

COMMENTARY

The Challenge of Diagnosis and Classification in Early Ankylosing Spondylitis

Do We Need New Criteria?

Martin Rudwaleit,¹ Muhammad A. Khan,² and Joachim Sieper¹

Introduction

Ankylosing spondylitis (AS) is a common chronic inflammatory disease with an estimated prevalence of 0.2–1.2% (1–4). The disease typically starts during the third decade of life and has a substantial socioeconomic impact on the patient and society (5–7). Until recently, the treatment options for AS were limited. The mainstays of treatment were regular physical therapy and nonsteroidal antiinflammatory drugs (NSAIDs) (8). In contrast, disease-modifying antirheumatic drugs (DMARDs) as well as corticosteroids, which are quite effective in some of the other chronic inflammatory diseases such as rheumatoid arthritis (RA), show only very limited or no efficacy in AS (9–11). Thus, in the past, a delayed diagnosis did not have much of an adverse consequence because of the lack of highly effective therapeutic choices.

Most recently, it has been convincingly demonstrated that the tumor necrosis factor α (TNF α)-blocking agents infliximab and etanercept have a strong and prompt effect on almost all features of AS, such as clinical disease activity, physical function, spinal mobility, peripheral arthritis, enthesitis, and levels of acute-phase reactants (12–19). In several studies of AS patients whose disease was refractory to NSAIDs and physical therapy, ~50% of the patients have demon-

strated at least a 50% improvement when treated with either of the two TNF α -blocking compounds. It has also been shown that active juxtaarticular bony inflammation (“bone edema”), as detected by magnetic resonance imaging (MRI), can be suppressed (20,21), and it is hoped that this kind of treatment will also favorably influence long-term outcome, including reduction or prevention of radiologic progression. Recent data also show that AS patients with a short disease duration and good functional status are more likely to respond to TNF α -blocking agents than patients with longstanding disease and impaired function (22). Thus, an early and reliable diagnosis of AS has now become an important and very relevant issue.

Evolution of criteria for AS

The Rome criteria (23) from 1961 were the first set of criteria developed for the classification of AS, and on their subsequent evaluation (24), 2 items from these criteria were deleted: thoracic pain (because of its low specificity) and uveitis (because of its low sensitivity). This resulted in the New York classification criteria (25) in 1966. In 1977, Calin et al had suggested criteria for chronic inflammatory back pain (IBP) to help differentiate IBP from other causes of chronic back pain (26). This led in 1984 to the modified New York criteria for the classification of AS (27). The modified New York criteria incorporate the IBP components in place of the rather nonspecific clinical symptom of chronic low back pain that was used in both the Rome and the New York criteria. A patient can be classified as having definite AS if at least 1 clinical criterion (inflammatory back pain, limitation of mobility of the lumbar spine, or limitation of chest expansion) *plus* the radiologic criterion (radiographic sacroiliitis of grade 2 bilaterally or grade 3–4 unilaterally) are fulfilled. These modified criteria are cur-

Supported by grant FKZ 01 GI 9946 from the BMBF (Kompetenznetz Rheuma).

¹Martin Rudwaleit, MD, Joachim Sieper, MD: Charité, Campus Benjamin Franklin, Berlin, Germany; ²Muhammad A. Khan, MD: MetroHealth Medical Center, Case Western Reserve University, Cleveland, Ohio.

Address correspondence and reprint requests to Martin Rudwaleit, MD, Rheumatologie, Medizinische Klinik I, Charité, Campus Benjamin Franklin, Hindenburgdamm 30, 12200 Berlin, Germany. E-mail: martin.rudwaleit@charite.de.

Submitted for publication September 16, 2004; accepted in revised form January 5, 2005.

rently the most widely used classification criteria, and clinicians are also using it for the clinical diagnosis of AS.

The early disease stage of AS

Over the last decades, it has become increasingly evident that in many patients with AS, it takes years from the onset of IBP until the appearance of radiographic sacroiliitis. For example, in a study of 88 patients with clinical features compatible with early AS (IBP plus additional features such as peripheral arthritis, heel pain, acute uveitis, or elevated levels of acute-phase reactants) but with radiographically normal sacroiliac joints, 36% had developed radiographic evidence of sacroiliitis after 5 years and 59% had done so after 10 years (28).

Thus, time (duration of symptoms/disease duration) is an important factor in determining the presence or absence of radiographic sacroiliitis in predisposed individuals. This was also shown in a study of HLA-B27-positive relatives of patients with AS (1), where radiographic sacroiliitis was noted in 16% of patients younger than 45 years and increased to 38% of patients older than 45 years. In another family study, radiographic evidence of sacroiliitis was found in 40% of patients with a symptom duration of <10 years, 70% with symptoms for 10–19 years, and 86% with symptoms for ≥ 20 years (29).

The absence of radiographic sacroiliitis during the early stage of disease must certainly not imply that there is no inflammation in the sacroiliac joints and/or other parts of the axial skeleton. Recent application of MRI techniques has demonstrated (and confirmed) that ongoing active (“acute”) inflammation in fact does occur in the sacroiliac joints and/or spine prior to the appearance of changes detectable radiographically. In a prospective study of 9 patients with normal or equivocal findings on radiographs in the presence of active inflammation on MRI at baseline, 6 developed definite radiographic sacroiliitis (at least grade 2 bilaterally) after 3 years of followup (30). MRI seems to represent the most recent milestone in diagnostic imaging of the preradiographic phase of AS (30–32) (Figure 1), although a clear definition of positive and negative findings on MRI is needed, along additional data on the sensitivity and specificity of MRI findings in patients with early disease. An international effort to address these questions is under way.

Burden of disease in early AS

Remarkably, the presence or absence of radiographic sacroiliitis does not determine the burden of

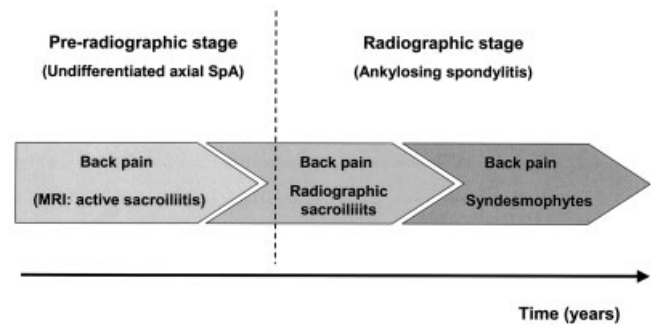


Figure 1. Unifying concept of axial spondylarthritis (SpA), showing schematically the transition from early to late axial SpA. Axial SpA usually starts without sacroiliitis detectable by conventional radiography, but with time (over years), sacroiliitis and possibly syndesmophytes will become detectable by conventional radiography. Back pain as the leading symptom may be present throughout the disease course. During the early disease stage, magnetic resonance imaging (MRI) may detect acute inflammatory lesions in the absence of radiographic sacroiliitis. The vertical dotted line indicates the borderline that is used, according to established criteria, to separate axial undifferentiated SpA without radiographic changes from ankylosing spondylitis.

disease during the early stage of AS. A study of the German Spondyloarthritis Inception Cohort (GESPIC) has recently shown that patients with early disease without radiographic sacroiliitis (undifferentiated spondylarthritis [SpA] with axial involvement) do not differ in this regard from patients with definite AS (with radiographic sacroiliitis) of short duration (<10 years) with respect to disease activity (as evaluated by the Bath Ankylosing Spondylitis Disease Activity Index), level of global pain, level of pain at night, patient’s global assessment of disease activity, need for treatment, response to treatment, and quality of life (33). In other words, the presence or absence of changes detected by conventional radiography does not primarily determine the impact of the disease on daily life.

Radiographic and preradiographic AS, a single disease entity

Given the compelling amount of data suggesting that the occurrence of radiographic sacroiliitis in patients with axial SpA is mainly a function of time, with some influence of severity factors, the presence and absence of radiographic sacroiliitis in patients with SpA represent different stages of a single disease continuum and, therefore, the same disease entity. Radiographic sacroiliitis is often followed at a later stage by the formation of syndesmophytes (Figure 1). Thus, the presence of radiographic changes should be seen as a

marker of chronicity and/or severity, rather than as an essential diagnostic criterion. Conceptually, radiographic sacroiliitis is one of the consequences of an inflammatory process and does not, by itself, reflect the ongoing active inflammation.

We therefore propose that SpA patients with predominantly axial symptoms should be considered as having the same disease entity as AS patients, independently of the presence of radiographic sacroiliitis. The concept of considering patients with and patients without radiographic sacroiliitis as having a single disease entity makes it also less important to distinguish between grade 1 and grade 2 radiographic sacroiliitis (see below). Such a concept is reminiscent of that in early RA, in which bony changes (erosions) detected radiographically are not regarded as an essential feature for early diagnosis or classification (34), although they will develop in the majority of untreated RA patients over time. Moreover, in early RA, MRI may often detect evidence of inflammation and even bony erosions that are not detectable by conventional radiography, and the presence and extent of bony erosions are regarded as an indicator of disease severity, rather than an essential diagnostic feature (35). As discussed above, radiographic changes probably occur even later in the disease course of AS than in the disease course of RA.

Limitations of the currently used criteria for AS

The Rome criteria (23), New York criteria (25), and the modified New York criteria (27) for the classification of AS include the presence of spinal/thoracic pain, restriction of spinal mobility, and radiologic sacroiliitis. While the latter 2 components mainly reflect damage (irreversible changes), the first component (spinal/thoracic pain) may reflect both an ongoing active inflammation and the resultant damage. It is self-evident that such criteria will perform less well if applied to the diagnosis of patients with early AS in whom radiographically detectable damage has yet not occurred.

Problems with the presence of radiographic sacroiliitis as a requirement for the diagnosis of AS. The requirement of radiographic sacroiliitis for AS in particular results in a low sensitivity for the criteria if applied to patients with early disease because patients who present with clinical symptoms but without radiographic sacroiliitis will not be recognized as having AS (28,36,37). This circumstance was already taken into account at the time of the development of the Rome criteria because, in the absence of radiographic sacroiliitis, the presence of 4 of the 5 clinical criteria was

regarded as being sufficient to make a classification of probable AS (23). The modified New York criteria also allow for such a classification (probable disease) if *all 3* clinical criteria are present (27). But, as discussed below, restriction of spinal mobility and restriction of chest expansion are absent in a very large proportion of AS patients during the early course of disease. Therefore, the absence of both radiographic sacroiliitis and impaired spinal mobility at the patient's first presentation to a physician contributes to the long delay (5–10 years) in the ultimate diagnosis of AS in many patients (7).

Difficulties in grading radiographic sacroiliitis.

There is also some concern about the specificity of radiographic sacroiliitis, especially grade 2 sacroiliitis, which is regarded in all criteria as being sufficient for the classification of definite AS. As discussed back in 1966, it is difficult to differentiate between grades 1 and 2, which currently constitute the borderline between "disease" and "no disease," and it was estimated that in ~20% of cases, designations of grades 1 and 2 might be wrong (24). Nevertheless, the presence of grade 2 sacroiliitis bilaterally by radiography remained part of these criteria for 2 reasons: the sensitivity would have dropped considerably if only patients with grade 3 sacroiliitis could have been included, and there was no better imaging procedure available at that time for early diagnosis.

This issue was investigated again more recently in a Dutch study (38), which reported that the specificity and sensitivity of sacroiliitis on radiographs as read by trained radiologists or rheumatologists are only ~80% each, resulting in a relatively low likelihood ratio (LR) of ~4 ($LR = \text{sensitivity}/1 - \text{specificity}$) for the presence of AS. Although no differentiation between grade 2 and higher grades of sacroiliitis (grade 3 or 4) was made in that study, one can assume that major problems must have arisen from difficulties in differentiating between grade 1 and grade 2 sacroiliitis, as discussed above. Thus, requiring the presence of grade 3 or 4 sacroiliitis would greatly increase the specificity of the radiologic criterion and, thus, would be preferable, provided that the accompanying loss of sensitivity is compensated for by other means.

Restriction of spinal mobility and chest expansion. Similar to the radiologic criterion, the 2 clinical criteria "restriction of spinal mobility" and "restriction of chest expansion" are likely to perform reasonably well in advanced disease (the modified New York criteria for AS were evaluated in patients with established/advanced disease) but seem to be less useful for application in early disease. In a Dutch study from 1985 (39), limitation

of lumbar spinal mobility in “all 4 planes” was found in 45% of AS patients with a symptom duration of 0–4 years, while reduced chest expansion (<3 cm) was present in only 24%.

We analyzed spinal mobility in 190 AS patients from the GESPIC who had disease symptoms for <10 years (40). We observed that only 34% of the patients had limitation of lumbar spine mobility in all 4 planes and that only 18% had chest expansion of <3 cm. In another group of 76 patients from the GESPIC who had axial undifferentiated SpA (without radiographic sacroiliitis; duration of symptoms <5 years), only 14% showed limitation of mobility of the lumbar spine in both the frontal and sagittal planes, and only 1.3% showed reduced chest expansion (40).

Thus, restricted spinal mobility and restricted chest expansion reflect disease duration (chronicity) and/or severity, and are not *sensitive* enough to be used in criteria sets that aim to encompass not only advanced, but also early, cases of AS. Moreover, and of equal importance, the *specificity* of the criterion “restriction of spinal mobility” ranges only between 37% and 75%, as several studies have revealed when patients with mechanical back pain were used as a comparator group (41–43).

The challenge of making a diagnosis in early disease

The physician largely depends on the presence of radiographic sacroiliitis to diagnose AS in a patient with axial SpA, and as discussed above, radiographic sacroiliitis is mainly a function of disease duration and is influenced by severity factors. In daily practice, patients who have chronic IBP and radiographic sacroiliitis at the time of presentation to the physician will be diagnosed as having AS, whereas those who have chronic IBP but do not have radiographic sacroiliitis at first presentation may or may not be diagnosed as having preradiographic axial SpA. This situation is depicted with hypothetical frequencies in Figure 2, where group A represents patients with radiographic sacroiliitis at first presentation and groups B, C, and D represent those without radiographic sacroiliitis at first presentation. The majority of patients in groups B and C will develop radiographic sacroiliitis and possibly syndesmophytes with time, and only a small proportion of patients in group D may never develop radiographic sacroiliitis despite having IBP for many years.

In a patient with early AS at the preradiographic stage, making the diagnosis is particularly challenging for the following 3 reasons. First, chronic back pain as

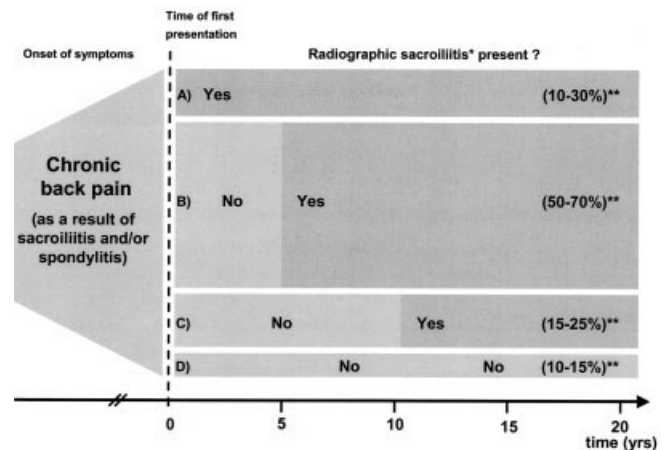


Figure 2. Variation in the development of radiographic sacroiliitis in patients with axial spondylarthritis (SpA) (hypothetical frequencies). The initial symptom in most patients is back pain, which after a certain time, may lead the patient to seek medical help. Depending mostly on the duration of symptoms and the time at which medical help is sought, radiographic sacroiliitis may be present. After a very short period of symptoms at first presentation, only a few patients will have radiographic sacroiliitis (group A). In the majority of patients, radiographic sacroiliitis will be present after 5–10 years of symptoms (groups B and C). A further small percentage of patients will never develop radiographic changes (group D). Asterisk indicates presence or absence of syndesmophytes. Double asterisk indicates the estimated proportion of patients of the total patient population. The percentages of patients shown at the 4 time points are hypothetical estimates and need to be investigated in larger studies.

the leading symptom is very common in the general population, whereas AS (preradiographic and radiographic) accounts for not more than 5% of all patients with chronic back pain (44). Second, the type of back pain that is typical of early AS (i.e., IBP) is present in 70–80% of AS patients, but it is also present in 20–25% of patients with “mechanical” back pain (26,43,45). Thus, given only moderate values for sensitivity and specificity, the presence of IBP alone does not suffice for making the diagnosis. Third, there are, up to now, no widely accepted diagnostic guidelines for early, preradiographic AS.

Diagnosis versus classification

Can we use classification criteria for diagnostic purposes? In the absence of diagnostic criteria, classification criteria (46) are often used to aid the diagnostic process in daily practice. Obviously, in the field of SpA as well as other rheumatic diseases, the same or similar clinical, laboratory, or imaging parameters are used—and appropriately—for both diagnostic and classifica-

Table 1. Diagnostic versus classification criteria

Diagnostic criteria	Classification criteria
Used by a physician to make a diagnosis	Applied to patients in whom the diagnosis has already been made
When making the diagnosis, the value of diagnostic tests/parameters depends on the prevalence of the disease (pretest probability)	Prevalence of the disease is not important, since all patients should have the disease (have been previously diagnosed)
The purpose of diagnostic criteria/algorithms is to help diagnose individual patients	The purpose of classification criteria is to provide a unique language for researchers to evaluate homogeneous groups of patients, which facilitates comparisons of clinical or experimental studies
Criteria for diagnosis should have a high <i>sensitivity</i> in order to identify as many patients with the disease as possible	Criteria for classification should have a high <i>specificity</i> (close to 100%) in order to avoid misclassification (inclusion of patients who do not have the disease)
Should allow for flexibility in diagnostic confidence (definite, probable, possible)	Gives a yes or no answer (criteria fulfilled or not fulfilled)
Applies to the individual patient	Applies to groups of patients

tion purposes. However, the way these parameters are incorporated differs between diagnosis and classification. When making a diagnosis, the clinical value of any diagnostic test or criterion is highly dependent on the pretest probability or, in the case of an individual patient, the prevalence of the disease (Bayes' theorem) (47). This is in contrast to the application of classification criteria, which are meant to be applied only in patients who have already been diagnosed so as to ascertain the validity or certainty of their classification as having the disease under study (Table 1). As a consequence, the application of classification criteria as diagnostic criteria can easily represent misuse of the criteria, as has been demonstrated in the field of systemic vasculitis (46,48,49). Making a diagnosis because a patient fulfills certain classification criteria runs the risk of misdiagnosis, especially if applied in a population with a low pretest probability for the particular disease.

Furthermore, the application of classification criteria usually yields a yes (criteria fulfilled) or no (criteria not fulfilled) answer. In contrast, only a yes or no option for diagnostic purposes would not be acceptable in clinical care, where the diagnostic process should also provide the degree of certainty, which is often categorized as definite, probable, or possible presence of the disease (Table 1). Making a diagnosis of probable or possible SpA reflects a lower level of clinical confidence

about the presence of the disease and is not an uncommon situation in clinical practice.

A novel approach to early diagnosis of AS. We recently published an approach (diagnostic algorithm) to the early diagnosis of axial SpA, particularly at the preradiographic stage, to be applied by the rheumatologist and experienced physician (50). In our diagnostic algorithm, we took into account the different diagnostic weights of some of the clinical, laboratory, and imaging parameters. According to these calculations, a disease probability of $\geq 90\%$ will be reached in a patient if, for example, IBP plus 3–4 further features of SpA (clinical, laboratory, or imaging) are present (50). The easiest way to calculate the disease probability in an individual patient would be to multiply the individual likelihood ratios of all SpA features that are identified in the patient (see Appendix A). The LR combines into 1 value the sensitivity and specificity of a diagnostic test, and it is an indicator of the diagnostic value of the respective test: the higher the LR, the better the diagnostic value of the test (refs. 51 and 52 and Appendix A).

Given a prevalence of 5% for axial SpA (AS and preradiographic AS) among patients with chronic back pain (44), an LR product of ~ 200 will give at least an $\sim 90\%$ probability of disease. A disease probability of $\sim 80\%$ is achieved with an LR product of ~ 80 , and a disease probability of $\sim 50\%$ is achieved with an LR product of ~ 20 (Figure 3). It can easily be seen from Figure 3 that 3–4 parameters need to be positive for a disease probability of 90%. The disease probability drops to 80% or 50% if 3 or only 2 parameters, respectively, are present (see Appendix A as well as ref. 50 for more details about these calculations). We suggest here that grade 3 radiographic sacroiliitis (unilateral or bilateral) be added as a further parameter (in addition to those we previously proposed) for use in the diagnosis of early, preradiographic AS (50). This would enable one to use this diagnostic approach for the whole group of patients with axial SpA (preradiographic and radiographic AS). For grade 3 sacroiliitis, we assumed a sensitivity of 40% for the total group of radiographic plus preradiographic axial SpA, with a specificity of 98%. This results in a positive LR of 20 for grade 3 sacroiliitis.

The sensitivities and specificities of each of the diagnostic parameters for axial SpA were taken from different studies in different populations and can only be regarded as best estimates until they are confirmed in prospective studies (50). In addition, we have proposed an algorithm for early referral (by less experienced

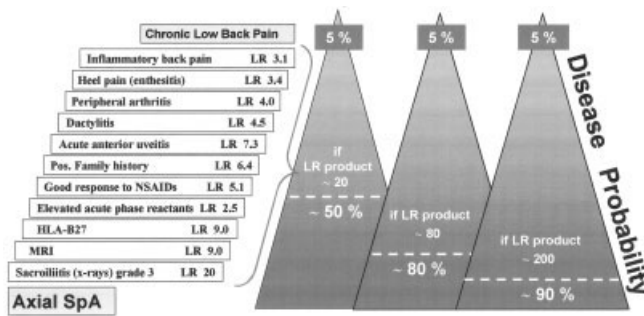


Figure 3. Disease probabilities of the presence of axial spondylarthritis (SpA) according to the presence of individual SpA parameters in individual patients. The prevalence (pretest probability) of having axial SpA among patients with chronic back pain is ~5%. To calculate the disease probability for an individual patient, the likelihood ratios (LRs) of the parameters that are present in the patient are multiplied, resulting in an individual LR product. Thus, the resulting LR product depends on both the number of parameters present and the LR of the parameters present. The 3 pyramids represent 3 possible scenarios where different LR products result in different disease probabilities: if the LR product is ~20, the resulting disease probability will be ~50%, if the LR product is ~80, the disease probability will be ~80%, and if the LR product is ~200, the disease probability will be ~90%. A disease probability of 90% or more is regarded by us as definite disease. (For further details, see text, Appendix A, and ref. 50.) Pos. = positive; NSAIDs = nonsteroidal antiinflammatory drugs; MRI = magnetic resonance imaging.

physicians) for rheumatologic consultation of selected patients with chronic back pain (53).

Classification criteria for SpA. The purpose of classification is to provide a standardization of approaches to understanding the etiology and course of diseases (46,49,54). Classification criteria should serve as a tool for research and communication, providing uniform criteria for the scientific community by which to classify patients with the same disease (homogeneous population of patients). They also aid in the selection of patients for clinical and therapeutic trials, and make the data obtained by different researchers in different cohorts of patients comparable. Thus, classification criteria in particular are required to have a high specificity (as close to 100% as possible) in order to reliably separate patients with the disease from patients without the disease but with similar symptoms or manifestations.

Two sets of criteria for classifying patients with SpA, both with and without radiologic sacroiliitis, were introduced about 15 years ago: the European Spondylarthropathy Study Group (ESSG) criteria (55) and the Amor criteria (56). Both of these criteria sets were very important steps toward a more comprehensive classification of this group of diseases. The criteria were

developed not only for the purpose of classification, but also to encompass a wider clinical spectrum by adding the subgroup of undifferentiated SpA (57) and by stressing the concept of SpA, which was first introduced by Moll et al in 1974 (58). The ESSG criteria focused on the 2 leading symptoms of all types of SpA, the presence of IBP or asymmetric peripheral (oligo)arthritis, and stressed that these symptoms are more important than their assignment to a specific SpA subtype. As a consequence of the ESSG concept, patients with IBP would have the same disease (i.e., SpA) whether they have radiographic sacroiliitis or not.

Is there a need for new classification criteria?

There are several issues that underscore the need for new classification criteria for early AS. First, as outlined above, the modified New York criteria are not applicable to preradiographic AS. Second, the specificity of the available SpA classification criteria is not high enough. This seems to be true particularly for the ESSG criteria. Third, MRI as a new and valuable diagnostic tool is not part of any of the existing sets of criteria. Fourth, classifying axial and peripheral SpA separately may be advantageous over a single set of unifying criteria. Fifth, clinical trials in patients with early disease without radiographic changes are already under way.

Specificity of existing classification criteria for SpA. The lower the specificity of the classification criteria, the higher the proportion of misclassified patients (49,54). This appears to be a major problem with the application of the ESSG criteria to patients with early, preradiographic AS. If only 2 parameters—as required by the ESSG criteria—are present in a patient with chronic back pain, the probability that axial SpA is in fact present in this patient is only ~50%. This view is supported by a recent study from Spain showing that only 53.6% of patients that had initially fulfilled the ESSG criteria were considered to have SpA after 5 years of followup (59). Although this study was small, the results are consistent with our disease probability calculations (50) and with our experience from clinical practice. However, a more comprehensive long-term study of a larger number of patients would be necessary before the long-term accuracy of the ESSG criteria can be appropriately judged.

The Amor criteria (56,60) appear to be superior to the ESSG criteria since they require the presence of at least 3 or 4 SpA parameters for definite classification of axial SpA (59,61). However, with respect to early, preradiographic axial SpA, these criteria also have some

disadvantages. For example, IBP as the leading symptom is not particularly well-defined in the Amor criteria, and the presence of HLA-B27 and a family history of SpA are considered as entirely dependent parameters, which may not be the case. Moreover, MRI of the sacroiliac joints as a more recent diagnostic tool is not part of these criteria.

Separate classification of axial SpA? The unifying concept of considering SpA as a single disease entity, as has been put forward by the ESSG criteria and the Amor criteria, was very helpful in underscoring the shared genetic, pathophysiologic, and clinical components among the different SpA subgroups. However, despite these shared components, the subgroups differ substantially from each other. For example, a 20-year-old HLA-B27-negative patient with reactive arthritis is likely to have self-limiting disease, whereas a 20-year-old HLA-B27-positive patient who already has sacroiliitis and syndesmophytes at presentation is likely to run a highly progressive course of AS. Thus, from a clinical point of view, knowing that a certain patient fulfills the ESSG or the Amor criteria for SpA gives too little information about the clinical manifestations in this individual patient. It is difficult to have even a vague idea about the composition of the patients that take part in the studies that use these criteria. In contrast, knowing that a patient fulfills not only the ESSG or the Amor criteria, but also the modified New York criteria gives a much more precise picture of that patient. Such a precise picture is in fact needed when interpreting therapeutic trials or outcome studies in SpA.

Therefore, having separate criteria for axial SpA (AS and preradiographic AS) and for peripheral SpA is certainly preferable over a unifying set of criteria for the whole group of SpA, since the patient's clinical picture would be more precise than that portrayed by the ESSG and the Amor criteria. The subdivision into preradiographic and radiographic axial SpA adds further relevant clinical information about the clinical status/stage of the patients with axial SpA.

Proposal for new classification criteria

We strongly feel a need for either new or revised classification criteria for axial SpA. Such criteria will comprise all parameters relevant to axial SpA including MRI findings. The new or revised classification criteria should adequately encompass the whole spectrum of axial SpA, i.e., the preradiographic and the radiographic stages. Parameters considered to be relevant to such criteria for axial SpA are those shown in Figure 3, which

are also used for diagnostic purposes. As we have pointed out above, the diagnostic weight is different for these parameters (shown as LRs in Figure 3). Accordingly, it remains to be seen if such a weight should also be assigned to each parameter for classification purposes, as well as how many parameters will be needed for classification. The sensitivity and specificity of new classification criteria for early AS/axial SpA have to be tested against SpA (sensitivity) and non-SpA (specificity) as diagnosed by an expert. Based on the considerations of disease probabilities made above, we assume that at least 3–4 SpA parameters must be present in order to reach a very good specificity of the criteria in order to avoid misclassification. An international effort to validate the superiority of a potential new set of classification criteria over existing ones is clearly needed. The Assessment in Ankylosing Spondylitis (ASAS) International Working Group should be a suitable organization for conducting such studies.

Proposal for new terminology

If we consider that patients with SpA at the preradiographic and the radiographic stages represent a single disease continuum, the term AS would be misleading for some patients, since in essence, this term implies that ankylosis has already taken place. The state of "ankylosis," however, does not apply to early disease, and equally important, it does not even apply to all patients with longstanding disease. Patients with longstanding disease who have not yet developed sacroiliitis that can be detected by conventional radiography and who may run a mild disease course would unnecessarily be stigmatized by the term "ankylosing."

Considering the whole spectrum of patients, the term "axial SpA" seems preferable to us, since it stresses the predominant involvement of the sacroiliac joints and spine but does not automatically imply the presence of radiographic changes or ankylosis. For lack of a better name, one of us previously used the term "spondylitic disease" (36), and for patients and nonspecialist physicians, the term could simply be "inflammatory spine disease." Whatever term is chosen, it could (and probably should) be further split into a preradiographic stage and a radiographic stage to further characterize the condition. Importantly, our main intention is to stress a new disease concept, not to create a new name. Although the term "AS" is well established, it should be reserved for those patients who have radiographic evidence of ankylosis.

Conclusion

There are compelling data demonstrating that many patients with AS at the start of their illness and for many years thereafter often do not show radiographic evidence of sacroiliitis, since it takes time to evolve. There is a need for an early diagnosis in all patients with AS/axial SpA, especially now that more effective treatment options are available. Moreover, better classification criteria are needed in order to avoid misclassification in clinical studies/drug trials of patients with early disease, particularly those at the preradiographic stage. Any set of new classification criteria needs to be validated internationally and needs to be compared with existing criteria as well.

REFERENCES

1. Van der Linden S, Valkenburg HA, de Jongh BM, Cats A. The risk of developing ankylosing spondylitis in HLA-B27 positive individuals: a comparison of relatives of spondylitis patients with the general population. *Arthritis Rheum* 1984;27:241-9.
2. Gran JT, Husby G, Hordvik M. Prevalence of ankylosing spondylitis in males and females in a young middle-aged population in Tromsø, northern Norway. *Ann Rheum Dis* 1985;44:359-67.
3. Braun J, Bollow M, Remlinger G, Eggens U, Rudwaleit M, Distler A, et al. Prevalence of spondylarthropathies in HLA-B27 positive and negative blood donors. *Arthritis Rheum* 1998;41:58-67.
4. Saraux A, Guedes C, Allain J, Devauchelle V, Valls I, Lamour A, et al, and the Société de Rhumatologie de l'Ouest. Prevalence of rheumatoid arthritis and spondyloarthropathy in Brittany, France. *J Rheumatol* 1999;26:2622-7.
5. Khan MA. Update on spondyloarthropathies. *Ann Intern Med* 2002;136:896-907.
6. Sieper J, Braun J, Rudwaleit M, Boonen A, Zink A. Ankylosing spondylitis: an overview. *Ann Rheum Dis* 2002;61 Suppl 3:iii8-18.
7. Feldtkeller E, Khan MA, van der Heijde D, van der Linden S, Braun J. Age at disease onset and diagnosis delay in HLA-B27 negative vs. positive patients with ankylosing spondylitis. *Rheumatol Int* 2003;23:61-6.
8. Dougados M, Dijkmans B, Khan M, Maksymowich W, van der Linden S, Brandt J. Conventional treatments for ankylosing spondylitis. *Ann Rheum Dis* 2002;61 Suppl 3:iii40-50.
9. Chen J, Liu C. Methotrexate for ankylosing spondylitis. *Cochrane Database Syst Rev* 2004;3:CD004524.
10. Braun J, Alten R, Burmester G, Ebner W, Gomor B, Grasedyck K, et al. Efficacy of sulfasalazine in undifferentiated spondyloarthritis and early ankylosing spondylitis: a multicenter randomized controlled trial. *Ann Rheum Dis* 2004;63 Suppl 1:413.
11. Van Denderen JC, van der Paardt M, Nurmohamed MT, de Ryck YM, Dijkmans BA, van der Horst-Bruinsma IE. Double-blind, randomised, placebo-controlled study of leflunomide in the treatment of active ankylosing spondylitis. *Ann Rheum Dis* 2004;63 Suppl 1:397.
12. Braun J, Brandt J, Listing J, Zink A, Alten R, Golder W, et al. Treatment of active ankylosing spondylitis with infliximab: a randomised controlled multicentre trial. *Lancet* 2002;359:1187-93.
13. Van den Bosch F, Kruijthof E, Baeten D, Herrens A, de Keyser F, Mielants H, et al. Randomized double-blind comparison of chimeric monoclonal antibody to tumor necrosis factor α (infliximab) versus placebo in active spondylarthropathy. *Arthritis Rheum* 2002;46:755-65.
14. Marzo-Ortega H, McGonagle D, O'Connor P, Emery P. Efficacy of etanercept in the treatment of the enthesal pathology in resistant spondylarthropathy: a clinical and magnetic resonance imaging study. *Arthritis Rheum* 2001;44:2112-7.
15. Gorman JD, Sack KE, Davis JC Jr. Treatment of ankylosing spondylitis by inhibition of tumor necrosis factor α . *N Engl J Med* 2002;346:1349-56.
16. Maksymowich WP, Jhangri GS, Lambert RG, Mallon C, Buenviaje H, Pedrycz E, et al. Infliximab in ankylosing spondylitis: a prospective observational inception cohort analysis of efficacy and safety. *J Rheumatol* 2002;29:959-65.
17. Breban M, Vignon E, Claudepierre P, Devauchelle V, Wendling D, Lespessailles E, et al. Efficacy of infliximab in refractory ankylosing spondylitis: results of a six-month open-label study. *Rheumatology (Oxford)* 2002;41:1280-5.
18. Brandt J, Khariouzov A, Listing J, Haibel H, Sorensen H, Grassnickel L, et al. Six-month results of a double-blind, placebo-controlled trial of etanercept treatment in patients with active ankylosing spondylitis. *Arthritis Rheum* 2003;48:1667-75.
19. Davis JC Jr, van der Heijde D, Braun J, Dougados M, Cush J, Clegg DO, et al. Recombinant human tumor necrosis factor receptor (etanercept) for treating ankylosing spondylitis: a randomized, controlled trial. *Arthritis Rheum* 2003;48:3230-6.
20. Braun J, Baraliakos X, Golder W, Brandt J, Rudwaleit M, Listing J, et al. Magnetic resonance imaging examinations of the spine in patients with ankylosing spondylitis, before and after successful therapy with infliximab: evaluation of a new scoring system. *Arthritis Rheum* 2003;48:1126-36.
21. Rudwaleit M, Baraliakos X, Brandt J, Sieper J, Braun J. Magnetic resonance imaging (MRI) of the spine and the sacroiliac joints (SIJ) in ankylosing spondylitis (AS) before and during therapy with the anti-TNF agent etanercept. *Ann Rheum Dis* 2004;63 Suppl 1:397.
22. Rudwaleit M, Listing J, Brandt J, Braun J, Sieper J. Prediction of a major clinical response (BASDAI 50) to tumour necrosis factor α blockers in ankylosing spondylitis. *Ann Rheum Dis* 2004;63:665-70.
23. Kellgren JH, Jeffrey MR, Ball J. The epidemiology of chronic rheumatism. Vol. I. Oxford: Blackwell Scientific Publications; 1963. p. 326-7.
24. Gofton JP, Lawrence JS, Bennett PH, Burch TA. Sacroiliitis in eight populations. *Ann Rheum Dis* 1966;25:528-33.
25. Bennett PH, Wood PH. Population studies of the rheumatic diseases. Amsterdam: Excerpta Medica Foundation; 1968. p. 456-7.
26. Calin A, Porta J, Fries JF, Schurmann DJ. Clinical history as a screening test for ankylosing spondylitis. *JAMA* 1977;237:2613-4.
27. Van der Linden SM, Valkenburg HA, Cats A. Evaluation of the diagnostic criteria for ankylosing spondylitis: a proposal for modification of the New York criteria. *Arthritis Rheum* 1984;27:361-8.
28. Mau W, Zeidler H, Mau R, Majewski A, Freyschmidt J, Stangel W, et al. Clinical features and prognosis of patients with possible ankylosing spondylitis: results of a 10-year followup. *J Rheumatol* 1988;15:1109-14.
29. Said-Nahal R, Miceli-Richard C, Berthelot JM, Duche A, Demis-Laboue E, Le Blevec G, et al, on behalf of the Groupe Français D'Étude Génétique Des Spondylarthropathies. The familial form of spondylarthropathy: a clinical study of 115 multiplex families. *Arthritis Rheum* 2000;43:1356-65.
30. Oostveen J, Prevo R, den Boer J, van de Laar M. Early detection of sacroiliitis on magnetic resonance imaging and subsequent development of sacroiliitis on plain radiography: a prospective, longitudinal study. *J Rheumatol* 1999;26:1953-8.
31. Ahlstrom H, Feltelius N, Nyman R, Hallgren R. Magnetic resonance imaging of sacroiliac joint inflammation. *Arthritis Rheum* 1990;33:1763-9.
32. Braun J, Bollow M, Eggens U, König H, Distler A, Sieper J. Use of dynamic magnetic resonance imaging with fast imaging in the

- detection of early and advanced sacroiliitis in spondylarthropathy patients. *Arthritis Rheum* 1994;37:1039–45.
33. Rudwaleit M, Listing J, Marker-Hermann E, Zeidler H, Braun J, Sieper J. The burden of disease in patients with ankylosing spondylitis (AS) and pre-radiographic axial spondyloarthritis [abstract]. *Arthritis Rheum* 2004;50 Suppl 9:S211.
 34. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatology Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
 35. Ostergaard M, Hansen M, Stoltenberg M, Jensen KE, Szkudlarek M, Pedersen-Zbinden B, et al. New radiographic bone erosions in the wrists of patients with rheumatoid arthritis are detectable with magnetic resonance imaging a median of two years earlier. *Arthritis Rheum* 2003;48:2128–31.
 36. Khan MA, van der Linden SM, Kushner I, Valkenburg HA, Cats A. Spondylitic disease without radiologic evidence of sacroiliitis in relatives of HLA-B27 positive patients. *Arthritis Rheum* 1985;28:40–3.
 37. Dougados M. Diagnostic features of ankylosing spondylitis. *Br J Rheumatol* 1995;34:301–5.
 38. Van Tubergen A, Heuft-Dorenbosch L, Schulpen G, Landewe R, Wijers R, van der Heijde D, et al. Radiographic assessment of sacroiliitis by radiologists and rheumatologists: does training improve quality? *Ann Rheum Dis* 2003;62:519–25.
 39. Goei The HS, Steven MM, van der Linden SM, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis: a comparison of the Rome, New York and modified New York criteria in patients with a positive clinical history screening test for ankylosing spondylitis. *Br J Rheumatol* 1985;24:242–9.
 40. Rudwaleit M, Listing J, Marker-Hermann E, Zeidler H, Zink A, Braun J, et al. Spinal mobility in early ankylosing spondylitis (AS) and in axial SpA without definite radiographic sacroiliitis. *Ann Rheum Dis* 2004;63 Suppl 1:400.
 41. Sadowska-Wroblewska M, Filipowicz A, Garwolinska H, Michalski J, Rusiniak B, Wroblewska T. Clinical signs and symptoms useful in the early diagnosis of ankylosing spondylitis. *Clin Rheumatol* 1983;2:37–43.
 42. Thomas E, Silman AJ, Papageorgiou AC, Macfarlane GJ, Croft PR. Association between measures of spinal mobility and low back pain. *Spine* 1998;23:343–7.
 43. Rudwaleit M, Metter A, Listing J, Sieper J, Braun J. Clinical parameters in the differentiation of inflammatory back pain from non-inflammatory back pain. *Ann Rheum Dis* 2002;61 Suppl 1:57.
 44. Underwood MR, Dawes P. Inflammatory back pain in primary care. *Br J Rheumatol* 1995;34:1074–7.
 45. Van der Linden SM, Fahrner H. Occurrence of spinal pain syndromes in a group of apparently healthy and physically fit sportsmen (orienteers). *Scand J Rheumatol* 1988;17:475–81.
 46. Fries JF, Hochberg MC, Medsger TA Jr, Hunder GG, Bombardier C, and The American College of Rheumatology Diagnostic and Therapeutic Criteria Committee. Criteria for rheumatic disease: different types and different functions. *Arthritis Rheum* 1994;37:454–62.
 47. Griner PF, Mayewski RJ, Mushlin AI, Greenland P. Selection and interpretation of diagnostic tests and procedures: principles and applications. *Ann Intern Med* 1981;94:553–600.
 48. Rao JK, Allen NB, Pincus T. Limitations of the 1990 American College of Rheumatology classification criteria in the diagnosis of vasculitis. *Ann Intern Med* 1998;129:345–52.
 49. Hunder GG. The use and misuse of classification and diagnostic criteria for complex diseases. *Ann Intern Med* 1990;129:417–8.
 50. Rudwaleit M, van der Heijde D, Khan MA, Braun J, Sieper J. How to diagnose axial spondyloarthritis early. *Ann Rheum Dis* 2004;63:535–43.
 51. Pauker SG, Kassirer JP. Decision analysis. *N Engl J Med* 1987;316:250–8.
 52. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. Clinical epidemiology: a basic science for clinical medicine. 2nd ed. Boston: Little, Brown; 1991. p. 69–152.
 53. Sieper J, Rudwaleit M. Early referral recommendations for axial spondyloarthritis (including pre-radiographic and radiographic ankylosing spondylitis) in primary care. *Ann Rheum Dis* 2004. E-pub ahead of print.
 54. Vitali C. Classification criteria for Sjögren's syndrome [letter]. *Ann Rheum Dis* 2003;62:94.
 55. Dougados M, van der Linden S, Juhlin R, Huitfeldt B, Amor B, Calin A, et al, and the European Spondylarthropathy Study Group. The European Spondylarthropathy Study Group preliminary criteria for the classification of spondylarthropathy. *Arthritis Rheum* 1991;34:1218–27.
 56. Amor B, Dougados M, Mijiyawa M. Critères de classification des spondylarthropathies. *Rev Rhum Mal Osteoartic* 1990;57:85–9.
 57. Khan MA, van der Linden SM. A wider spectrum of spondyloarthropathies. *Semin Arthritis Rheum* 1990;20:107–13.
 58. Moll JM, Haslock I, Macrae IF, Wright V. Associations between ankylosing spondylitis, psoriatic arthritis, Reiter's disease, the intestinal arthropathies, and Behçet's syndrome. *Medicine (Baltimore)* 1974;53:343–64.
 59. Collantes E, Veroz R, Escudero A, Munoz E, Munoz MC, Ciscal A, et al, and the Spondyloarthropathy Study Group. Can some cases of 'possible' spondyloarthropathy be classified as 'definite' or 'undifferentiated' spondyloarthropathy? Value of criteria for spondyloarthropathies. *Joint Bone Spine* 2000;67:516–20.
 60. Amor B, Dougados M, Lustrat V, Menkes CJ, Dubost JJ, Roux H, et al. Évaluation des critères des spondylarthropathies d'Amor et de l'European Spondylarthropathy Study Group (ESSG). *Ann Med Interne (Paris)* 1991;142:85–9.
 61. Amor B. Usefulness of criteria for spondyloarthropathies. *Joint Bone Spine* 2000;67:502–3.

APPENDIX A: DISEASE PROBABILITY CALCULATIONS

The likelihood ratio (LR) combines both the sensitivity and the specificity of a given parameter in a single value and is defined as follows:

$$\text{Positive LR} = \text{sensitivity}/1 - \text{specificity}$$

if the parameter is present (52). The disease probability (posttest probability) can be calculated by using the LR, as follows:

$$\text{Posttest probability} = \text{posttest odds}/1 + \text{posttest odds}$$

The posttest odds is derived from the product of the pretest odds and the LR product of the various test parameters that are present or positive, using the following formula:

$$\text{Pretest odds} \times \text{LR}_1 \times \text{LR}_2 \times \text{LR}_3 \times \dots = \text{posttest odds}$$

where LR_1 , LR_2 , LR_3 , ... represent the LR of the various diagnostic parameters (tests) applied. The pretest odds can be calculated from the pretest probability (disease prevalence) by the formula:

$$\text{Pretest odds} = \text{pretest probability}/1 - \text{pretest probability}$$

In the calculations, we assume a pretest probability (prevalence of axial SpA among patients with chronic back pain) of 5% (44). Thus, a pretest probability of 5% corresponds to a pretest odds of 0.05, and a posttest probability of ~90% corresponds to a posttest odds of ~10. Using this approach and the LRs shown in Figure 3, one can understand how many and which parameters need to be present in order to reach a predefined level of disease probability (i.e., level of diagnostic certainty). As shown in Figure 3, different combinations of parameters can be chosen to reach an LR product of 200, which is needed in order to reach a level of diagnostic certainty of at least 90% ($0.05 \times 200 = 10$).