Reprinted from:
The Journal of Rheumatology 1994;21:2286–91

The Journal of Rheumatology

A New Approach to Defining Disease Status in Ankylosing Spondylitis: The Bath Ankylosing Spondylitis Disease Activity Index

SARAH GARRETT, TIM JENKINSON, L. GAIL KENNEDY, HELEN WHITELOCK, PENNY GAISFORD, and ANDREI CALIN

ABSTRACT. ***Objective.*** Disease status, in terms of disease activity, disease progression and prognosis is difficult to define in ankylosing spondylitis (AS). No gold standard exists. Therefore, the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), a self-administered instrument, has been developed as a new approach to defining disease activity in patients with AS.

Methods. The index, designed by a multidisciplinary team with input from patients, consists of six 10 cm horizontal visual analog scales to measure severity of fatigue, spinal and peripheral joint pain, localized tenderness and morning stiffness (both qualitative and quantitative). The final BASDAI score has a range of 0 to 10. The index was distributed to a cross section of patients, including inpatients receiving 3 weeks of intensive physiotherapy treatment and hospital outpatients. BASDAI was completed by a total of 154 patients. Validation of the new instrument was achieved through analysis of user friendliness, reliability (consistency), score distribution and sensitivity to change. Comparisons were made with a previous Bath disease activity index (DAI) and the Newcastle Enthesis Index.

Results. The BASDAI was found by patients to be quick and simple to complete (mean: 67 s). Test-retest reliability was good ($r = 0.93$; $p < 0.001$), as was the distribution of scores across the scale (score range: 0.5–10; mean: 4.31). BASDAI was sensitive to change, reflecting a 16% (mean) improvement in inpatient scores after 3 weeks of treatment. It is superior to the DAI in terms of construct and content validity and to the Enthesis Index in all aspects.

Conclusion. In summary, BASDAI is user friendly, reliability, sensitive to change and reflects the entire spectrum of disease. It is a comprehensive self-administered instrument for assessing disease activity in AS. (*J Rheumatol 1994;***21:2286–91**)

Key Indexing Terms:
ANKYLOSING SPONDYLITIS
SELF-ADMINISTERED INSTRUMENT
DISEASE ACTIVITY
VALIDITY

Disease status, in terms of disease activity, disease progression and prognosis is difficult to define in ankylosing spondylitis (AS)[1]. Fundamental to investigating the natural history of the disease is the assessment of outcome[2], for which radiology, metrology, and measures of functional (dis)ability are tools[1,3]. Such measurements of damage and its functional consequences should, however, be distinguished from measures of disease activity[1]. The assessment of disease activity in a predominantly axial disease such as AS is notoriously difficult and, as yet, no gold standard exists[1,4]. In contrast to the situation in rheumatoid arthritis (RA), laboratory indicators of disease activity reflect neither clinical activity nor radiological progression, and their use in AS is controversial[1] In RA, a core set of disease activity measures has been introduced[5–7], providing an advance over previous techniques[8–10]. In AS, such experience is limited. The Newcastle Enthesis Index[11] and a previous Disease Activity Index (DAI)[12] have been described, but these either fail as true measurements of disease activity or have not been adequately validated. In addition, neither are fully comprehensive. For example, recent research at the Royal National Hospital for the Rheumatic Diseases (RNHRD) has demonstrated that fatigue is a major component of AS for many patients[13] and this should now be incorporated into any measurement of disease activity. A new and comprehensive index is therefore necessary. To this end, the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), a self-administered instrument, has been developed.

Combining individual variables, which may have little value as single measures, into an index bestows a number of advantages, such as improved validity[3], avoidance of duplicity and increased sensitivity to change[6,14], producing a more powerful indicator of outcome[15]. Increased sensitivity of an index can provide statistical advantages and may

From The Royal National Hospital for Rheumatic Diseases, Bath, UK.

Supported by The Arthritis and Rheumatism Council, the National Ankylosing Society, the Pilkington Trust and the Coates Trust.

S.L. Garrett, BA, Research Associate; T.R. Jenkinson, MRCP, Senior Registrar; L.G. Kennedy, BSc, Research Associate; H.C. Whitelock, DipPhys, MCSP HT, Superintendent Physiotherapist; P.M. Gaisford, GradDipPhys, MCSP HT; A. Calin, MD, FRCP, Consultant Rheumatologist.

Address reprint requests to Dr. A. Calin, RNHRD, Upper Borough Walls, Bath, BA1 1RL, UK.

Submitted February 1, 1994 revision accepted May 18, 1994.

substantially reduce the size of the required sample[3,15]. As stressed by Bombardier and Tugwell[16], measurement indices must satisfy 5 recognized validity criteria: content (the choice and relative importance of each component is appropriate for the purpose of the index); face (the methods of weighting and aggregating components into an index are sensible); criterion (the index produces consistent results that reflect the true clinical state of the patient); discriminant (the index detects the smallest clinically significant difference between and within patients); and construct (the index agrees with expected results based on the hypothesis of the investigator). A self-assessment instrument should be reliable, reproducible and reflect the entire spectrum of the disease severity. It needs, in addition, to be quick and simple to complete. Finally it has the advantage of providing an inexpensive method of obtaining clinical information that can be safely and frequently repeated.

MATERIALS AND METHODS

The index BASDAI was developed, on the basis of clinical experience, by a team of physiotherapists, research associates and rheumatologists, with a major input from patients with AS. The resulting instrument consists of 6 questions relating to 5 major symptoms relevant to AS: fatigue, spinal pain, joint pain/swelling, areas of localized tenderness and morning stiffness (Figure 1). The last is measured in terms of both quality (degree of stiffness) and quantity (length of time for which stiffness persists). BASDAI

BASDAI

The Bath Ankylosing Spondylitis Disease Activity Index

PLEASE PLACE A MARK ON EACH LINE BELOW TO INDICATE YOUR ANSWER TO EACH QUESTION, RELATING TO THE PAST WEEK.

(1) How would you describe the overall level of fatigue / tiredness you have experienced?

NONE _____ VERY SEVERE

(2) How would you describe the overall level of AS neck, back or hip pain you have had?

NONE _____ VERY SEVERE

(3) How would you describe the overall level of pain/swelling in joints other than neck, back or hips you have had?

NONE _____ VERY SEVERE

(4) How would you describe the overall level of discomfort you have had from any areas tender to touch or pressure?

NONE _____ VERY SEVERE

(5) How would you describe the overall level of morning stiffness you have had from the time you wake up?

NONE _____ VERY SEVERE

(6) How long does your morning stiffness last from the time you wake up?

0 hrs — 1/2 — 1 — 1 1/2 — 2 or more hrs

Fig. 1. The BASDAI.

requires patients to indicate the degree to which they have experienced these symptoms over the past week.

Ten centimetre visual analog scales (VAS) were used to measure the patient's response to each question as they allow maximum reliability, and sensitivity to change and improve the capacity of an index to elicit a range of responses across the entire scale. In accordance with previous work[17] the VAS were unmarked, except by the words "none" at the start and "very severe" at the end of each line. The exception was a 0–2 h time scale (marked at every quarter of an hour) used to measure quantity of morning stiffness. The time scale given for this measurement was derived from an analysis of retrospective data regarding duration of morning stiffness, in about 2000 patient questionnaires (unpublished data). From the range of possible answers ("over 4 hours" to "do not have morning stiffness") the mean response was found to be 30 min to 1 h. Thus 1 h was taken as the midpoint of the time scale, with 2 or more h being given the maximum score.

Each visual analog scale was scored from 0 to 10. The mean of the 2 scores relating to morning stiffness was taken, providing an aggregate score. Thus each symptom is given equal weighting. The resulting 0–50 score for the overall index was converted to a 0–10 scale to give the final BASDAI score.

A pilot questionnaire was given to a group of patients on a 3-week intensive physiotherapy course. Modifications to the index, such as the wording of questions and the inclusion of quality of morning stiffness were made as a result of patient feedback.

The final version of the BASDAI was completed by 4 sets of inpatients (n = 46) on 4 separate occasions during their physiotherapy course: Days 0, 1, 8 and 18. One hundred and eight other spondylitics, including RNHRD outpatients and members of various NASS (National Ankylosing Spondylitis Society) self-help groups also completed the instrument. This resulted in a total of 292 questionnaires completed by 154 patients.

As a means of comparison, the earlier Bath DAI[12] was given to patients each time they completed the BASDAI. In addition, the Newcastle Enthesis Index[11] was carried out by a physiotherapist on Days 8 and 18 of treatment with 25 inpatients. Both the DAI and the Newcastle Enthesis Index scores were converted to a 0–10 scale to enable direct comparison with the BASDAI.

User friendliness, test-retest reliability, score distribution and sensitivity to change were all analyzed for the BASDAI and were compared to results from the DAI and the Newcastle Enthesis Index. Reliability was assessed by testing the correlation between BASDAI scores taken at the same time of day on Days 0 and 1 of the inpatient course; appropriate use of the scale was ascertained from the range and mean of scores given by the total patient population (using the Day 0 score for the inpatient cohort); while sensitivity of the BASDAI to change was analyzed by a comparison of inpatient scores on Day 0 and Day 18 of treatment.

Possible redundancy within the index was tested for by analyzing the degree of association between scores given for each of the 6 questions.

Analysis of results was carried out using the UNISTAT statistical software on an IBM compatible PC. The Pearson correlation coefficient was used to perform correlations; the Wilcoxon signed rank test and the Kruskal-Wallis one way ANOVA for analyses of difference.

RESULTS

The mean age of the 154 patients (115 men: 39 women — a 2.9:1 ratio) who completed the BASDAI was 47.7 (SD 11.29; inpatients: 47.1, outpatients 47.9), with a mean age at disease onset of 23.0 (SD 7.81; inpatients: 23.8, outpatients: 22.6), and a mean disease duration of 24.7 years (inpatients: 23.3, outpatients 25.3).

The BASDAI was found by patients to be both a quick and simple index, taking between 30 s and 2 min to complete (mean: 67 s). The index proved to be highly reliable in terms of the consistency of inpatient scores measured 24 h apart (Day 0 of treatment: mean score = 5.434, SD 2.38; Day 1: mean score = 5.438, SD 2.24; r = 0.93; p <0.001).

The capacity of the BASDAI to elicit a range of responses across the scale was good, with a score range of 0.5–10 from the whole patient sample (n = 154), with a mean score of 4.31, SD 2.12 (Figure 2). The mean BASDAI score for the inpatients (5.06) was significantly higher than that of the outpatients (4.0; p = 0.005).

All of the individual symptoms showed good score distribution, with scores spread across at least 95% of the scale. Correlations between each set of symptom scores ranged between r = 0.34 ("fatigue" versus "joint pain") and r = 0.66 ("spinal pain" versus "localized tenderness"). The correlation between quality and quantity of morning stiffness was closer: r = 0.79.

The comparison of inpatient scores on Day 0 with those on Day 18 of physiotherapy showed the BASDAI to be sensitive to change. The mean scores for Days 0 and 18 were 5.31 (SD 1.74) and 4.46 (SD 2.21), respectively, reflecting a significant improvement over this period of treatment (p = 0.009; mean score change = −0.85 [16.4% improvement]; range = −4.0 to +2.1).

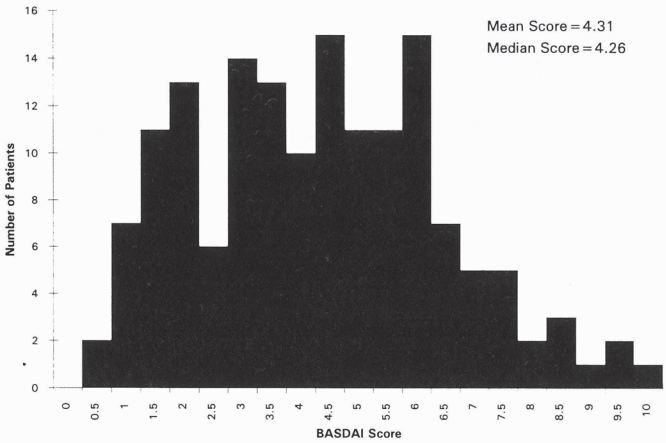

Fig. 2. The distribution of BASDAI scores among 154 patients (mean score = 4.31).

Table 1. *Comparison of DAI vs BASDAI in terms of the validity criteria analyzed*

	DAI	versus	BASDAI
Time (mean seconds)	75	NS	67
Reproducibility (Day 0 vs Day 1)	r = 0.96	NS	r = 0.93
Score distribution			
Mean:	4.12	NS	4.31
Range:	0–9.5		0.5–10
Sensitivity (Day 1 vs Day 18)			
Mean change:	−1.22 (p = 0.002)	NS	−0.87 (p = 0.009)
Range:	−6.3–+2.1		−4.0–+2.1
% Improvement:	22.8%		16.4%

The BASDAI correlated well with the DAI in all aspects of validity criteria (Table 1). However, a significantly higher number of patients ($\chi^2 = 7.21$, $p = 0.009$) felt that the BASDAI contained the most suitable questions for obtaining information on the symptoms of those with AS. Inpatients took a mean of 75 s (SD: 34.59) to complete the DAI which, like the BASDAI, was highly reliable (mean score on Day 0 of treatment = 5.38, SD 2.29; mean score on Day 1 = 5.66, SD 2.46; $r = 0.96$, $p < 0.001$). Score distribution across the scale was equal to that of the BASDAI (mean score = 4.12, SD 2.10; range = 0–9.5; Figure 3), and there was a good correlation between the 2 sets of scores ($r = 0.75$; $p < 0.001$). The mean inpatient DAI score (5.13) was again significantly higher than that of the outpatients (3.79; $p = 0.001$). The DAI reflected the sensitivity to change shown by the BASDAI, with mean inpatient scores improving from 5.34 on Day 0 of the 3 week course to 4.12 by Day 18 ($p = 0.002$; mean score change = −1.22 [22.8% improvement]; range = −6.3 to +2.1). There was no significant partiality among the patients for either the DAI or the BASDAI in terms of which questionnaire was the easiest to understand and complete, or overall questionnaire preference.

Among the cohort of 25 inpatients on whom the Newcastle Enthesis Index was also carried out, the BASDAI showed superior distribution of scores across the scale and greater sensitivity to change over 10 days of physiotherapy. The mean Enthesis Index score on Day 8 of treatment was 1.96, with a range of 0–5.33, compared to a mean BASDAI score of 5.06, with a range of 0.83–8.79. While the mean BASDAI score improved significantly from 5.06 to 4.15 by Day 18 of treatment ($p = 0.009$; mean score change = −0.91; range = −3.7 to +1.66), there was no comparable change in the Enthesis Index scores over this period (mean score on Day 8 = 1.96 vs mean on Day 18 = 1.79, $p = 0.35$; mean score change = −0.18; range = −1.73 to +1.22). Reliability of the Newcastle Enthesis Index has already been shown to be poor[11].

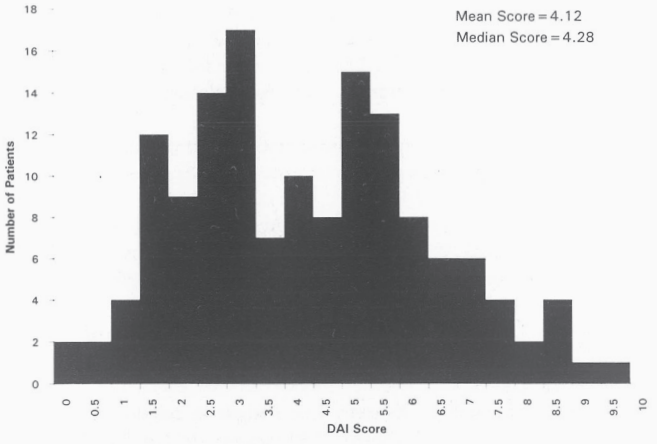

Fig. 3. DAI: The distribution of DAI scores among 154 patients (mean score 4.12).

DISCUSSION

The development of the BASDAI was stimulated by a dissatisfaction with existing measurements of disease activity. For example, the Newcastle Enthesis Index is inadequate. Specifically, it is very limited in content, focussing purely on the entheses, and thus does not fully address the range of symptoms in AS. It exhibits neither sufficient reliability, score range, nor sensitivity to change. A further (major) disadvantage of the measurement is that it requires a trained clinician or physiotherapist to perform the assessment and is thus expensive in terms of time and finance. Likewise, the earlier Bath DAI is not fully comprehensive. Specifically, it omits reference to fatigue, quality of morning stiffness and localized tenderness.

Naturally, the components included in any assessment instrument are not exhaustive, but represent only a few of the possible questions. If there is a very high correlation between the responses to 2 different questions, then the information obtained from one question mirrors that obtained from the other (internal redundancy). It is therefore possible to develop indices which adequately represent a dimension of disease (e.g., disease activity) without specifically addressing every possible question[18]. This is an important concept since there is an inverse correlation between the number of components included in an instrument and the accuracy of responses obtained[19].

The five components of the BASDAI were regarded as vital in ascertaining a comprehensive picture of a patient's disease activity. Fatigue, previously overlooked in AS, was included in the light of recent research concluding that it is an important and common symptom in this disease[8]. This research was a direct result of patient feedback. It was recognized that there are 2 main sources of pain in AS, spinal pain and pain in peripheral joints. These constitute separate symptoms and need to be measured as such. Localized tenderness was included in order to assess severity of enthesitis. Morning stiffness was recognized as having two significant aspects: not only is the length of time for which morning stiffness persists important, but the degree of stiffness should also be taken into account. For example, a patient with only 15 min of stiffness each morning but who suffers severe debility during this period is (arguably) as severely affected as a patient with stiffness for over 2 h each day but who suffers very little reduction of function as a result. The inclusion of both the quality and quantity of morning stiffness does not however give this symptom excess weighting in the index since an aggregate (mean) score is taken.

None of the individual components of the BASDAI correlated closely, signifying that none of the questions in the index are redundant and thus vindicating the inclusion of each symptom. The highest correlation in the index was, as expected, between the quality and quantity of morning stiffness ($r = 0.79$, $p < 0.001$). However, there was sufficient

noncorrelation to justify the inclusion of both aspects of the symptom (Figure 4).

Although, after validation, the DAI showed greater change over 3 weeks of treatment than did the BASDAI (22.8 vs 16.4% score improvement; NS), this may be a result of its bias towards pain and its inclusion of a scale measuring the patient's well-being. The score change is, arguably, more likely to be a reflection of the effects of the intensive physiotherapy program on these aspects than a greater sensitivity of the DAI to change in disease activity as a whole. The BASDAI does not differ from the DAI in any of the other areas analyzed (reliability, use of the scale and user friendliness), but is superior, from the perspective of both patients and clinicians, in terms of face and content validity, comprehensiveness of symptoms and their weighting. The BASDAI has, in addition, proved to be superior in all aspects to the Newcastle Enthesis Index as a measure of disease activity.

The difference apparent between the mean scores of the inpatients and outpatients for both the BASDAI and the DAI were predictable, since patients on an intensive hospital course would naturally tend to be those with worse disease.

The ability of the BASDAI to distinguish clinically significant differences in both inter and intrapatient responses (discriminant validity)[9] was shown by the total range of scores among patients, reflecting the entire disease spectrum, and further by the sensitivity of the index to change within patients on the 3-week physiotherapy course. The 16% (mean) improvement in the symptoms of those patients after 3 weeks of physiotherapy is likely to be of clinical significance. Further study is required however into the sensitivity of the index in relation to drug therapy. This may have a more dramatic effect on results than did the inpatient program of physiotherapy. Although there is no disease modifying drug for AS (in contrast to the perceived situation in RA), nonsteroidal antiinflammatory drugs do have a marked effect on symptoms and therefore on disease activity. BASDAI could thus be incorporated into clinical pharmaceutical studies.

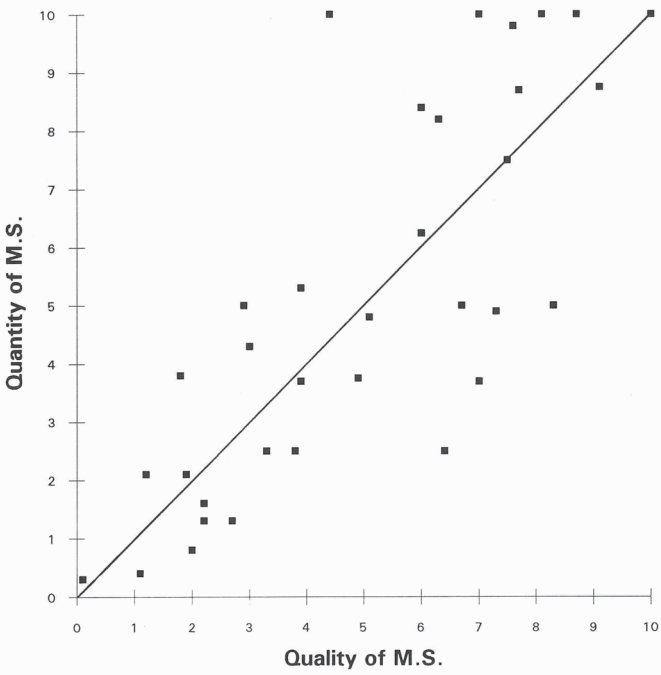

Fig. 4. Morning Stiffness: The degree of association between the quality and quantity.

In summary, BASDAI is a comprehensive new index for the measurement of disease activity in AS. It is user friendly, highly reliable, reflects the entire spectrum of disease and exhibits the ability to be sensitive to clinical changes. Most striking is its recognition of the five major symptoms experienced by patients with AS, some of which have not been addressed in previous measurements of disease activity. A more precise definition of disease activity will lead to an enhanced understanding of outcome and prognosis in AS[20].

REFERENCES

1. Taylor HG, Wardle T, Beswick EJ, *et al:* The relationship of clinical and laboratory measurements to radiological change in AS. *Br J Rheumatol 1991;*30:330–5.
2. Rigby AS, Silman AJ: Editorial. Outcome assessment in clinical trials of AS. *Br J Rheumatol 1991;*30:321–2.
3. van der Heijde DMFM, van't Hof MA, van Reil PLCM, van de Putte LBA: Validity of single variables and indices to measure disease activity in rheumatoid arthritis. *J Rheumatol 1993;*20:538–41.
4. Calin A: Editorial. Assessing disease activity in AS. *Lancet 1987;*1:1072.
5. Felson DT: Outcome measures for rheumatoid arthritis. *J Rheumatol 1993;*20:531–4.
6. Boers M, Tugwell P: The validity of pooled outcome measures (indices) in rheumatoid arthritis clinical trials. *J Rheumatol 1993;*20:568–74.
7. Tugwell P, Boers M: OMERACT conference on outcome measures in rheumatoid arthritis clinical trials: conclusion. *J Rheumatol 1993;*20:590.
8. Davis MJ, Dawes PT, Fowler PD, *et al:* Comparison and evaluation of a disease activity index for use in patients with rheumatoid arthritis. *Br J Rheumatol 1990;*29:111–5.
9. Haataja M, Kalliomak JL: Laboratory scale for evaluating the activity of rheumatoid arthritis. *Rheumatol Rehabil 1978;*17:83–5
10. Mallya RK, Mace BEW: The assessment of disease activity in rheumatoid arthritis using a multivariate analysis. *Rheumatol Rehabil 1981;*20:14–7.
11. Mander M, Simpson JM, McLellan A, Walker D, Goodacre JA, Carson Dick W: Studies with an enthesis index as a method of clinical assessment in ankylosing spondylitis. *Ann Rheum Dis 1987;*46:197–202.
12. Kennedy LG, Edmunds L, Calin A: The natural history of ankylosing spondylitis. Does it burn out? *J Rheumatol 1993;*20:688–92.
13. Calin A, Edmunds L, Kennedy LG: Fatigue in ankylosing spondylitis - why is it ignored? *J Rheumatol 1993;*20:991–5.
14. Roberts R: Pooled outcome measures in arthritis: the pros and cons. *J Rheumatol 1993;*20:566–7.

15. Goldsmith CH, Smyth HA, Helewa A: Interpretation and power of a pooled index. *J Rheumatol 1993;*20:575–8.
16. Bombardier C, Tugwell P: A methodical framework to develop and select indices for clinical trials: statistical and judgmental approaches. *J Rheumatol 1982;*9:753–7.
17. Bird HA, Dixon JS: Measurement of pain. In: Wright V, ed. *Balliere's Clin Rheumatol 1987;*1:75–80.
18. Fries JF, Spitz PW, Young DY: The dimensions of health outcomes: the Health Assessment Questionnaire, disability and pain scales. *J Rheumatol 1982;*9:789–93.
19. Levy P, Lemeshow S: *Sampling for Health Professionals.* Belmont: Wadsworth Publishing, 1980.
20. Calin A: Editorial. Can we define the outcome of ankylosing spondylitis and the effect of physiotherapy management? *J Rheumatol 1994;*21:184–5.