	1.00	4.00
Riel index	1.97	0.62	1.00	4.00		—	1.95	0.64	1.00	4.00
General health	30.7	22.6	0	100	0.53	-0.38				
Pain*	31.6	23.0	0	100	0.59	-0.31	28.5	24.9	0	100
Tender joints*	10.2	8.9	0	44	0.78	-0.11	12.9	12.3	0	48
Swollen joints	12.1	7.3	0	37	0.51	—	7.5	7.1	0	40
Ritchie index*	8.3	7.5	0	46	1.1	-0.06	11.3	10.9	0	52
Morning stiffness†	40	67	0	360	2.78	0.16	36	69	0	360
ESR‡	29	24	1	140	1.42	-0.43	31	26	1	140
C reactive protein†	28	39	0.1	260	2.46	-0.61	24	36	0.1	398
Haemoglobin	7.9	1.0	4.7	10.8	-0.41	—	8.0	1.0	4.6	11.0
Grip strength*	40	22	0	138	0.90	0.07	47	26	0	184

*Root transformation.

†Double root transformation.

‡Log transformation; ESR=erythrocyte sedimentation rate.

mean over all patients are calculated, leading to a mean correlation matrix. The correlation matrix is the basis on which the mean correlations with the other 12 variables are then calculated. Table 3 shows that the results of the correlational validity of the individual patient data are fully comparable in the two clinics. The mean correlations of the three composite indices are comparable and substantially higher than the single variables. The variable haemoglobin has an especially low correlational validity (0.29). The calculations are also performed on the raw data. Firstly, the correlation of each variable with every other variable is determined. Thereafter, for every variable we calculated the mean (minimum, maximum) correlation with the other variables. Table 3 shows that the overall correlations are similar to those on the individual patient data.

CRITERION VALIDITY

Criterion validity implies that a good disease activity variable measures the true clinical status of the patient. Two theoretical aspects are distinguished.

The first aspect of the criterion validity of all variables is determined by calculating the Pearson correlations between the variables and the physical disability on the individual patient data. This shows how the disease activity variables reveal what is the most important

Table 4 Criterion validity I: the mean and median individual Pearson correlation of each variable with the physical disability (clinic I, n=133)

Variable	Mean	Median
Disease activity score	0.44	0.70
Mallya index	0.43	0.60
Riel index	0.33	0.50
General health	0.36	0.49
Pain	0.39	0.58
Tender joints	0.37	0.58
Swollen joints	0.36	0.57
Ritchie index	0.42	0.68
Morning stiffness	0.29	0.47
ESR*	0.34	0.45
C reactive protein	0.38	0.60
Haemoglobin	-0.16	-0.32
Grip strength	-0.35	-0.58

*ESR=erythrocyte sedimentation rate.

aspect for the patients. Table 4 gives the mean and the median correlations. The medians can be interpreted best. The disease activity score and the Ritchie index have the highest median correlations (0.70 and 0.68, respectively); haemoglobin shows the lowest correlation (0.32).

Clinical judgment by rheumatologists is another procedure to evaluate actual disease activity. We therefore used the decisions of rheumatologists in clinical practice to start or stop treatment with SAARDs. Decisions to begin treatment with SAARDs or to stop SAARDs because of a lack of efficacy were equated with times of high disease activity; withdrawal as a result of remission, or not starting or changing SAARDs for at least one year as times of low disease activity. The disease activity score is the formalisation of the decisions of rheumatologists on the start and withdrawal of treatment with SAARDs in clinical practice and was constructed for the group of rheumatologists in clinic I. Therefore, only the records of the patients of clinic II were used for this analysis. Patients were divided into two groups, one with high and one with low disease activity, according to the above explicit rules.⁹ The group with high disease activity consists of 129 records and the group with low disease activity 115 records.

For all variables the mean and standard deviations of the groups with high and low disease activity are determined. The larger the difference between the means of the two groups,

Table 3 Correlational validity: mean Pearson correlation of each variable with the other 12 variables on the raw and individual data of all patients in the two clinics

Variable	Mean individual correlations				Raw correlations	
	Clinic I (n=133)		Clinic II (n=100)		Clinic I (n=2937)	
	Mean	Range	Mean	Range	Mean	Range
Disease activity score	0.63	0.35-0.89	0.64	0.34-0.87	0.61	0.32-0.87
Mallya index	0.65	0.51-0.83	0.64	0.43-0.87	0.62	0.44-0.86
Riel index	0.61	0.39-0.83	0.63	0.51-0.87	0.59	0.40-0.86
General health	0.38	0.14-0.73			0.36	0.06-0.80
Pain	0.42	0.18-0.73	0.42	0.06-0.70	0.38	0.03-0.80
Tender joints	0.50	0.18-0.92	0.50	0.13-0.93	0.47	0.09-0.94
Swollen joints	0.43	0.24-0.81	0.41	0.21-0.70	0.38	0.18-0.72
Ritchie index	0.50	0.19-0.92	0.51	0.11-0.93	0.48	0.08-0.94
Morning stiffness	0.42	0.18-0.75	0.35	0.02-0.61	0.33	0.07-0.64
ESR*	0.47	0.28-0.70	0.38	0.10-0.74	0.36	0.09-0.70
C reactive protein	0.41	0.23-0.70	0.40	0.13-0.74	0.38	0.09-0.70
Haemoglobin	0.32	0.14-0.59	0.30	0.02-0.58	0.29	0.03-0.59
Grip strength	0.41	0.32-0.61	0.46	0.28-0.71	0.41	0.23-0.64

*ESR=erythrocyte sedimentation rate.

the larger the discriminating power of the variable. To compare the variables, the differences are standardised (i.e. the difference is divided by the pooled standard deviation). Table 5 shows that the disease activity score has the most discriminating power (SD score 1.66), followed by the Riel index (SD score 1.46) and the Mallya index (SD score 1.37). The swollen joints count is the only single variable with a considerable discriminating power (SD score 1.35).

Table 5 Criterion validity II: mean and standard deviation (SD) for high and low disease activity (DA) in clinic II patients (n=100) and standardised difference

Variable	Criterion validity II				Standardised difference*
	High DA		Low DA		
	Mean (n=129)	SD	Mean (n=115)	SD	
Disease activity score	4.27	1.11	2.50	1.01	1.66
Mallya index	2.62	0.58	1.89	0.48	1.37
Riel index	2.41	0.62	1.60	0.47	1.46
Pain	6.14	2.30	3.34	2.79	1.10
Tender joints	4.01	1.63	2.19	1.76	1.08
Swollen joints	12.60	7.25	4.16	4.93	1.35
Ritchie index	3.82	1.56	2.05	1.51	1.15
Morning stiffness	2.21	1.31	1.06	1.14	0.93
ESR†	3.52	0.88	2.72	0.93	0.89
C reactive protein	2.27	0.80	1.59	0.62	0.94
Haemoglobin	7.44	1.07	7.88	0.69	0.49
Grip strength	7.18	1.78	5.48	2.32	0.82

*Standardised difference: (mean of high DA group - mean of low DA group) / pooled SD of high and low DA group.

†ESR=erythrocyte sedimentation rate.

Table 6 Construct validity: correlations between increase in joint damage in periods of six months and two years and mean clinical and laboratory variables over the same period (clinic I, n=77)

Variable	Six monthly periods* (n=345)			Two year periods* (n=133)		
	E	N	T	E	N	T
	Disease activity score	0.33	0.28	0.33	0.31	0.26
Mallya index	0.30	0.30	0.34	0.25	0.30	0.31
Riel index	0.30	0.25	0.32	0.22	0.21	0.24
General health	0.07	0.17	0.13	0.12	0.23	0.20
Pain	0.12	0.19	0.17	0.18	0.26	0.26
Tender joints	0.18	0.13	0.16	0.15	0.06	0.12
Swollen joints	0.42	0.27	0.39	0.54	0.39	0.48
Ritchie index	0.17	0.14	0.17	0.13	0.07	0.11
Morning stiffness	0.15	0.07	0.14	0.10	0.03	0.10
ESR†	0.29	0.38	0.39	0.19	0.36	0.29
C reactive protein	0.35	0.43	0.46	0.40	0.52	0.50
Haemoglobin	-0.24	-0.19	-0.25	-0.10	-0.14	-0.14
Grip strength	-0.29	-0.28	-0.32	-0.32	-0.39	-0.38

*E=erosions, N=joint space narrowing, T=total joint score.

†ESR=erythrocyte sedimentation rate.

Table 7 Summary of validation criteria for the 13 variables

	Correlational validity	Criterion validity I	Criterion validity II	Construct validity
Disease activity score	++	++	++	+
Mallya index	++	+	++	+
Riel index	++	±	++	+
General health	±	±	±	-
Pain	±	+	+	-
Tender joints	+	+	+	-
Swollen joints	±	+	++	++
Ritchie index	+	++	+	-
Morning stiffness	±	±	±	-
ESR*	+	±	±	+
C reactive protein	±	+	±	++
Haemoglobin	-	-	-	±
Grip strength	±	+	±	+

+ = very good; ++ = good; ± = moderate; - = poor.

*ESR=erythrocyte sedimentation rate.

CONSTRUCT VALIDITY

Construct validity evaluates whether the variable which measures the process corresponds with the ultimate result of the process: the outcome. Longstanding disease activity leads to (largely) irreversible joint damage, visible on radiographs as erosion and joint space narrowing. Thus the damage caused by the disease process and visible on radiographs is cumulative. For this analysis only those patients of clinic I who have been followed for at least two years have been included (n=77). The mean of each variable is calculated over adjacent periods of six months. This mean is correlated with the increase in joint damage in the same period. In total, 345 of those periods were available for analysis. The mean of the variables and the increase of the joint damage over a period of two years were also calculated. The total of the two year periods in the analysis is 133. Table 6 summarises the results of the Spearman rank correlations of the mean of the variables with the increase in number of erosions, joint space narrowing and total score. The results of the six monthly periods and the two year periods are comparable. The C reactive protein, swollen joints, ESR, disease activity score, Mallya index, Riel index, and grip strength have the highest correlations. The other variables have substantially lower correlations.

Discussion

Surprisingly, no systematic studies on the comparison of the validity of 'process' measurements have been made, although these variables are widely used in clinical trials and in clinical practice. In this study we were able to evaluate various aspects of the validity of 13 variables frequently used to measure disease activity. Not only single variables were included, but also indexes such as the Mallya, Riel, and disease activity score. Owing to the lack of a standard for disease activity we used an explorative method to evaluate validity, as is used widely in the social sciences. Physical disability and radiographically detected damage in the hands and feet are mostly used as 'outcome' measurements.³⁻⁵ Variables which measure the 'process' (disease activity) have to correspond with the 'outcome'.¹ Therefore we used the physical disability and the radiographically detected damage in the hands and feet to validate the disease activity variables.

Table 7 summarises all types of validity and gives a review of the relative quality of the variables. The limits of '+' are determined arbitrarily. However, this was carried out to make an overall comparison of all aspects of the validity of all variables more easily. This table provides a practical guide for the choice of a process variable. The disease activity indices (especially the disease activity score and the Mallya index) score best overall. The validity of various single variables was very poor (especially haemoglobin, morning stiffness, general health, pain, and grip strength). The most valid single variable is the number of swollen joints. The relation between swollen joints and radiographic progression confirms

earlier reports.^{22 23} The disease activity score is composed of the Ritchie index, the number of swollen joints, the ESR, and general health.⁹ The Mallya index consists of morning stiffness, pain, grip strength, articular index, haemoglobin, and ESR.⁸ Thus the combination of a few variables, which have little value as a single measurement, into an index greatly enhances the validity.

The disease activity score is based on judgments in clinical practice. This disease activity score turned out to be one of the most valid variables to measure disease activity compared with widely used variables. Therefore it might be concluded that clinical judgment by rheumatologists correlates with the ultimate outcome, namely physical disability and radiographically detectable damage.

In clinical trials of SAARDs, many disease activity variables are used to evaluate the effect. This introduces the problem of multiple testing, which may lead to conflicting results in several variables. In addition, more patients are needed to diminish a false positive result (type I error). This study shows that many of the single variables have a (very) low validity in measuring disease activity and that the disease activity score and Mallya index are the most valid variables in measuring disease activity. Therefore these indices may be used as single endpoint measures in clinical trials, preventing conflicting results and multiple testing (and fewer patients are needed). Not only in trials, but also in patient management, the problem arises of the discordance of the courses of individual variables of disease activity. In this situation, the indices may also be helpful.

In conclusion, we recommend using a disease activity index in evaluating trials and in clinical practice. Most of the so called disease activity variables only partly measure disease activity. It is questionable whether these variables, such as haemoglobin and morning stiffness, should be used as *single* disease activity variables. However, in combination with others variables in a disease activity index, these variables may be useful.

This investigation was supported by a grant from the Program for the Stimulation of Health Research (SGO) and the Netherlands League against Rheumatism. We thank H Theunisse, D Lubberts, and W Lolkema, research nurses, for the collection of the clinical data and Dr N Houtman for referral of patients.

1 Fries J F. Toward an understanding of patient outcome measurement. *Arthritis Rheum* 1983; 26: 697-704.

- 2 Sharp J T, Bluhm G B, Brook A, *et al*. Reproducibility of multiple-observer scoring of radiologic abnormalities in the hands and wrists of patients with rheumatoid arthritis. *Arthritis Rheum* 1985; 28: 16-24.
- 3 Sharp J T. Radiologic assessment as an outcome measure in rheumatoid arthritis. *Arthritis Rheum* 1989; 32: 221-9.
- 4 Meenan R F, Gertman P M, Mason J H. The arthritis impact measurement scales. *Arthritis Rheum* 1982; 25: 1048-53.
- 5 Fries J F, Spitz P, Kraines R G, Holman H R. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980; 23: 137-45.
- 6 Bombardier C, Tugwell P, Sinclair A, Dok C, Anderson G, Buchanan W. Preference for endpoint measures in clinical trials: results of structured workshops. *J Rheumatol* 1982; 9: 798-801.
- 7 Scott D L, Spector T D, Pullar T, McConkey B. What should we hope to achieve when treating rheumatoid arthritis? *Ann Rheum Dis* 1989; 48: 256-61.
- 8 Mallya R K, Mace B E W. The assessment of disease activity in rheumatoid arthritis using a multivariate analysis. *Rheumatol Rehabil* 1981; 20: 14-7.
- 9 van der Heijde D M F M, van't Hof M A, van Riel P L C M, van Leeuwen M A, van Rijswijk M H, van de Putte L B A. Judging disease activity in clinical practice in rheumatoid arthritis. The first step in the development of a 'disease activity score'. *Ann Rheum Dis* 1990; 49: 916-20.
- 10 van Riel P L C M, van de Putte L B A, Gribnau F W J, Macrae K D. Comparison of auranofin and aurothioglucose in the treatment of rheumatoid arthritis: a single-blind study. *Clin Rheumatol* 1984; 3 (suppl 1): 51-6.
- 11 Tugwell P, Bombardier C. A methodologic framework for developing and selecting endpoints in clinical trials. *J Rheumatol* 1982; 9: 758-62.
- 12 Kirwan J R, Chapat de Saintonge D M, Joyce C R B, Currey H L F. Clinical judgment in rheumatoid arthritis I. Rheumatologists' opinions and the development of paper patients. *Ann Rheum Dis* 1983; 42: 644-7.
- 13 Kirwan J R, Chapat de Saintonge D M, Joyce C R B, Holmes J, Currey H L F. Inability of rheumatologists to describe their true policies for assessing rheumatoid arthritis. *Ann Rheum Dis* 1986; 45: 156-61.
- 14 March R E, Kirwan J R, Reeback J S, Holborow E J. IgM, IgG and IgA rheumatoid factors (antiglobulins) in early rheumatoid arthritis and their production of articular index over one year. *Scand J Rheumatol* 1987; 16: 407-11.
- 15 Thompson P W, Silman A J, Kirwan J R, Currey H L F. Articular indices of joint inflammation in rheumatoid arthritis. *Arthritis Rheum* 1987; 30: 618-23.
- 16 Bull B S, Levy W C, Westengard J C, *et al*. Ranking of laboratory tests by consensus analysis. *Lancet* 1986; ii: 377-80.
- 17 Bull B S, Farr M, Meyer P J, Westengard J C, Bacon P A, Stuart J. Efficacy of tests used to monitor rheumatoid arthritis. *Lancet* 1989; ii: 965-7.
- 18 Ritchie D M, Boyle J A, McInnes J M, *et al*. Clinical studies with an articular index for the assessments of joint tenderness in patients with rheumatoid arthritis. *Q J Med* 1968; 147: 393-406.
- 19 van der Heijde D M F M, van Riel P L C M, van de Putte L B A. Sensitivity of a Dutch health assessment questionnaire in a trial comparing hydroxychloroquine vs. sulphasalazine. *Scand J Rheumatol* 1990; 19: 407-12.
- 20 Sharp J T, Young D Y, Bluhm G B, *et al*. How many joints in the hands and wrists should be included in a score of radiologic abnormalities used to assess rheumatoid arthritis. *Arthritis Rheum* 1985; 28: 1326-35.
- 21 van der Heijde D M, van Riel P L, Nuver-Zwart H H, Gribnau F W, van de Putte L B. Effects of hydroxychloroquine and sulphasalazine on progression of joint damage in rheumatoid arthritis. *Lancet* 1989; i: 1036-8.
- 22 Sharp J T, Lidsky M D, Duffy J. Clinical responses during gold therapy for rheumatoid arthritis. Changes in synovitis, radiologically detectable erosive lesions, serum proteins, and serologic abnormalities. *Arthritis Rheum* 1982; 25: 540-9.
- 23 Ingeman-Nielsen M, Haslkov O, Hansen T M, Halberg P, Stage P, Lorenzen I. Clinical synovitis and radiological lesions in rheumatoid arthritis. A prospective study of 25 patients during treatment with remission-inducing drugs. *Scand J Rheumatol* 1983; 12: 237-40.