RHEUMATOLOGY

Rheumatology 2014;53:353–360
doi:10.1093/rheumatology/ket359
Advance Access publication 5 November 2013

Original article

Performance of the Assessment of Spondyloarthritis International Society criteria for the classification of spondyloarthritis in early spondyloarthritis clinics participating in the ESPERANZA programme

Eva Tomero[1], Juan Mulero[2,3], Eugenio de Miguel[4], Cruz Fernández-Espartero[5], Milena Gobbo[6], Miguel A. Descalzo[6], Eduardo Collantes-Estévez[7,8], Pedro Zarco[9], Santiago Muñoz-Fernández[10], Loreto Carmona[11] and the ESPERANZA Study Group*

Abstract

Objective. The objective of this study was to analyse the performance of the Assessment of SpondyloArthritis International Society (ASAS) criteria for the classification of SpA in early SpA clinics.

Methods. We used a cross-sectional study of patients referred to early SpA units within the ESPERANZA programme (a Spanish nationwide health management programme designed to provide excellence in diagnosis and care for early SpA). Patients were eligible if they were <45 years of age and had any of the following: (i) a 2-year history of inflammatory back pain; (ii) back or joint pain with psoriasis, anterior uveitis, radiographic sacroiliitis, family history of SpA or positive HLA-B27; or (iii) asymmetric arthritis. We excluded patients for whom imaging (X-rays/MRI) or HLA-B27 results were not available. We analysed the performance (sensitivity and specificity) of different classification criteria sets, taking the rheumatologist's opinion as the gold standard.

Results. The analysis included 775 patients [mean age 33 (S.D. 7) years; 55% men; mean duration of symptoms 11 (S.D. 6) months]; SpA was diagnosed in 538 patients (69.5%). A total of 274 (67.9%) patients with chronic back pain met the ASAS axial criteria, 76 (56.3%) patients with arthritis but not chronic back pain fulfilled the ASAS criteria for peripheral SpA and 350 (65.1%) fulfilled all the ASAS criteria. The sensitivity and specificity of the ASAS criteria set were 65% and 93%, respectively (axial criteria: sensitivity 68%, specificity 95%). The sensitivity and specificity for the ESSG and Amor criteria were 58% and 90% and 59% and 86%, respectively.

Conclusion. Despite performing better than the Amor or ESSG criteria, the ASAS criteria may be limited to detection of early forms, particularly in populations in which MRI is not extensively available or in populations with a low prevalence of HLA-B27.

Key words: early spondyloarthritis, classification criteria, clinical practice, validation studies.

[1]Rheumatology Department, Hospital Universitario de la Princesa, Instituto de Investigación La Princesa, Madrid, [2]Rheumatology Department, Hospital Universitario de Puerta de Hierro-Majadahonda, Majadahonda, [3]School of Medicine, Universidad Autónoma de Madrid, [4]Rheumatology Department, Hospital Universitario de la Paz, Madrid, [5]Rheumatology Department, Hospital de Móstoles, Móstoles, [6]Research Unit, Fundación Española de Reumatología, Madrid, [7]Rheumatology Department, Hospital Universitario Reina Sofía, Granada, [8]School of Medicine, Universidad de Córdoba, Córdoba, [9]Rheumatology Department, Fundación Hospital Universitario Fundación Alcorcón, [10]Rheumatology Unit, Hospital Infanta Sofia, San Sebastián de los Reyes and [11]Health Sciences School, Universidad Camilo José Cela and Institute for Musculoskeletal Health, Madrid, Spain.

Submitted 6 December 2012; revised version accepted 10 September 2013.

Correspondence to: Eva Tomero, Hospital Universitario de la Princesa, Servicio de Reumatología, C/Diego de León 62, Madrid 28006, Spain. E-mail: tomeroeva@yahoo.es

*For a list of the members of the ESPERANZA Study Group please see supplementary data, available at *Rheumatology* Online.

Introduction

The term SpA refers to a heterogeneous group of diseases that is considered a single entity according to the classification criteria of the ESSG and of Amor [1–4]. However, applying these criteria in patients with a recent onset of symptoms or early SpA results in underclassification, thus warranting the development of new criteria [5–7].

The study of the German cohort of early axial disease introduced the concept of non-radiological axial disease, which represents a developmental stage in the clinical spectrum of axial disease [8–11]. This concept has been consolidated thanks to MRI, which can detect sacroiliac inflammation before any radiological damage is visible and thus identify patients with early-stage axial predominance [12–15]. Non-radiological axial disease was included in the new classification criteria for axial SpA of the Assessment of SpondyloArthritis International Society (ASAS) [16]. Subsequently ASAS published criteria for SpA in patients with predominantly peripheral disease; all clinical forms can be classified using the new ASAS criteria [16–18]. With the local expert's opinion as the gold standard, these criteria have 82.9% sensitivity and 84.4% specificity for classification of patients with a mean disease duration of 6 years, while in the same dataset the sensitivity and specificity of previous criteria are lower [17].

The ASAS criteria were recently tested in different settings. Peripheral SpA criteria in particular were tested in a cohort of patients with recent-onset arthritis and the full criteria in a cohort with long-standing SpA [18, 19]. It would be desirable to assess the behaviour of the ASAS criteria in a population of early SpA including both axial and peripheral forms.

In Spain, the national programme for the care of early SpA is known as ESPERANZA [20], whose key goals are to promote and establish early SpA units and to facilitate early diagnosis and treatment. Patients with suspected SpA and a maximum 2 years from the onset of symptoms are referred from primary or specialized care to the programme, where they are attended by rheumatologists specifically trained in SpA. The population served by the programme is ideal for validating the different classification criteria in early SpA, with the additional advantage that it includes both axial and peripheral SpA.

Based on the ESPERANZA programme, the objectives of the present study were to assess the performance of the ASAS classification criteria in early SpA. We also aimed to analyse the predictive capacity of the ASAS, Amor, and ESSG criteria for identifying rheumatologist-diagnosed SpA and evaluate specific clinical features, laboratory findings and imaging tests results that increase the likelihood of a patient being classified as SpA either by the ASAS criteria or by a rheumatologist.

Methods

We performed a cross-sectional study within the framework of the ESPERANZA programme, which has been described in detail elsewhere [20]. Briefly, patients are referred to early SpA units in participating areas following agreed referral criteria and specific protocols and training. Once patients are accepted in the programme, their data are entered into a web-based medical record that makes it possible to monitor the quality of care as well as conduct research. The programme generates an ongoing early SpA cohort from which the study sample was drawn. The programme was reviewed and approved by the Research Ethics Committee of Hospital Reina Sofía, Cordoba, Spain. The approval covers the analysis of data described in this study.

Patients

Patients were eligible for the programme if they were referred from primary or specialized care (rheumatology, gastroenterology, ophthalmology, orthopaedics and emergency services), were aged <45 years, had 3 months to 2 years from the onset of the symptoms and had at least one of the following conditions: (i) inflammatory back pain (IBP), (ii) asymmetrical arthritis, predominantly in lower limbs, or (iii) back pain or arthralgia with psoriasis, IBD, uveitis, radiographic sacroiliitis, positivity for HLA-B27 or a family history of SpA. We used the information of all patients included from the start of the programme in April 2008 to June 2011. For the present analysis we excluded patients if the rheumatologist considered that the referral criteria were not satisfied and if HLA-B27 or imaging (sacroiliac X-ray or MRI) data were not available. Patients with no HLA-B27 or imaging data (patients excluded) were compared with patients for whom data were available (patients included).

Variables and measurements

The ESPERANZA programme makes it possible to record demographic data, symptoms related to SpA included in each of the classification criteria (ASAS, Amor and ESSG), specific metrology parameters recommended for this group of diseases [21–27] and analytical parameters, including HLA-B27, and imaging (X-rays and/or sacroiliac joint MRI). The decision to perform an MRI was at the discretion of the unit rheumatologist. The definition of positive MRI was based on the ASAS/OMERACT definition of active sacroiliitis on MRI [28]. Definitions of the variables used in the ESPERANZA programme are included in supplementary Table S1, available at *Rheumatology* Online. All these data were included in the online information system after the patient had signed the informed consent.

Reference standard

The reference standard was the clinical diagnosis made by the unit rheumatologist. All participating rheumatologists were specifically trained in SpA, had access to constantly updated materials and are considered experts in the field. Additionally, data from patients included in the analysis were reviewed by four rheumatologists with extensive experience in SpA (E.C.-E., E.M., J.M. and P.Z.) and who assigned a secondary diagnosis. Since

concordance between experts and reviewers was good ($\kappa = 0.71$) for both positive and negative diagnoses of SpA, the rheumatologist's diagnosis was considered valid, assuming that he/she possessed information not accessible to the reviewers. In cases where the unit rheumatologist had not yet made a final diagnosis, the expert's clinical judgment was applied, given that at least three of them agreed on the following options: SpA, non-SpA and unlikely SpA. The option of unlikely SpA was only accepted in cases where the patients did not have enough features to make a diagnosis of SpA, but reviewers still could not exclude this option at that time. Patients classified as unlikely SpA were reclassified as non-SpA. We carried out a sensitivity analysis including and excluding unlikely SpA from the group SpA.

Criteria for classification

Each patient was classified according to the ESSG, Amor and ASAS criteria with all available data during follow-up. For analysis of the axial ASAS criteria we considered only those patients with low back pain, and for the peripheral criteria only those patients with peripheral arthritis, dactylitis or enthesitis in the absence of chronic axial pain. We thus prevented the same patient from being classified as both an axial or peripheral case.

Statistical analysis

The study sample was described according to the distribution of the variables using measures of central tendency and dispersion for quantitative variables and the frequency and percentages for qualitative variables. The performance and validity of the different sets of classification criteria (ESSG, Amor and ASAS) were analysed in terms of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and likelihood ratio (LR). We used a Venn diagram to represent the percentage of patients who met each of the criteria (ESSG, Amor and ASAS) within the group of patients diagnosed with SpA.

Bivariate and multivariate logistic regression models were constructed to identify the parameters (clinical, laboratory or imaging) that contributed the most to the diagnosis of SpA by the rheumatologist and to the classification of SpA according to ASAS criteria. All parameters were estimated with a 95% CI. Statistical significance was set at 0.05. The analyses were performed with Stata 11.1 (StataCorp, College Station, TX, USA).

Results

The eligible population comprised 1179 patients enrolled from 25 centres. Of these, 98 did not meet the referral criteria according to the rheumatologist and were excluded, as were 130 patients for whom HLA-B27 data were not available and 176 patients whose X-rays and/or sacroiliac MRI data were not available (Fig. 1). In total, 775 patients were considered for analysis: 538 (69.5%) were diagnosed as SpA, 182 (23.5%) were not SpA and 55 (7%) were unlikely SpA (later reclassified to non-SpA). In 131 patients the diagnosis of the experts was accepted over a missing diagnosis. Comparisons between the 306 excluded patients and 775 included patients are presented in supplementary Table S2, available at *Rheumatology* Online. Although some differences emerge, for example, in the percentage of patients with IBP, the groups are quite similar.

Most patients had been referred from primary care (68%), mainly because of IBP (74%); 13% were referred because of asymmetric arthritis. The mean age of the patients was 33.1 (S.D. 7.1) years and 54% were men. Time from onset of symptoms was short, 11.9 (S.D. 6.6) months. The mean follow-up until the end of this study was 0.9 (S.D. 0.8) months with a median of two visits per patient. The most common musculoskeletal manifestation was chronic low back pain (73.4%) and the most common extra-articular manifestation was psoriasis (13.9%). The prevalence of HLA-B27 in this population was 55.6%, and 19% of the population had established radiological sacroiliitis (bilateral grade II or unilateral grade III or IV). The remaining descriptive parameters are summarized in Table 1.

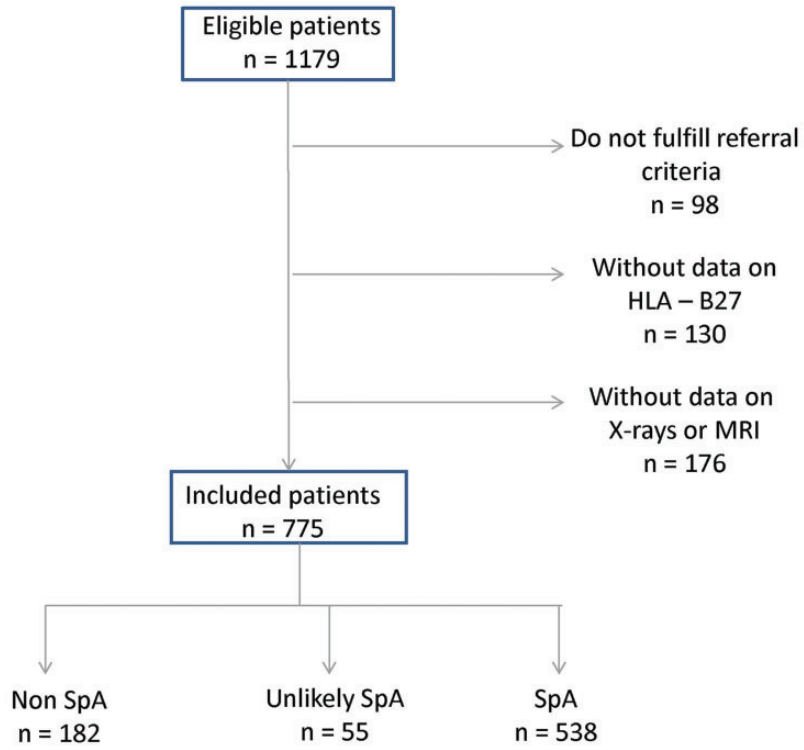
As for the ASAS criteria, 65.1% of the patients fulfilled the complete set; 67.9% of the patients with chronic low back pain ($n = 403$) fulfilled the axial ASAS criteria and 56.3% of the patients without chronic low back pain ($n = 135$) fulfilled the peripheral ASAS criteria. The sensitivity and specificity of the ASAS criteria in this early SpA sample were higher (65% and 93%, respectively) than those of the ESSG (58% and 90%, respectively) and of Amor (59% and 86%, respectively). The performance of each criteria set is summarized in Table 2. The sensitivity analysis including unlikely SpA patients as SpA instead of non-SpA, as well as excluding them from all analyses, did not affect the results (data not shown).

The multivariate analysis showed that the most influential parameters in the diagnosis of SpA by the rheumatologist were the presence of sacroiliitis in imaging tests [odds ratio (OR) 28.3, $P < 0.001$], peripheral arthritis (OR 7.2, $P < 0.001$), HLA-B27 positivity (OR 5.9, $P < 0.001$) and psoriasis (OR 4.7, $P < 0.001$) (Table 3). The parameters that had a major influence on the ASAS classification of SpA were the presence of sacroiliitis in imaging tests ($OR > 100$, $P < 0.001$) and positivity for HLA-B27 (OR 69.3, $P < 0.001$). When only the sacroiliac MRI data were taken into account and not the X-ray data, the OR was 38.3 ($P < 0.001$). The most significant predictors were peripheral arthritis and psoriasis (OR 19.4, $P < 0.001$; and OR 10.7, $P < 0.001$, respectively) (supplementary Table S3, available as supplementary data at *Rheumatology* Online). In the bivariate analysis the presence of IBD and uveitis did not reach statistical significance for the diagnosis or for ASAS classification.

Discussion

Our study compares the performance of three sets of criteria for SpA in a clinical setting of early SpA. The early SpA clinics participating in the ESPERANZA programme took into account suspected SpA, together with IBP and asymmetrical arthritis. This approach differed from that

Fig. 1 Flow-chart of the patients included in the study.



Non-SpA: patients not diagnosed with SpA; unlikely SpA, patients with no definitive diagnosis; SpA: patients diagnosed with SpA. The final diagnosis was the clinical diagnosis made by the unit rheumatologist, except in those patients in whom a diagnosis was not available; in these patients, the clinical judgment of the SpA experts was applied.

adopted in previous validation settings [16, 17]. The results show a rather low sensitivity of all three sets of criteria, including the recent ASAS criteria, and good specificity. Although PPV in this study (elevated prevalence, such as this programme, of SpA) is high for all criteria, a negative classification by any of the criteria sets does not rule out a diagnosis. Nevertheless, the fact the rheumatologists in the programme made a clinical judgment of SpA in negative patients means that the criteria can be improved.

Curiously, the weight of those variables with better predictability of a diagnosis by an expert rheumatologist differs little from the weight of the variables that better predict a classification based on the ASAS criteria. In both cases the variables were similar, with the exception of male gender and elevated ESR (both of which were present in the model for diagnosis by a rheumatologist but not in the ASAS classification model) and family history of SpA (present in the ASAS classification model but not in the model with diagnosis by a rheumatologist as the dependent variable). Although this observation may reflect slightly the different results when making a diagnosis and classifying, it cannot explain the low sensitivity of the ASAS criteria. Diagnosing SpA in its early stages is difficult, since clinical presentation is very heterogeneous. No single features (clinical, laboratory or imaging) can distinguish SpA from other rheumatic diseases, therefore, for the diagnosis to be confirmed in daily practice, it is necessary to combine assessment of symptoms, physical examination and imaging and laboratory analysis. According to Khan [29], diagnosis of the early stages of SpA may often be related more to the patient's clinical presentation and the clinician's personal experience and intuition than to precise diagnostic criteria [30].

Lower sensitivity values for the ASAS criteria were expected in early SpA populations, as disease probability was based on the LR of each parameter included in the criteria. These parameters were obtained from populations with well-established disease, which in many cases was diagnosed as AS [9]. In a cohort of <1 year of progress, it is reasonable to expect that patients have fewer clinical features typical of SpA already present from the beginning. This explains the low percentages of patients with uveitis, IBD, dactylitis or urethritis.

In our study, the ASAS criteria performed better than the ESSG and Amor criteria for sensitivity, specificity and predictive ability. However, even though the ASAS criteria were designed to identify patients with early SpA, 35% of patients with early SpA remained unclassified. Sensitivity is quite low for three sets of criteria, with the result that 17% of patients cannot be classified as SpA by any of them. Consequently the ASAS criteria may not

Table 1 Characteristics of the patients included in the study ($n = 775$)

Variable	SpA ($n = 538$)	Non-SpA ($n = 182$)	Unlikely ($n = 55$)
Age, mean (S.D.), years	33 (7)	33 (7)	34 (7)
Caucasians, *n* (%)	462 (86)	168 (92)	46 (84)
Men, *n* (%)	328 (61)	73 (40)	28 (51)
Time from onset of symptoms, mean (S.D.), months	12.1 (6.8)	10.9 (6.0)	12.4 (6.5)
Chronic low back pain, *n* (%)	403 (75)	143 (79)	48 (87)
Inflammatory low back pain, *n* (%)	363 (67)	78 (43)	45 (82)
Pain in buttocks, *n* (%)	164 (30)	30 (16)	9 (16)
Asymmetrical oligoarthritis, *n* (%)	108 (20)	8 (4)	0 (0)
Peripheral arthritis, *n* (%)	97 (18)	8 (4)	2 (4)
Heel pain, *n* (%)	123 (23)	14 (8)	6 (11)
Other enthesitis, *n* (%)	36 (7)	4 (2)	2 (4)
Dactylitis, *n* (%)	43 (8)	1 (1)	0 (0)
Psoriasis, *n* (%)	75 (14)	9 (5)	2 (4)
IBD, *n* (%)	27 (5)	5 (3)	3 (5)
Anterior uveitis, *n* (%)	28 (5)	5 (3)	0 (0)
Diarrhoea in the previous month, *n* (%)	10 (2)	0 (0)	0 (0)
Urethritis/cervicitis, *n* (%)	12 (3)	0 (0)	0 (0)
Family history of SpA, *n* (%)	166 (31)	30 (16)	14 (25)
HLA-B27 positive, *n* (%)	299 (56)	37 (20)	2 (4)
Sacroiliitis in imaging (X-rays or MRI), *n* (%)	199 (37)	4 (2)	0 (0)
Sacroiliitis in X-rays, *n* (%)	101 (19)	0 (0)	0 (0)
Sacroiliitis in MRI, *n* (%) ($N = 359$)	131 (24)	4 (2)	0 (0)
Good response to NSAIDs, *n* (%)	342 (64)	75 (41)	30 (55)
Elevated CRP, *n* (%)	129 (24)	16 (9)	30 (55)
Elevated ESR, *n* (%)	112 (21)	14 (8)	0 (0)

Table 2 Performance of different sets of criteria for early SpA with mixed population (axial and peripheral arthritis)

Criteria	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	LR positive (95% CI)	LR negative (95% CI)
ASAS axial[a]	68 (63, 73)	95 (91, 98)	97 (94, 98)	58 (53, 64)	13 (7.1, 23.8)	0.3 (0.3, 0.4)
ASAS axial imaging	43 (38, 48)	98 (95, 99)	98 (94, 99)	45 (40, 50)	20.4 (7.7, 54.1)	0.6 (0.5, 0.6)
ASAS axial clinical	50 (45, 55)	96 (93, 99)	97 (93, 99)	48 (43, 53)	13.5 (6.5, 28.2)	0.5 (0.5, 0.6)
ASAS peripheral[b]	56 (48, 65)	85 (71, 94)	92 (83, 97)	40 (30, 50)	3.7 (1.8, 7.4)	0.5 (0.4, 0.6)
ASAS total	65 (61, 69)	93 (89, 96)	95 (93, 97)	54 (49, 59)	9.1 (5.7, 14.4)	0.4 (0.3, 0.4)
ESSG	58 (54, 62)	90 (86, 94)	93 (90, 96)	49 (44, 53)	6.0 (4.0, 8.8)	0.5 (0.4, 0.5)
Amor	59 (55, 63)	86 (81, 90)	90 (87, 93)	48 (43, 53)	4.1 (3.0, 5.7)	0.5 (0.4, 0.5)

[a]For the analysis of the axial ASAS criteria we considered only those patients with chronic low back pain ($n = 403$). [b]For the analysis of the peripheral ASAS criteria we considered only those patients with peripheral arthritis, dactylitis or enthesitis in the absence of chronic axial pain ($n = 135$).

be suitable for use as diagnostic criteria in this population of early SpA (Fig. 2). However, the ASAS criteria could be considered as a classification tool for axial forms.

Exclusion of patients for whom HLA-B27 or imaging data were not available could have led to selection bias (a possibility that cannot be ruled out completely) and is therefore a limitation of the study. Nevertheless, although differences were found between excluded and included patients, they are understandable, and there is no clear reason to assume non-comparability or selection bias (supplementary TableS2, available at *Rheumatology* Online).

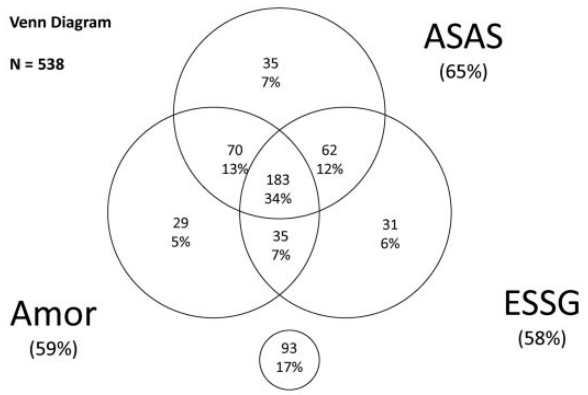
The specificity values of the ASAS criteria are satisfactory and higher than those reported in the study where they were validated [17], except for axial imaging. In contrast, all sensitivities found in the present study are lower than those reported in the original cohort, as well as in patients with long-standing disease of the COchin SPondyloArthritis (COSPA) cohort [19]. This finding can be explained by the low prevalence of HLA-B27 and the results of imaging tests. In the ESPERANZA population, the prevalence of HLA-B27 positivity was 55.6% [well below that of the ASAS and COSPA populations (65.9% and 77%, respectively)], which is consistent with recently

TABLE 3 Results of the bivariate and multivariate logistic regression analysis of parameters that contribute to the diagnosis of SpA

Variables	Bivariate OR (95% CI)	Multivariate OR (95% CI)
Male gender	2.1 (1.5, 2.9)**	1.4 (0.9, 2.2)
Inflammatory low back pain	2.9 (2.1, 4.1)**	3.8 (2.4, 6.0)**
Heel pain	3.2 (1.9, 5.3)**	2.3 (1.2, 4.4)*
Other enthesitis	2.1 (1.5, 2.9)**	—
Arthritis	5.3 (3.3, 8.6)**	7.2 (3.9, 13.5)**
Psoriasis	3.3 (1.7, 6.4)*	4.7 (2.1, 10.5)**
Family history of SpA	1.9 (1.3, 2.8)**	—
Uveitis	2.5 (0.9, 6.7)	—
IBD	1.5 (0.7, 3.4)	—
Good response to NSAIDs	2.2 (1.6, 3.0)**	2.3 (1.5, 3.6)**
Elevated CRP	3.6 (2.2, 6.1)*	2.8 (1.4, 5.5)*
HLA-B27 positive	6.4 (4.3, 9.3)**	5.9 (3.6, 9.5)**
Sacroiliitis on MRI/X-ray	34.2 (12.5, 93.3)**	28.3 (9.8, 81.5)**

*$P < 0.01$, **$P < 0.001$.

FIG. 2 Venn diagram representing the number of patients from the ESPERANZA study fulfilling the different criteria for SpA.



Amor: Amor criteria.

published results of the emAR II study on the characteristics of patients with SpA followed in Spanish rheumatology units (58.8%), whose mean time with the disease was 105 months [31]. The prevalence of the HLA-B27 not only affects the percentage of patients in the clinical arm, but also indirectly affects the imaging arm in the classification algorithm. As was recently demonstrated in the DESIR (Devenir des Spondyloarthropathies Indifférenciées Récentes) cohort, the presence of HLA-B27 is independently associated with radiological damage and inflammation detected on MRI [32, 33]. On the other hand, differences were found in the percentage of patients with radiological sacroiliitis (19%), an observation that is clearly associated with differences in the duration of disease progression from other studies (29.7% and 82.3%) [19, 32]. Finally, in a population with very early disease, it is expected that structural lesions will be limited and inflammatory lesions detected by MRI will predominate; however, the low number of MRI scans performed (36.3%) in this population may limit the sensitivity of the ASAS criteria in our study.

As for the peripheral ASAS criteria, the lower sensitivity detected has a plausible explanation in our population, namely, the low presence of enthesitis. There is a clear difference in the percentage of patients with enthesitis between the ESPERANZA study (27.3%) and the original ASAS validation study (56.8%) [17].

In conclusion, in this cohort of early SpA patients, the ASAS criteria performed slightly better than those of the ESSG and Amor, as their sensitivity and specificity are superior. However, the sensitivity of the ASAS criteria in a population with early axial and peripheral forms is poorer than in previous studies, therefore the ability of the ASAS criteria to detect early forms is limited in daily clinical practice, especially in populations where MRI is not readily available or the prevalence of HLA-B27 is low.

Rheumatology key messages

- The sensitivity of all classification criteria is low in very early SpA.
- The new ASAS criteria perform slightly better than other criteria in early SpA.
- HLA-B27 positivity, arthritis and imaging tests are determinant for a diagnosis in early SpA.

Acknowledgements

The scientific committee wants to acknowledge the support of Drs Juan C. López-Robledillo and Javier Rivera, and the excellent technological support by SER's programmers Francisco Javier Quesada and Juan Manuel Barrio.

Funding: This work was supported by Pfizer. Pfizer sponsored the Spanish Foundation of Rheumatology to run the ESPERANZA programme. The sponsor had no role in the

study design, collection, analysis or interpretation of data, the writing of the report or in the decision to submit the article for publication. The researchers were independent from the funders. This work was partially supported by the RETICS programme, RD08/0075 (RIER) from Instituto de Salud Carlos III (ISCIII), within the VI PN I+D+I 2008–2011 (FEDER).

Disclosure statement: M.G. works for the Spanish Society of Rheumatology and develops studies supported by different laboratories (Pfizer, Lilly, Grünental, Roche, MSD, Abbot). E.D.M. has received consultancy fees from Pfizer, Abbvie and Abbott, grants and/or research support from Abbott and Pfizer and payment for lectures including service on speakers bureaus from Abbvie, Pfizer, Menarini, BMS, MSD, Schering-Plough, Roche, UCB, Wyeth and Abbott. All other authors have declared no conflicts of interest.

Supplementary data

Supplementary data are available at *Rheumatology* Online.

References

1 Amor B. Usefulness of criteria for spondyloarthropathies. Joint Bone Spine 2000;67:502–3.

2 Amor B, Dougados M, Mijiyawa M. [Criteria of the classification of spondylarthropathies]. Rev Rhum Mal Osteoartic 1990;57:85–9.

3 Dougados M, Hochberg MC. Why is the concept of spondyloarthropathies important? Best Pract Res Clin Rheumatol 2002;16:495–505.

4 Dougados M, van der Linden S, Juhlin R *et al*. The European Spondylarthropathy Study Group preliminary criteria for the classification of spondylarthropathy. Arthritis Rheum 1991;34:1218–27.

5 Collantes E, Veroz R, Escudero A *et al*. Can some cases of 'possible' spondyloarthropathy be classified as 'definite' or 'undifferentiated' spondyloarthropathy? Value of criteria for spondyloarthropathies. Spanish Spondyloarthropathy Study Group. Joint Bone Spine 2000;67:516–20.

6 Collantes-Estevez E, Cisnal del Mazo A, Munoz-Gomariz E. Assessment of 2 systems of spondyloarthropathy diagnostic and classification criteria (Amor and ESSG) by a Spanish multicenter study. European Spondyloarthropathy Study Group. J Rheumatol 1995;22:246–51.

7 Rojas-Vargas M, Munoz-Gomariz E, Escudero A *et al*. First signs and symptoms of spondyloarthritis—data from an inception cohort with a disease course of two years or less (REGISPONSER-Early). Rheumatology 2009;48:404–9.

8 Rudwaleit M, Haibel H, Baraliakos X *et al*. The early disease stage in axial spondylarthritis: results from the German Spondyloarthritis Inception Cohort. Arthritis Rheum 2009;60:717–27.

9 Rudwaleit M, Khan MA, Sieper J. The challenge of diagnosis and classification in early ankylosing spondylitis: do we need new criteria? Arthritis Rheum 2005;52:1000–8.

10 Rudwaleit M, van der Heijde D, Khan MA *et al*. How to diagnose axial spondyloarthritis early. Ann Rheum Dis 2004;63:535–43.

11 Sieper J, Rudwaleit M. Early referral recommendations for ankylosing spondylitis (including pre-radiographic and radiographic forms) in primary care. Ann Rheum Dis 2005;64:659–63.

12 Bennett AN, McGonagle D, O'Connor P *et al*. Severity of baseline magnetic resonance imaging-evident sacroiliitis and HLA-B27 status in early inflammatory back pain predict radiographically evident ankylosing spondylitis at eight years. Arthritis Rheum 2008;58:3413–8.

13 Braun J, Baraliakos X. Imaging of axial spondyloarthritis including ankylosing spondylitis. Ann Rheum Dis 2011;70(Suppl 1):i97–103.

14 Maksymowych WP, Chiowchanwisawakit P, Clare T *et al*. Inflammatory lesions of the spine on magnetic resonance imaging predict the development of new syndesmophytes in ankylosing spondylitis: evidence of a relationship between inflammation and new bone formation. Arthritis Rheum 2009;60:93–102.

15 Weber U, Lambert RG, Ostergaard M *et al*. The diagnostic utility of magnetic resonance imaging in spondylarthritis: an international multicenter evaluation of one hundred eighty-seven subjects. Arthritis Rheum 2010;62:3048–58.

16 Rudwaleit M, Landewe R, van der Heijde D *et al*. The development of Assessment of SpondyloArthritis International Society classification criteria for axial spondyloarthritis (part I): classification of paper patients by expert opinion including uncertainty appraisal. Ann Rheum Dis 2009;68:770–6.

17 Rudwaleit M, van der Heijde D, Landewe R *et al*. The development of Assessment of SpondyloArthritis International Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 2009;68:777–83.

18 van den Berg R, van Gaalen F, van der Helm-van Mil A *et al*. Performance of classification criteria for peripheral spondyloarthritis and psoriatic arthritis in the Leiden early arthritis cohort. Ann Rheum Dis 2012;71:1366–9.

19 Cheung P, Paternote S, Burki V *et al*. Performance of the assessment in spondyloarthritis international society classification for axial and peripheral spondyloarthritis in an established clinical cohort: comparison criteria sets of Amor and European spondyloarthropathy study group. J Rheumatol 2012;39:816–3.

20 Munoz-Fernandez S, Carmona L, Collantes E *et al*. A model for the development and implementation of a national plan for the optimal management of early spondyloarthritis: the ESPERANZA program. Ann Rheum Dis 2011;70:827–30.

21 Calin A, Jones SD, Garrett SL *et al*. Bath Ankylosing Spondylitis Functional Index. Br J Rheumatol 1995;34:793–4.

22 Garrett S, Jenkinson T, Kennedy LG *et al*. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol 1994;21:2286–91.

23 Heuft-Dorenbosch L, Spoorenberg A, van Tubergen A *et al*. Assessment of enthesitis in ankylosing spondylitis. Ann Rheum Dis 2003;62:127–32.

24 Jenkinson TR, Mallorie PA, Whitelock HC *et al*. Defining spinal mobility in ankylosing spondylitis (AS). The Bath AS Metrology Index. J Rheumatol 1994;21: 1694–8.

25 Jones SD, Porter J, Garrett SL *et al*. A new scoring system for the Bath Ankylosing Spondylitis Metrology Index (BASMI). J Rheumatol 1995;22:1609.

26 Sieper J, Rudwaleit M, Baraliakos X *et al*. The Assessment of SpondyloArthritis International Society (ASAS) handbook: a guide to assess spondyloarthritis. Ann Rheum Dis 2009;68(Suppl 2):ii1–44.

27 van der Heijde D, Landewe R, Feldtkeller E. Proposal of a linear definition of the Bath Ankylosing Spondylitis Metrology Index (BASMI) and comparison with the 2-step and 10-step definitions. Ann Rheum Dis 2008;67: 489–93.

28 Rudwaleit M, Jurik AG, Hermann KG *et al*. Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. Ann Rheum Dis 2009; 68:1520–7.

29 Khan MA. Ankylosing spondylitis: introductory comments on its diagnosis and treatment. Ann Rheum Dis 2002; 61(Suppl III):iii3–7.

30 van Tubergen A, Weber U. Diagnosis and classification in spondyloarthritis: identifying a chameleon. Nat Rev Rheumatol 2012;8:107–13.

31 Casals-Sánchez JL, Garcia de Yébenes MJ, Descalzo MA *et al*. Characteristic of patients with spondyloarthritis followed in rheumatology units in Spain. emAR II study. Reumatol Clin 2012;8:107–13.

32 Dougados M, d'Agostino MA, Benessiano J *et al*. The DESIR cohort: a 10-year follow-up of early inflammatory back pain in France: study design and baseline characteristics of the 708 recruited patients. Joint Bone Spine 2011;78:598–603.

33 Chung HY, Machado P, van der Heijde D *et al*. HLA-B27 positive patients differ from HLA-B27 negative patients in clinical presentation and imaging: results from the DESIR cohort of patients with recent onset axial spondyloarthritis. Ann Rheum Dis 2011;70:1930–6.