Annals of the Rheumatic Diseases 1990; **49**: 916–920

Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score

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

Abstract

An investigation of clinical and laboratory variables which might form the basis for judging disease activity in clinical practice was made by six rheumatologists in a prospective study of up to three years' duration of 113 patients with early rheumatoid arthritis. Decisions to start treatment with slow acting antirheumatic drugs were equated with moments of high disease activity. If treatment with slow acting antirheumatic drugs was not started or if the slow acting antirheumatic drug remained unchanged for at least one year or if treatment was stopped because of disease remission, this was equated with periods of low disease activity. Two groups, one with high and one with low disease activity according to the above criteria, were formed. Factor analysis was performed to enable easy handling of the large number of clinical and laboratory variables without loss of information; this resulted in five factors. Next, discriminant analysis was done to determine to what extent each factor contributed to discrimination between the two groups of differing disease activity. Finally, a multiple regression analysis was carried out to determine which laboratory and clinical variables underlie the factors of the discriminant function, resulting in a 'disease activity score'. This score consisted of the following variables: Ritchie index, swollen joints, erythrocyte sedimentation rate, and general health, in declining importance. The rheumatologists' decisions to prescribe slow acting antirheumatic drugs, or not, were mainly based on articular symptoms.

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

Medical Statistical Department, University Nijmegen, 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 Désirée M F M van der Heijde, Department of Rheumatology, University Hospital Nijmegen, Geert Grooteplein Zuid 8, 6525 GA Nijmegen, The Netherlands.

Accepted for publication 22 January 1990

In rheumatoid arthritis disease activity cannot be measured by one single variable. In clinical practice an opinion of disease activity is formed from a combination of information, such as laboratory and clinical variables, radiological assessments, and overall impression of the patient. This clinical judgment of disease activity varies considerably among different rheumatologists as has been shown by Kirwan.[1] In addition, there is a discrepancy between what doctors believe their clinical behaviour to be and the way in which they really act in practice.[2 3]

If this clinical judgment can be formalised to provide a quantifiable disease activity index it would provide an opportunity to study and influence this process. In addition, such an instrument could be used to compare the efficacy of treatments in clinical trials.

In a large prospective study the decisions of rheumatologists to start treatment with a slow acting antirheumatic drug or to stop such treatment because of disease remission were equated with high and low disease activity respectively. The clinical and laboratory variables that explain most of the variance of the rheumatologists' decisions were composed by various statistical methods into a 'disease activity score'.

The result of such implicit judgment may be the first step in composing a disease activity index. The next step is validation: Is the judgment of the rheumatologists in practice a good reflection of 'real' disease activity?

Patients and methods

PATIENTS

All patients met the following criteria: they had classical or definite rheumatoid arthritis according to American Rheumatism Association criteria, disease duration of less than one year, and had not previously been treated with slow acting antirheumatic drugs. All consecutive patients eligible for the prospective follow up from January 1985 were asked to participate. Eight patients were not included (refusal, fatal accompanying disease). At the time of analysis 113 patients participated in the study and the follow up ranged from two to 39 months (number of check ups 1816).

METHODS

Two specially trained research nurses assessed all the patients every four weeks in the rheumatology outpatient department of the University Hospital, Nijmegen. Furthermore, all patients were followed up by their rheumatologists, independently of the evaluations of the research nurses, on average four to six times a year. The rheumatologist made all the decisions to start or withdraw slow acting antirheumatic drugs independently of the clinical assessments of the research nurses. The rheumatologists were not informed of the fact that their decisions were part of the investigation. Six rheumatologists were working at the outpatient department. The sequence of the start of the various slow acting antirheumatic drugs follows a fixed schedule: first step hydroxychloroquine or sulphasalazine, second step intramuscular gold, thereafter D-penicillamine or azathioprine or methotrexate. Corticosteroids and non-steroidal anti-inflammatory-drugs are allowed as adjuvants at all stages. The rheumatologists' policy is to start treatment with slow acting anti-

rheumatic drugs if the disease is not adequately controlled after two months' treatment with non-steroidal anti-inflammatory drugs alone.

ASSESSMENTS

The following assessments were made every four weeks: number of tender joints, number of swollen joints, Ritchie articular index,[4] morning stiffness (minutes), fatigue (hours after rising), 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 (mmHg) and the body weight (kg), erythrocyte sedimentation rate (ESR) according to Westergren (mm in 1st hour), haemoglobin (g/l), leucocyte count (10^9/l), thrombocyte count (10^9/l), total protein, albumin, α_1 globulin, α_2 globulin, β globulin, γ globulin (all g/l), C reactive protein (g/l), IgM rheumatoid factor (normal <5 IU/l). Albumin and glucose were determined in the urine. Creatinine (μmol/l), alkaline phosphatase (U/l), antinuclear antibodies, IgA, IgM, and IgG (g/l) were measured every three months and serum iron and total iron binding capacity (μmol/l) every six months. Every six months plain anterior radiographs of the hands and feet were obtained, all patients completed a questionnaire on their physical and psychosocial wellbeing, and the production of tears was assessed with a Schirmer test (normal >10 mm). Serum and plasma were stored at −20°C at every visit. Some measurements were made for control purposes because of the slow acting antirheumatic drugs, others to follow the course of the rheumatoid arthritis or to search for complications or systemic manifestations. Variables, previously reported as possible disease activity markers, were included in the analysis: number of tender joints, number of swollen joints, Ritchie articular index, morning stiffness, fatigue, pain, general health, grip strength right and left, ESR, haemoglobin, thrombocyte count, total protein, albumin, α_1 globulin, α_2 globulin, β globulin, γ globulin, C reactive protein, and IgM rheumatoid factor.

STATISTICAL PROCEDURE

Initially, all selected variables were assessed for their suitability for multivariate statistical analysis. If necessary, transformation to a reasonably normal distribution was performed. To find some structure in the large number of variables, factor analysis was performed, resulting in a few factors to be used in the further analysis. The patients' records were then divided (according to explicit rules) into a group with high and one with low disease activity. This selection was the basis of a discriminant analysis to determine how the factors might best be combined to produce a score which reflects the disease activity most accurately: the disease activity score. The reproducibility of the factors was studied with a correlation matrix analysis between periods. Subsequently, a multiple regression analysis was performed to determine how the disease activity score might best be measured in practice.

Results

Table 1 shows some general characteristics of the patients. Table 2 summarises the variables used in the analysis. Because of a high skewness some variables were transformed to approximate a normal distribution. These transformed values were used in the analysis. As the variable 'total protein' depends fully on the other components of the protein analysis, and the variable 'fatigue' showed no change over the time, these two variables were excluded from the analysis.

FACTOR ANALYSIS

Initially, a factor analysis on the complete data was performed. Three factors with Eigen values higher than 1 and a cumulative percentage of explained variance of 59% were analysed further. Table 3 shows the factors and corresponding variables. Factor 1 consists of inflammatory variables in the blood, factor 2 includes the semiobjective measures of joint examination, and factor 3 subjective variables and β globulin. To obtain a result which could be interpreted more easily a factor analysis on the individual data was performed. For this purpose Pearson correlations between all longitudinal series of observations in each individual patient were calculated. A correlation matrix of the median correlations over all patients was composed as the basis of the factor analysis. Five factors have an Eigen value higher than 1 with a cumulative percentage of explained variance of 66%. Table 4 summarises the five factors with the variables with the highest loadings on these factors. In conclusion, factor analysis of individual data gives the most easily interpreted results and the greatest explained variance. The factors can be labelled as variables of inflammation in the blood (factor 1), variables of joint examination (factor 2), protein analysis (factor 3), subjective complaints (factor 4), and grip strength (factor 5). Two factors (1 and 3) are objective measurements, factor 2 is a semiobjective assessment, factor 4 reflects the subjective complaints of the patient, and factor 5 combines disease activity and structural (irreversible) damage of the hand.

CALCULATION OF THE FACTOR VALUES

The standard deviation score (Z score=(observed value−mean)/standard deviation) was calculated for all variables at each visit. The mean of the Z scores of the variables corresponding with one factor gave the factor value for that visit. The factor values are the basis of the subsequent calculation. (For the sake of simplicity we used

Table 1 Characteristics of the 113 patients at the start of the study

Male/female (n)	43/70
Mean (SD) age (years)	54·2 (14·5)
Range	16–81
Median	55·3
Mean (SD) disease duration (months)	5·5 (3·6)
Range	0–21
Median	4·9
IgM RF* >5 IU/ml (n)	89

*RF=rheumatoid factor.

Table 2 Minimum, maximum, and skewness of the variables

Variable	Mean	SD	Range	Skewness	Skewness after transformation
ESR* (mm/1st h)	31	25	(1–140)	1·367	−0·460
Albumin (g/l)	44·4	5·0	(26·5–64·6)	−0·474	−0·474
α_1 Globulin (g/l)	2·4	0·8	(0·5–8·1)	0·980	−0·249
α_2 Globulin (g/l)	7·1	1·7	(1·7–17·6)	0·734	−0·182
β Globulin (g/l)	7·3	1·4	(4·0–14·7)	0·515	0·175
γ Globulin (g/l)	12·7	4·1	(4·0–37·2)	1·146	−0·177
Haemoglobin (g/l)	127	16	(76–174)	−0·214	−0·214
Thrombocytes ($\times 10^9$/l)	313	100	(103–845)	1·126	0·102
CRP* (g/l)	30	40	(3–260)	2·455	0·231
IgM RF*	215	412	(<5–3200)	3·078	0·238
Morning stiffness (min)	52	74	(0–360)	2·394	0·284
Pain (100 mm scale)	33	22	(0–100)	0·433	0·433
General health (100 mm scale)	31	22	(0–100)	0·423	0·423
Grip strength right (mmHg)	40	22	(0–158)	1·031	−0·037
Grip strength left (mmHg)	38	22	(0–131)	0·830	−0·228
Ritchie score	9	8	(0–46)	0·997	−0·152
Tender joints (n)	11	9	(0–40)	0·618	−0·252
Swollen joints (n)	14	7	(0–36)	0·381	0·381

*ESR=erythrocyte sedimentation rate; CRP=C reactive protein; RF=rheumatoid factor.

Table 3 Three factor model of all data

Factor 1	Factor 2	Factor 3
ESR*	Ritchie score	Pain
Thrombocytes	Tender joints	General health
Haemoglobin	Swollen joints	β Globulin
Albumin	Morning stiffness	
α_1 Globulin	Grip strength right	
α_2 Globulin	Grip strength left	
γ Globulin		
CRP*		
IgM RF*		

*ESR=erythrocyte sedimentation rate; CRP=C reactive protein; RF=rheumatoid factor.

sum scores rather than the actual factor scores, resulting in slightly different values: the factor values.)

RELIABILITY OF THE FACTOR VALUES

Cronbach's alpha,[5] computed to determine the reliability of the factor values, ranged from 0·73 (factor 2) to 0·92 (factor 5), which may be considered fairly reliable.

COURSE OF THE FACTOR VALUES

To determine the progression of the factor values with time the mean autocorrelations of the five factors were determined. Factor 1 had an autocorrelation of 0·61, which means that the factor proceeded smoothly with time. On the other hand, factor 3 had an autocorrelation of 0·21 and accordingly, a capricious course. The autocorrelations of the other three factors varied from 0·42 to 0·48.

REPRODUCIBILITY OF THE FACTOR VALUES IN THE LONGITUDINAL STUDY

The measurement quality of the factors is very important in longitudinal studies. Fortunately, owing to the repetition of measurements, it was possible to estimate this quality in the study using the interperiod correlation matrix. Thus the intercorrelation of five periods (months) was plotted against the intervening time intervals (one to four months). When the short time interval was taken into account the correlations were linearly related to time, which may be represented by a well fitting regression line (correlation *v* time interval). Extrapolation of this line to a time interval of zero gave the direct measurement-remeasurement correlation, which may be interpreted as a quality measure of the factor. Table 5 presents the results of the estimations.

DEFINITION OF DISEASE ACTIVITY

The overall judgment of the rheumatologist was the starting point in assessing disease activity. Indeed, the decision to start or terminate slow acting antirheumatic drugs was used as the criterion. The decision to start slow acting antirheumatic drugs was taken by the rheumatologists independently of the clinical assessments of the research nurses.

The definition of high disease activity was (*a*) start of a slow acting antirheumatic drug; (*b*) termination of treatment with slow acting antirheumatic drugs because of lack of effect. The definition of low disease activity was (*a*) termination of treatment with slow acting antirheumatic drugs because of remission of the rheumatoid arthritis; (*b*) not changing a slow acting antirheumatic drug for at least one year; (*c*) not starting treatment with a slow acting antirheumatic drug for at least one year. If a patient met the above criteria more than once only the observations made at an interval of at least five months were included. Ultimately, 78 patients met these conditions, with the numbers

Table 4 Five factor model of the individual data

Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
ESR*	Ritchie score	Albumin	Pain	Grip stength, right, left
Thrombocytes	Tender joints	α_1 Globulin	General health	
Haemoglobin	Swollen joints	α_2 Globulin	Morning stiffness	
CRP*		β Globulin		
IgM RF*		γ Globulin		

*ESR=erythrocyte sedimentation rate; CRP=C reactive protein; RF=rheumatoid factor.

Table 5 Measurement-remeasurement correlations of the factors estimated from the interperiod correlation matrix

Factor	r_0
1	0·94
2	0·87
3	0·70
4	0·74
5	0·94

of records ranging from one to four. In all, 177 assessments were analysed: 138 observations in the group with high disease activity and 39 in those with low disease activity.

DISCRIMINANT ANALYSIS

The factor values of the five factors of the assessments were used in the discriminant analysis, restricted to the group with high disease activity and the group with low disease activity. Moreover, an analysis without factor 5 was made because grip strength reflects not only disease activity but also (irreversible) destruction. A discriminant analysis on three factors, leaving out factors 3 and 5, was also performed, as factor 3 (protein analysis) has the lowest reproducibility (table 5). Table 6 shows the relative importance of the variables used in three, four, and five factor groups. When factor 3 or 5 or both, was omitted this did not affect the canonical correlation or the discriminating power. In the subsequent analysis the discriminant function based on three and four factors was used. The joints scores (factor 2) contributed most to the discriminant score. The other factors were equally important but at a lower level.

MULTIPLE REGRESSION ANALYSIS

The discriminant function contains factor values which in turn are composed of several clinical and laboratory variables. A multiple regression analysis was used to determine which variables contribute most to the discriminant function, to obtain a disease activity score which could be used easily in practice.

A stepwise forward multiple regression analysis was performed on all 1816 records, with the discriminant function on three factors as the dependent variable and ESR, haemoglobin, thrombocytes, morning stiffness, number of tender joints, number of swollen joints, Ritchie score, pain, general health, C reactive protein, and IgM rheumatoid factor as independent variables. As the first step the Ritchie score was included (multiple R=0·852), thereafter the number of swollen joints (multiple R=0·951), ESR (multiple R=0·977), general health (multiple R=0·984), thrombocytes (multiple R=0·988), and number of tender joints (multiple R=0·993). On the basis of these results we decided to compose a disease activity score using the variables Ritchie score, number of swollen joints, ESR, and general health. In practice the variable of general health is not always determined. Therefore the regression comparison was also calculated with the remaining three variables. Both the Ritchie score and the ESR were transformed. Table 7 shows the disease activity scores of four and three variables after conversion. From these data the relative contribution to the equation of each variable is not clear. Therefore we determined the partial correlations of the variables when used as the first step and when used as the final step in the regression analysis. Table 8 shows these correlations.

Table 8 Relative contribution of the variables to the equation of disease activity, expressed as the partial correlations of the initial and final steps in the regression analysis

Variable	*Initial partial correlation*	*Final partial correlation*
Ritchie score	0·80	0·85
Swollen joints	0·69	0·74
ESR	0·49	0·63
General health	0·50	0·39

The constant of the function of three variables (D3) was chosen so that the mean of this score equalled the score based on four variables without a constant to ensure the possibility of exchange.

The mean disease activity scores of all records were 3·57 (SD 1·18, range 0·51–7·26, four variables) and 3·57 (SD 1·12, range 0·59–7·05, three variables).

REPRODUCIBILITY OF THE DISEASE ACTIVITY SCORE

Reproducibility of the disease activity score was determined by an interperiod correlation matrix of five periods (months), as described for the reproducibility of the factor values. The measurement-remeasurement correlation was 0·89 for the disease activity scores with both three and four variables.

Table 6 Pooled correlation of discriminant functions of three, four, and five factors

	5 Factors	*4 Factors*	*3 Factors*
Factor 1	0·51	0·51	0·53
Factor 2	0·86	0·86	0·91
Factor 3	0·51	0·51	
Factor 4	0·47	0·47	0·51
Factor 5	−0·54		
Canonical correlation	0·658	0·662	0·630
Correctly classified (%)	83·2	83·9	83·6

Table 7 Computation of the disease activity scores. The units used in the formulas are given in table 2

Disease activity score (four variables)=
D4=0·53938×sq rt (Ritchie score)+0·06465×(number of swollen joints)+
0·330×ln(ESR)+0·00722×(general health)

Disease activity score (three variables)=
D3=0·53938×sq rt (Ritchie score)+0·06465×(number of swollen joints)+
0·330×ln(ESR)+0·224

Discussion

Clinical judgment of disease activity in rheumatoid arthritis is a complex process of combining clinical and laboratory variables as well as radiological assessments and overall impression. This clinical judgment is difficult to formalise. Kirwan described the great difference between what doctors believe about their clinical behaviour and their actual behaviour in pratice.[2 3] Most rheumatologists are unable to describe their policies in judging disease activity. Kirwan's studies were executed on 'paper patients' on single occasions. Though these 'paper patients' correlated well with the real situation, a survey of judging disease activity on such patients does not necessarily reflect the process of actual decision making in practice.[1 6] This study describes judging disease activity in

actual practice. We opted for a real decision point in patient management: the moment when the rheumatologist considered the rheumatoid arthritis so active that the patient had to start treatment with or change slow acting drugs was marked as high disease activity. Conversely, patients who were not treated with slow acting antirheumatic drugs or continued to take the same slow acting antirheumatic drug during at least one year were placed in the group of low disease activity. In addition to disease activity other factors may lead to the start or withdrawal of a slow acting antirheumatic drug—for example, the refusal of the patient. These factors might have interfered with the appropriate classification of patients. This probably played a part in only a few patients, so we did not correct for possible misclassification as this might have introduced a subjective interpretation.

We were able to describe the real process of decision making in practice because the rheumatologists were unaware that their decisions were part of an investigation. It was possible to develop the disease activity score because of the prospective monthly follow up of a large number of patients, during periods of up to more than three years. At each visit a large number of variables was collected by the same observer.

A disease activity score based on the variables collected by the research nurses was composed using different statistical procedures. This disease activity score shows which variables best explain the decisions of the rheumatologists. The disease activity score includes four variables: Ritchie score (semiobjective, clinical variable), number of swollen joints (objective, clinical variable), ESR (laboratory variable), and general health (subjective variable). In practice, the variable, general health, is not always determined. Therefore we calculated a constant to be added to the equation of the disease activity score with three variables so that the means of the scores based on three and four variables were equal, making the two disease activity scores interchangeable.

Although not chosen for this reason, the variables are part of the whole spectrum: clinical, laboratory, objective, and subjective. The Ritchie index gives most weight to the equation, indicating that a high Ritchie score plays a major part in the decision about starting treatment with a slow acting antirheumatic drug. Secondly, swollen joints are important, followed by ESR and general health. The rheumatologists of our department were guided more by joint symptoms than by laboratory abnormalities (table 8). This is in accordance with the results of structured workshops on preferences for endpoint measures in clinical trials in Canada, where joint count, pain, and overall scores accounted for 60% of the total score for all measures.[7] These workshops used (among other things) patients' profiles before and after a non-steroidal anti-inflammatory drug trial. Recently, Anderson described the results of an analysis of pooled raw data from three placebo controlled trials of slow acting antirheumatic drugs in rheumatoid arthritis.[8] The joint tenderness count, ESR, joint swelling score, doctor's assessment of disease activity, and joint tenderness score showed, in descending order, the highest rankings of adjusted t statistic for all patients who took an active drug compared with all who took placebo. These results greatly resemble our disease activity score. The next step in the further development of the disease activity score is validation: Do the rheumatologists' clinical judgments really reflect disease activity? One way of answering this is to use the disease activity score in a clinical trial comparing two slow acting antirheumatic drugs and relate the results to individual clinical and laboratory variables. Correlation with radiological progression or functional capacity is another possible method of validation. If the disease activity score proves to be valid it may serve as a single outcome variable in a clinical trial comparing slow acting antirheumatic drugs. This would avoid conflicting results with various outcome variables. Moreover, smaller numbers of patients are needed since no correction has to be made for multiple statistical testing. In practice the validated disease activity score might be used to assess the disease activity of an individual patient and determine objectively when to start using a slow acting antirheumatic drug. In addition, the efficacy of the slow acting antirheumatic drug may be determined: an improvement of the score by 1·08 (=2× standard error) or more is a statistically significant improvement. This provides clear information, helpful in the management of individual patients.

The value of the disease activity score is currently under investigation and will become apparent in the future.

This investigation was supported by a grant from the Programme for the Stimulation of Health Research (SGO).

1 Kirwan J R, Chaput de Saintonge D M, Joyce C R B, Currey H L F. Clinical judgment in rheumatoid arthritis. I. Rheumatologists' opinion and the development of 'paper patients'. *Ann Rheum Dis* 1983; **42**: 644–7.
2 Kirwan J R, Chaput de Saintonge D M, Joyce C R B, Currey H L F. Clinical jugment in rheumatoid arthritis. II. Judging 'current disease activity' in clinical practice. *Ann Rheum Dis* 1983; **42**: 648–51.
3 Kirwan J R, Chaput 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.
4 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.
5 Cronbach L J. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; **16**: 297–334.
6 Kirwan J R, Bellamy N, Condon H, Watson Buchanan W, Barnes C G. Judging 'current disease activity' in rheumatoid arthritis. An international comparison. *J Rheumatol* 1983; **10**: 901–5.
7 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.
8 Anderson J J, Felson D T, Meenan R F, Williams H J. Which traditional measures should be used in rheumatoid arthritis clinical trials? *Arthritis Rheum* 1989; **32**: 1093–8.