Rheumatology 2005;44:789–795 doi:10.1093/rheumatology/keh595
Advance Access publication 9 March 2005

Measuring disease activity in ankylosing spondylitis: patient and physician have different perspectives

A. Spoorenberg[1,6], A. van Tubergen[1], R. Landewé[1], M. Dougados[2], S. van der Linden[1], H. Mielants[3], H. van de Tempel[4] and D. van der Heijde[1,5]

Objective. **There is no 'gold standard' to assess disease activity in patients with ankylosing spondylitis (AS). It is known that patients and physicians have different opinions about disease activity. The objective was therefore to investigate on which criteria patients with AS and physicians base their judgement on disease activity.**

Methods. **A cohort of 203 AS out-patients fulfilling the modified New York criteria included in the ongoing long-term follow-up was analysed. The Assessment in Ankylosing Spondylitis (ASAS) International Working Group has established different domains relevant for outcome in AS. Each domain includes a number of instruments for making assessments, and all these instruments are included in the Outcome in Ankylosing Spondylitis International Study and were made every 6 months for 2 yr. Disease activity from the patient perspective as well as from the physician perspective was analysed using the patient's or the physician's global assessment of disease activity [visual analogue scale (VAS): 0 (best)–10 (worst)] by dichotomizing into 'high disease activity' (VAS $\geq$ 6.0) and 'low disease activity' (VAS $\leq$ 4.0). Data reduction by principal components analysis (PCA) was performed to distinguish factors capturing correlated instruments. Discriminant analysis with the factor loadings was performed to discriminate between a low and a high disease activity state from both the patient's and the physician's perspective. Multiple regression analysis on the discriminant scores was performed to prioritize the instruments.**

Results. **PCA revealed four factors: spinal mobility, physician assessments, patient assessments and laboratory assessments (Cronbach's alpha 0.52–0.80; explained variance 61%). Discriminant function analysis showed that the factor 'patient assessments' was most important (pooled correlation 0.85) in discriminating between a low and a high disease activity state as defined by the patient. The other three factors contributed marginally (pooled correlation <0.30). In contrast, the factors 'physician's assessments' (pooled correlation 0.62), 'spinal mobility' (pooled correlation 0.52) and 'laboratory assessments' (pooled correlation 0.48) contributed most to the physician's perspective. The factor 'patient assessments' did not contribute at all (pooled correlation 0.05). Multivariate analysis on the discriminant scores showed that the instruments 'pain spine', 'BASFI', 'pain joints' and 'BASDAI fatigue' explained more than 90% of variance in the case of the patient perspective. The instruments 'cervical rotation', 'swollen joint count', 'CRP' and 'intermalleolar distance' explained more than 90% of variance in case of physician perspective.**

Conclusion. **AS patients rate disease activity on the basis of complaints while physicians rate disease activity on the basis of instruments related to disease severity and inflammation.**

KEY WORDS: Ankylosing spondylitis, Disease activity.

In most rheumatological disorders, disease activity and outcome cannot be measured by a single variable. In clinical practice opinion about disease activity and outcome is based on different sources of information, such as patient complaints, clinical variables, laboratory variables (acute phase reactants) and imaging. All this information is compiled into an overall impression of disease activity.

Patients and physicians may think differently about how to define disease activity. In judging whether their disease is active or not, patients may rate their complaints higher than, for example, abnormal laboratory results or rapidly progressive damage on X-rays, whereas physicians will tend to give weight to the latter observations, irrespective of patient complaints.

In ankylosing spondylitis (AS) it is especially difficult to define disease activity because there is much variety in the clinical picture among different patients. Patients may only have axial involvement in all degrees of severity, but they may also have extraspinal manifestations, such as enthesitis and joint inflammation or inflammation of the gastrointestinal tract. The clinical diversity, both in severity and in localization, makes a high demand on the instruments that are supposed to measure disease activity and outcome in AS. Since there is no gold standard available for measuring disease activity and outcome, many different instruments have been developed to assess a variety of signs and symptoms in AS. Some of these instruments emphasize the patient's perspective of disease activity or outcome, whereas others

[1]University Hospital Maastricht, The Netherlands, [2]Hôpital Cochin, Paris, France, [3]University Hospital Gent, Belgium, [4]Maasland Hospital, Sittard, The Netherlands, [5]Limburg University Center, Diepenbeek, Belgium, [6]Medical Center Leeuwarden, The Netherlands, [7]University Hospital Leuven, Belgium.

Received 12 November 2003; revised version accepted 1 February 2005.

Correspondence to: D. van der Heijde, Department of Rheumatology, University Hospital Maastricht, PO Box 5800, 6202 AZ Maastricht, The Netherlands. E-mail: dhe@sint.azm.nl

represent the physician's point of view. Acute phase reactants such as C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR) are considered more objective. In AS their value in determining whether the disease is active or not is rather limited. Elevated CRP levels and ESR are frequently absent in AS and do not correlate well with clinical activity and radiological progression [1, 2].

In 1999, a core set to assess outcome in AS was selected by the 'Assessment in Ankylosing Spondylitis' (ASAS) International Working Group. This AS outcome core set consists of different domains with specific instruments for each domain [3]. It includes instruments that are supposed to primarily measure the patient's as well as the physician's perspective regarding disease activity and outcome.

The objective of this study was to explore differences between the perspective of the patient and the perspective of the physician with respect to disease activity. We hypothesized that patients with AS will rate the activity of their disease primarily on the basis of their complaints, whereas the treating physicians will use other parameters on which to base their judgement.

Patients and methods

For this study data derived from the OASIS cohort (Outcome in Ankylosing Spondylitis International Study) were used. The OASIS project is an international longitudinal observational multicentre study performed at the rheumatology out-patient department of the University Hospital in Maastricht (The Netherlands), the Maasland Hospital in Sittard (The Netherlands), Hôpital Cochin in Paris (France) and the University Hospital in Gent (Belgium). In total, 203 AS out-patients who satisfied the modified New York criteria [4] were included in this study. Of these, 73% were male, a distribution usually found in AS populations. The mean age at baseline was 43 yr (S.D. 13 yr). The mean duration of disease since diagnosis was 11 yr (S.D. 8 yr). Twenty-seven per cent of the patients had a history of peripheral arthritis diagnosed by the treating rheumatologist. For each country the same trained person (two rheumatologists, one for The Netherlands and one for Belgium, and one research nurse for France) assessed all patients every 6 months according to a prespecified protocol. All physician assessments of disease activity were done in France by one and the same rheumatologist. All patients were followed and treated by their rheumatologists according to common clinical practice, independent of the evaluations of the researchers.

Ethical approval was obtained for this study as well as written informed consent from all patients.

Assessments

The following assessments were made every 6 months for the first 2 yr of the OASIS study: physician spinal pain assessment (0 = no pain on firm palpation, percussion and on extreme motion of complete spine, no spasm; 1 = slight pain on firm palpation, percussion or motion of complete spine and no more than slight limitation of motion; 2 = moderate pain on moderate palpation, percussion or motion of complete spine and no more than slight limitation of motion; 3 = moderate to severe pain on light palpation, percussion or slight motion of complete spine and moderate to severe limitation of motion; 4 = extreme pain with inability to withstand even light palpation or percussion and essentially no mobility of spine) (FDA guidelines). The best of two tries was documented in case of: chest expansion (cm) [5], finger to floor (cm) [6], occiput to wall (cm) (FDA guidelines), tragus to wall (cm) [7], modified Schober test (cm) [8], cervical rotation (degrees) [9], lateral spinal flexion (cm) [10] and intermalleolar distance (cm) [11]. Tragus to wall, modified Schober test, cervical rotation, lateral spinal flexion and intermalleolar distance were combined to compute the Bath Ankylosing Spondylitis Metrology Index (BASMI) [12]. Other physician-derived assessments were: the articular index according to Dougados (range 0–30) [13], enthesis index according to Mander (0–90) [14], Maastricht Ankylosing Spondylitis Enthesis Score (MASES) (range 0–13) [15], physician assessment of disease activity on a visual analogue scale of 10 cm (VAS, 0 = not active, 10 = extremely active), physician assessment of the number of tender joints (range 0–44) and swollen joints (0–40) [8].

Patient-derived assessments were: duration of morning stiffness of the spine (min), duration of morning stiffness of the peripheral joints (min), pain of the spine (VAS, 0 = no pain, 10 = unbearable pain), pain of the peripheral (VAS, 0 = no pain, 10 = unbearable pain), patient assessment of disease activity (VAS, 0 = not active, 10 = extremely active), fatigue (VAS, 0 = not at all, 10 = extremely), Bath Ankylosing Spondylitis patient Global Score (BASG) (VAS, 0 = no influence of AS on global well-being during the past week, 10 = global well-being completely influenced by AS during the past week) [16], patient assessment of night pain (range 0 = no pain, 4 = extremely painful during the whole night), Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) (range 0–10) [17], Bath Ankylosing Spondylitis Functional Index (BASFI) (range 0–10) [18] and Dougados Functional Index (D-FI) (range 0–40) [13]. The fatigue question of the BASDAI was used as a single variable for further analysis of the domain fatigue. The laboratory assessments were: erythrocyte sedimentation rate (ESR, mm/h) and C-reactive protein (CRP, mg/l).

All measurements were obtained to follow the course of ankylosing spondylitis. Both the investigators and the patients were aware of the results of the physical assessments, and the assessments completed by the patients before they made their judgement of disease activity. Neither patients nor investigators knew the results of the blood tests or the previous examinations.

Analysis

First, all selected variables were assessed for their suitability for parametric statistical analysis. Variables with a skewness statistic ≥ 1 were logarithmically transformed in order to obtain a reasonably normal distribution, required in principal components analysis (PCA) (see below).

If mutual correlation of the selected instruments was high (Pearson correlation coefficient >0.6), it was assumed that these variables assess the same process (collinearity). To avoid collinearity, we chose one of the two variables with high Pearson correlation for further analyses (for specific selection rules see below).

Data reduction

When many different measures have been taken on the same patient, it is possible to determine if some of these tests are actually reflections of a smaller number of underlying constructs, or factors. Factor analysis explores the interrelationships among the variables, in order to discover these factors. To structure the large number of measures, and to identify underlying constructs (factors), PCA (SPSS 10.0.7 factor analysis, rotation-varimax) was performed here, resulting in a few factors to be used in further analyses. Factor extraction in PCA is based on the sizes of the eigenvalues. Eigenvalues represent variance accounted for by a factor.

According to the commonly used criteria, we selected only factors with an eigenvalue >1. We used varimax rotation as a technique to maximize the level of variance explained by each factor. The factor loadings were saved in order to use for further statistical analysis. The factor loadings express the level of

correlation between the measured variable and the underlying factor. Internal consistency of the resulting factors was investigated by calculating Cronbach's alpha. Cronbach's alpha, which is based on intervariable correlation, is a reliability statistic showing to what extent variables in the same factor measure the same underlying construct.

Discriminant function analysis and linear regression analysis

Using explicit rules the patients were divided into two groups of high and low disease activity: those with active disease from the patient perspective (1) and those with active disease from the physician perspective (2). Active disease from the patient and physician perspective was defined as a score of $\geq$6.0 (VAS) and low disease activity was defined as $\leq$4.0 (VAS) on the instruments 'patient's global assessment of disease activity' and 'physician's global assessment of disease activity' respectively [VAS range: 0 (not active) to 10 (extremely active)]. To investigate whether and how the constructs (factors) could discriminate between low and high disease activity from patient and physician point of view, discriminant function analyses using the factor loadings were performed for both perspectives. Pooled correlations were used to judge the relative contribution of each factor to the discriminant function. Correlations were also expressed as canonical correlations that are multiple correlations between the predictors (here: factors) and the resulting discriminant function, and provide information about the accuracy of classification on the basis of the predictors. To test the robustness of the observations, the same analyses were performed for different time points of the study.

Because a discriminant function with factors does not directly illustrate which instruments are most contributory, we performed linear regression analysis with stepwise forward selection of all the variables that were used in the PCA on the individual discriminant scores. Such an analysis, which in the end will select all variables by default, will give information on the contribution of every individual instrument by the order in which it appears. We decided to report those instruments that together explain more than 90% of variation as important.

Results

Selection of variables

The descriptive statistics of all assessments at baseline are presented in Table 1. Variables not normally distributed (skewness $\geq$1) were ln-transformed. A Pearson correlation matrix of all variables from the OASIS data set at baseline was constructed to trace collinearity ($\rho > 0.6$). We chose one of the two variables with high intervariable correlation for further analysis. BASFI was selected instead of D-FI because in contrast to BASFI, D-FI was also highly correlated ($\rho > 0.6$) with two other items; occiput-to-wall distance was selected instead of tragus-to-wall distance because it is more widely used in other AS studies; lateral spinal flexion was selected instead of the modified Schober test because logarithmic transformation of the latter did not result in a normal distribution; pain spine was selected instead of BASG-week (global well-being) because BASG-week was also highly correlated with two other items; physician assessment of painful joints was selected instead of MASES and Mander enthesis index because the latter two were also highly correlated with other variables. The stiffness variables were excluded because they were highly correlated with BASDAI-fatigue and with the pain variables.

Finally, the following instruments were selected for further analysis: physician spinal pain assessment (ln-transformed), physician assessment of painful peripheral joints (ln-transformed), physician assessment of swollen peripheral joints (ln-transformed), chest expansion, finger-to-floor distance, occiput-to-wall distance (ln-transformed), lateral spinal flexion, cervical rotation, intermalleolar distance, physician's assessment of disease activity (ln-transformed), Dougados articular index, patient's assessment of disease activity, patient night pain, patient pain joints, patient pain spine, BASFI, the BASDAI fatigue question, and ESR (ln-transformed) and CRP (ln-transformed).

Patients available for analysis

In all, 203 OASIS records with baseline data were available. In the case of the patient's perspective a total of 158 OASIS records fulfilled the criteria for low ($n = 112$) or for high ($n = 46$) disease

TABLE 1. Mean, standard deviation (S.D.), median (IQR), minimum, maximum and skewness before and after transformation of baseline data from OASIS population (patients, $n = 203$).

	Mean, S.D.	Median, IQR	Min.–max.	Skewness before transformation	Skewness after ln transformation
Chest expansion (cm)	4.7, 2.2	4.4, 3.0–6.0	0.4–12.5	0.63	
Finger to floor (cm)	14.6, 13.8	12.9, 1.0–22.7	0.0–56.5	0.74	
Cervical rotation (degrees)	63.9, 23.2	68.0, 50.3–80.8	3.3–107.0	−0.67	
Lateral spinal flexion (cm)	10.9, 5.9	10.8, 6.8–15.2	0.0–26.1	0.17	
Intermalleolar distance (cm)	104.5, 21.7	106.0, 93.0–118.0	15.2–150.0	−0.62	
Occiput to wall (cm)	3.8, 5.6	0.0, 0.0–5.8	0.0–26.1	1.87	0.18
Physician spinal pain assessment (0–4)	0.9, 0.9	1.0, 0.0–1.0	0–4.0	1.00	0.20
Physician swollen joints (0–40)	0.8, 2.6	0.0, 0.0–1.0	0.0–31.0	8.29	2.1
Physician painful joints (0–44)	3.3, 5.0	1.0, 0.0–5.0	0.0–41.0	3.04	0.56
Physician assessment of disease activity (VAS 0–10)	2.1, 1.5	1.3, 0.5–3.0	0.0–9.6	1.41	−0.42
Articular index Dougados (0–20)	2.7, 3.3	2.0, 0.0–4.0	0.0–20.0	2.08	0.22
Pain spine, patient (VAS 0–10)	3.5, 2.4	3.2, 1.7–5.2	0.0–9.5	0.36	
Pain joints, patient (VAS 0–10)	3.0, 2.6	2.6, 0.5–4.9	0.0–10.0	0.71	
Patient night pain (0–4)	1.2, 0.8	1.0, 1.0–2.0	0–3.0	0.27	
BASFI (0–10)	3.4, 2.9	3.3, 1.0–5.2	0.0–10.0	0.51	
BASDAI fatigue (VAS 0–10)	4.5, 2.9	4.6, 1.8–7.0	0.0–10.0	0.03	
Patient assessment of disease activity (VAS 0–10)	3.8, 2.8	3.5, 1.3–5.7	0.8–10.0	0.48	
ESR mm/h	14.3, 16.0	10.0, 4.0–18.0	0.0–118.0	2.96	−0.06
CRP (mg/l)	17.8, 24.9	7.0, 6.0–19.0	0.0–139.0	2.73	−0.18
BASDAI total (0–10)	3.5, 2.1	3.2, 1.6–5.3	0.6–9.7	0.39	

activity. The 45 records with an intermediate level of disease activity were not used for further analyses. For the physician perspective of disease activity 145 OASIS records were available for further analyses; 128 with low disease activity and 17 with high disease activity. The 58 records with an intermediate level of disease activity according to the physician were neglected.

We analysed the level of agreement between patients and physicians with respect to rating disease activity according to the above-mentioned criteria. Of the 138 patients with both patient and physician assessment available, disease activity was rated by both patient and physician as 'low' in 96 (70%) of the cases and as 'high' in 11 (8%) of the cases. Discordant opinions were found in 31 patients; 27 patients in whom disease activity was rated as 'high' by the patient and 'low' by the physician, and 4 patients with opposite ratings. These percentages were similar if the analysis was stratified per physician ($n = 3$).

Factor analyses

Factor analysis [SPSS-factor analysis (PCA) with varimax rotation for the final solution] on the OASIS baseline data of the above-described variables was performed to structure the high number of variables. For both patient and physician perspectives of disease activity four factors were extracted with eigenvalues >1, and a cumulative percentage of explained variance of 61%. Table 2 shows the distribution of the selected OASIS baseline variables in a four-factor model for each perspective of disease activity. Both perspectives gave the same factors that consisted of variables reflecting spinal mobility and function (1), variables reflecting assessments by the physician (2), variables reflecting assessments by the patient (3) and variables reflecting laboratory acute phase reactants (4). All separate variables were assigned to that factor for which the factor loading was optimal. The order in which the factors appear slightly differed across the perspectives. The internal consistency of the factors was moderate to good (Cronbach's α ranging from 0.52 to 0.80) (Table 2).

Discriminant analyses

The factor values of the four factors of both disease activity perspectives were used separately (as independent variables) in the discriminant function analyses. We performed discriminant function analyses on the baseline data, and subsequently on data of 6, 12, 18 and 24 months follow-up. The factor values (dependent variables) were restricted to the groups of patients with defined high and low disease activity for both perspectives. Table 3 shows the pooled and canonical correlation coefficients for all four factors for each perspective of disease activity.

Factor 3 (patients' assessments) and to a far less extent factor 2 (physicians' assessments) contributed most to the discriminant score describing disease activity from the patient's perspective, for all time periods analysed. The other two factors did not contribute at all. The pooled correlation for factor 3 remained high if tested at other time points. Factor 3 (physicians' assessments) contributed most to the discriminant score describing disease activity from the physician's perspective, but factor 2 (spinal mobility and function) as well as factor 4 (laboratory acute phase reactants) also contributed significantly. Interestingly, factor 1 (patients' assessments) was far less important. The contribution of factor 2 to the discriminant score was very consistent over time; the contribution of the factors 1, 3 and 4 was somewhat less consistent over time.

Regression analysis

Both discriminant functions (one for the patient perspective and one for the physician perspective) contain four factor values, which are each composed of several clinical variables. In order to get insight into which variables explain variation in the discriminant scores best, we performed stepwise forward multiple linear regression analyses for each disease activity perspective. These regression analyses were performed on the baseline OASIS records, with the two discriminant functions (individual discriminant scores) as dependent variables, and all assessments included in the PCA as explanatory variables.

Table 4 shows the relative contribution of the variables as a result of this stepwise analysis. 'Pain spine', followed by BASFI, 'pain joints', the BASDAI 'fatigue' question and 'physician global assessment of disease activity' together explained more than 90% of variance in discriminant scores for the patient perspective. 'Cervical rotation', followed by 'number of swollen joints', CRP, 'intermalleolar distance' and 'finger-to-floor distance' explained more than 90% of variance in discriminant scores for the physician perspective.

Discussion

The most important conclusion from this study is that patients with AS and physicians have different views on what active AS means. A simple 2×2 analysis that compared patient-reported disease activity and physician-reported disease activity revealed discordant opinions in a substantial proportion of patients (30%).

Patients base their judgement primarily on the presence and severity of complaints related to AS, which are registered by self-administered questionnaires. The impact of spinal mobility, the assessments made by the physician and the acute phase reactants can almost be neglected. The BASFI, an instrument primarily designed to assess function is also contributory in the patient perspective. Apparently, patients base part of their estimation of disease activity on what they are able to (physically) perform. This is a phenomenon that we know from rheumatoid arthritis patients in whom function measured by the Health Assessment Questionnaire (HAQ) is often strongly correlated with disease activity [19].

A few more conclusions can be derived from the observations regarding the patient perspective. First, patients express disease activity primarily in terms of subjective complaints. Objective impairments, such as impaired spinal mobility, are not contributory. Second, acute phase reactants are not important in explaining disease activity from the patient perspective. The physician's judgement about disease activity rests mainly on assessments made by the physician, and patient-derived scores are far less contributory. Actually, the physician's judgement is importantly influenced by assessments of spinal mobility and function, which seem to be related to both disease severity and disease activity. Looking at the combination of instruments that primarily explains the discriminant scores in the physician's perspective, it is remarkable that three of the five variables include instruments combining information on disease activity and severity (cervical rotation, intermalleolar distance and finger-to-floor distance) rather than pure activity (swollen joints and CRP). Both factor 2 (mobility and function) and factor 3 (physician assessments) include all measurements done by the physician. In a separate study designed to investigate on which criteria physicians judge whether a patient with AS should be treated with TNF-blocking drugs, we also found that physicians rated disease severity at least as important as disease activity [20].

It is rather intriguing that we, as physicians, now commonly use patient-derived instruments (such as BASDAI) as a 'gold standard' for measuring disease activity, both for including patients in clinical trials as well as establishing drug effects, whereas there is no appropriate evidence that patients and physicians perceive disease activity similarly. Another important consideration is the lack of evidence that either patient-derived or physician-derived assessments of disease activity are somehow associated with long-term

TABLE 2. Four-factor models with the factor loadings of the assessment belonging to each factor in bold, explained variance and internal consistency of variables reflecting the patient perspective and the physician perspective of disease activity

	Perspective: patient assessment of disease activity				Perspective: physician assessment of disease activity			
Assessments	Factor 1: 'mobility and function'	Factor 2: 'physician assessments'	Factor 3: 'patient assessments'	Factor 4: 'lab'	Factor 1: 'patient assessments'	Factor 2: 'mobility and function'	Factor 3: 'physician assessments'	Factor 4: 'lab'
Spinal pain assessment	0.39	**0.58**	0.30	0.07	0.36	0.41	**0.47**	0.05
Chest expansion	**−0.80**	0.09	−0.21	−0.13	−0.14	**−0.75**	0.03	−0.13
Fingers to floor	**0.66**	0.19	0.13	0.19	0.16	**0.69**	0.14	0.16
Cervical rotation	**−0.58**	−0.56	0.03	−0.11	−0.03	**−0.58**	−0.53	−0.13
Lateral spinal flexion	**−0.83**	−0.09	−0.10	−0.00	−0.12	**−0.82**	−0.09	−0.02
Intermalleolar distance	**−0.48**	−0.44	0.10	0.11	0.02	−0.45	**−0.49**	0.09
Pain spine	0.08	0.03	**0.80**	0.11	**0.81**	0.10	−0.05	0.08
Pain joints	0.00	0.17	**0.68**	0.21	**0.68**	−0.02	0.20	0.21
Night pain	0.10	0.24	**0.68**	0.06	**0.65**	0.12	0.22	−0.02
BASFI	0.46	0.39	**0.55**	0.14	**0.55**	0.45	0.41	0.13
BASDAI fatigue	0.03	0.12	**0.70**	−0.14	**0.72**	0.01	0.09	−0.12
Number of painful joints	0.05	**0.69**	0.31	0.06	0.34	0.03	**0.72**	0.10
Number of swollen joints	−0.04	**0.61**	−0.04	0.43	−0.01	−0.06	**0.68**	0.46
Occiput to wall	**0.78**	0.07	−0.03	0.16	−0.07	**0.77**	0.10	0.17
Patient assessment of disease activity	×	×	×	×	**0.76**	0.09	0.08	0.17
Physician assessment of disease activity	0.03	0.69	0.24	0.23	×	×	×	×
ESR	0.15	0.22	0.02	**0.80**	0.05	0.17	0.12	**0.80**
CRP	0.25	0.00	0.20	**0.80**	0.20	0.26	0.00	**0.80**
Articular index of Dougados	0.21	**0.68**	0.38	−0.13	0.42	0.24	**0.61**	−0.17
	α 0.52	α 0.80	α 0.73	α 0.71	α 0.80	α 0.52	α 0.75	α 0.71
		Explained variance: 61%				Explained variance: 61%		

Cronbach's alpha values refer to variables with bold factor loadings.

outcome in AS (long-term function, structural damage, loss of participation, etc.). Since the options for drug therapy in AS are steadily increasing it becomes more and more important to define measures assessing a uniform construct of disease activity and long-term outcome to be used in clinical trials.

One may question the validity of this study with respect to extrapolation of the findings. The OASIS cohort is an observational cohort with consecutive patients with AS from three different countries, from university hospitals as well as non-university general hospitals. Patients were included irrespective of age, gender, disease duration, severity or activity of their disease, and may therefore be considered an appropriate reflection of the average patient with AS. As shown in Table 1, the OASIS patients cover the entire range of scores of activity and severity variable, which may further add to this conclusion.

In order to be able to truly discriminate between active and inactive disease, we defined inactive disease at a value of $\leq$4.0 and active disease at a value of $\geq$6.0 (VAS from 0 to 10) for both patient and physician perspectives. By doing this, we may limit the interpretability of the findings in case of an 'indifferent' level of disease activity. The goal of this study, however, was to investigate how patients and physicians define inactive and active disease (which asks for a clear distinction), rather than classifying patients as active or inactive. However, the number of patients with high disease activity was rather low, especially in the case of physician perspective. The analyses at other time points revealed similar information as compared with the analyses performed at baseline data, which adds to the validity. These analyses are based on the same patient group and therefore do not contain independent information. Furthermore the patient perspective of disease activity is based on the opinion of a large number of patients; on the other hand the physician perspective is based on the opinion of three investigators only. Physicians and patients had the same access to information. Both investigators and patients were aware of the physical clinical assessments before they made their judgement of disease activity, and they were all unaware of the results of the previous examinations and blood tests. Nevertheless we must assume that the investigators judged disease activity differently because of their extended knowledge of the disease and interpretation of the assessments. Another remark is that one of the investigators assessing 'physician global disease activity' (in fewer than 25% of patients) is a rheumatologist who did not perform all the assessments himself. However, we found no important variation in results by centre. The similarity in results over the entire 2-yr follow-up, in that physician-derived assessments outweigh patient-derived assessments in discriminating between active and inactive disease, might serve as an external validation for the consistency in assessing disease activity by the assessors involved. Probably most important is the finding in another study that the treating physicians of the Dutch OASIS patients based their decisions on starting TNF-blocking agents more on spinal mobility and other disease severity measures than on disease activity measures [20], which is consistent with the results of the physician perspective in this study.

In summary, in this study we showed that patients with AS perceive disease activity differently from physicians. Patients rate complaints and to a lesser extent function as important values defining disease activity. Physicians rate variables reflecting inflammation and severity, such as their own assessments and acute phase reactants, as most important in assessing disease activity instead of patient perception.

The authors have declared no conflicts of interest.

TABLE 3. Pooled correlation of discriminant functions of four factors according to the patient's perspective and the physician's perspective of disease activity in AS at baseline, 6, 12, 18, 24 months follow-up

Perspective and discriminating factor	Baseline	6 months	12 months	18 months	24 months
Patient assessment of disease activity					
Factor 1: 'mobility and function'	0.13	0.12	0.24	0.04	0.14
Factor 2: 'physician'	0.27	0.43	0.42	0.34	0.43
Factor 3: 'patient'	0.85	0.84	0.79	0.81	0.80
Factor 4 'lab'	0.23	0.04	−0.01	−0.14	0.06
Canonical correlation	0.63	0.49	0.49	0.59	0.65
Physician assessment of disease activity					
Factor 1: 'patient'	0.05	0.41	0.07	−0.14	0.37
Factor 2: 'mobility and function'	0.52	0.47	0.55	0.99	0.81
Factor 3: 'physician'	0.62	0.72	0.74	−0.05	0.38
Factor 4: 'lab'	0.48	0.21	0.41	−0.14	0.18
Canonical correlation	0.39	0.44	0.35	0.22	0.25

TABLE 4. Stepwise multivariate regression analysis on discriminant scores obtained from a discriminant function analysis to discriminate between low and high disease activity from a patient perspective and from a physician perspective

		Model: standardized coefficient beta				
Variable	Corresponding factor	1	2	3	4	5
Perspective: patient's assessment of AS disease activity						
Pain spine	3	0.76	0.55	0.44	0.37	0.34
BASFI	3		0.52	0.42	0.36	0.31
Pain joints	3			0.30	0.29	0.29
BASDAI fatigue	3				0.22	0.22
Physician disease activity (logarithms)	2					0.18
Variance (*R* square)		0.58	0.81	0.86	0.90	0.93
Perspective: physician's assessment of AS disease activity						
Cervical rotation	2	−0.74	−0.59	−0.54	−0.44	−0.37
Swollen joints (logarithms)	3		0.46	0.35	0.37	0.38
CRP (logarithms)	4			0.33	0.34	0.30
Intermalleolar distance	2				−0.25	0.24
Finger to floor	2					0.18
Variance (*R* square)		0.54	0.73	0.83	0.88	0.91

Instruments that together contributed >90% of variation in discriminant scores were reported.

References

1. Spoorenberg A, van der Heijde D, de Klerk E *et al*. Ankylosing spondylitis. Relative value of erythrocyte sedimentation rate and C-reactive protein in ankylosing spondylitis. J Rheumatol 1999;26: 980–4.
2. Ruof J, Stucki G. Validity aspects of erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) in ankylosing spondylitis, a literature review. J Rheumatol 1999;26:766–70.
3. van der Heijde D, Calin A, Dougados M, Khan MA, van der Linden S, Bellamy N. Selection of instruments in the core set for DC-ART, SMARD, physical therapy and clinical record keeping in ankylosing spondylitis. Progress report of the ASAS Working Group. Assessments in Ankylosing Spondylitis. J Rheumatol 1999;26:951–4.
4. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis: a proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361–8.
5. The rheumatological physical examination. New York: Grune and Stratton, 1986:55.
6. Miller MH, Lee P, Smythe HA, Goldsmith CH. Measurements of spinal mobility in the sagittal plane: new skin contraction technique compared with established methods. J Rheumatol 1984;11:507–11.
7. Tomlinson M, Barefoot J, Dixon A. Physiotherapy 1996;72:238–40.
8. Spoorenberg A, van der Heijde D, Dougados M *et al*. The reliability of self-assessed joint counts in ankylosing spondylitis (AS). Ann Rheum Dis 2002;61:799–803.
9. Joint motion. Method of measuring and recording. Chicago: American Academy of Orthopedic Surgeons, 1965:44–64.
10. Pile KD, Laurent MR, Salmond CE, Best MJ, Pyle EA, Moloney RO. Clinical assessment of ankylosing spondylitis: a study of observer variation in spinal measurements. Br J Rheumatol 1991;30:29–34.
11. Kennedy LG, Jenkinson TR, Mallorie PA, Whitelock HC, Garrett SL, Calin A. Ankylosing spondylitis: the correlation between a new metrology score and radiology. Br J Rheumatol 1995;34:767–70.
12. Jenkinson TR, Mallorie PA, Whitelock H, Kennedy LG, Garrett Sl, Calin A. Defining spinal mobility in ankylosing spondylitis (AS). The Bath Ankylosing Spondylitis AS Metrology Index. J Rheumatol 1994;21:1694–8.
13. Dougados M, Gueguen A, Nakache JP, Nguyen M, Amor B. Evaluation of a functional index and an articular index in ankylosing spondylitis. J Rheumatol 1988;15:302–7.
14. Mander M, Simpson JM, McLellan A, Walker D, Goodacre JA, Dick WC. Studies with an enthesis index as a method of clinical assessment in ankylosing spondylitis. Ann Rheum Dis 1987;46:197–202.
15. Heuft-Dorenbosch L, Spoorenberg A, van Tubergen A *et al*. Assessment of enthesis in ankylosing spondylitis. Ann Rheum Dis 2003;62:20–6.
16. Jones SD, Steiner A, Garrett SL, Calin A. The Bath Ankylosing Spondylitis Patient Global Score (BAS-G). Br J Rheumatol 1996; 35:66–71.
17. Garrett S, Jenkison T, Kennedy LG, Whitelock H, Gaisford P, Calin A. A new approach to defining disease activity in ankylosing spondylitis: The Bath Ankylosing spondylitis Disease Activity Index. J Rheumatol 1994;21:2286–91.
18. Calin A, Garett S, Whitelock H *et al*. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. J Rheumatol 1994; 21:2281–5.
19. Drossaers-Bakker KW, de Buck M, van Zeben D, Zwinderman AH, Breedveld FC, Hazes JM. Long-term course and outcome of functional capacity in rheumatoid arthritis: the effect of disease activity and radiologic damage over time. Arthritis Rheum 1999; 42:1854–60.
20. Landewé R, Rump B, van der Heijde D, van der Linden S. Which patients with ankylosing spondylitis should be treated with tumor necrosis factor inhibiting therapy? A survey among Dutch rheumatologists. Ann Rheum Dis 2004;63:530–4.