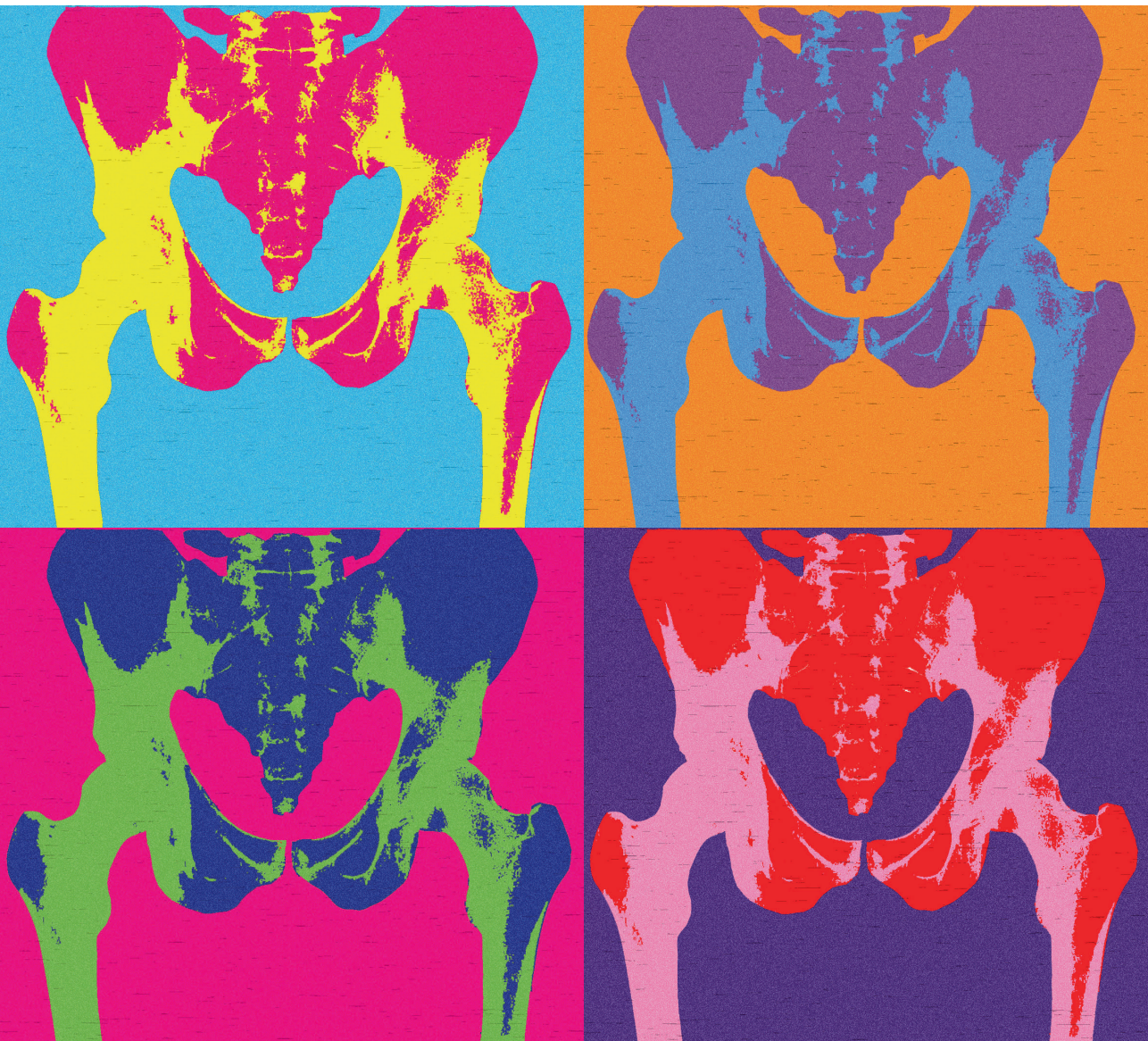
Spondyloarthritis

Recognition
Imaging
Treatment

Rosaline van den Berg



Spondyloarthritis

Recognition
Imaging
Treatment

Rosaline van den Berg

ISBN/EAN 9789461087836

Printed by: Gildeprint - Enschede

The printing of this thesis was financially supported by the Dutch Arthritis Foundation (het Reumafonds), Pfizer B.V., AbbVie and UCB Pharma B.V.

Spondyloarthritis

Recognition
Imaging
Treatment

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 29 oktober 2014
klokke 13:45 uur

door

Rosaline van den Berg

geboren te Alkmaar
in 1984

Promotiecommissie

Promotor:

Prof. dr. D.M.F.M. van der Heijde

Co-promotores:

Dr. F.A. van Gaalen
Dr. M. Reijnierse

Overige leden:

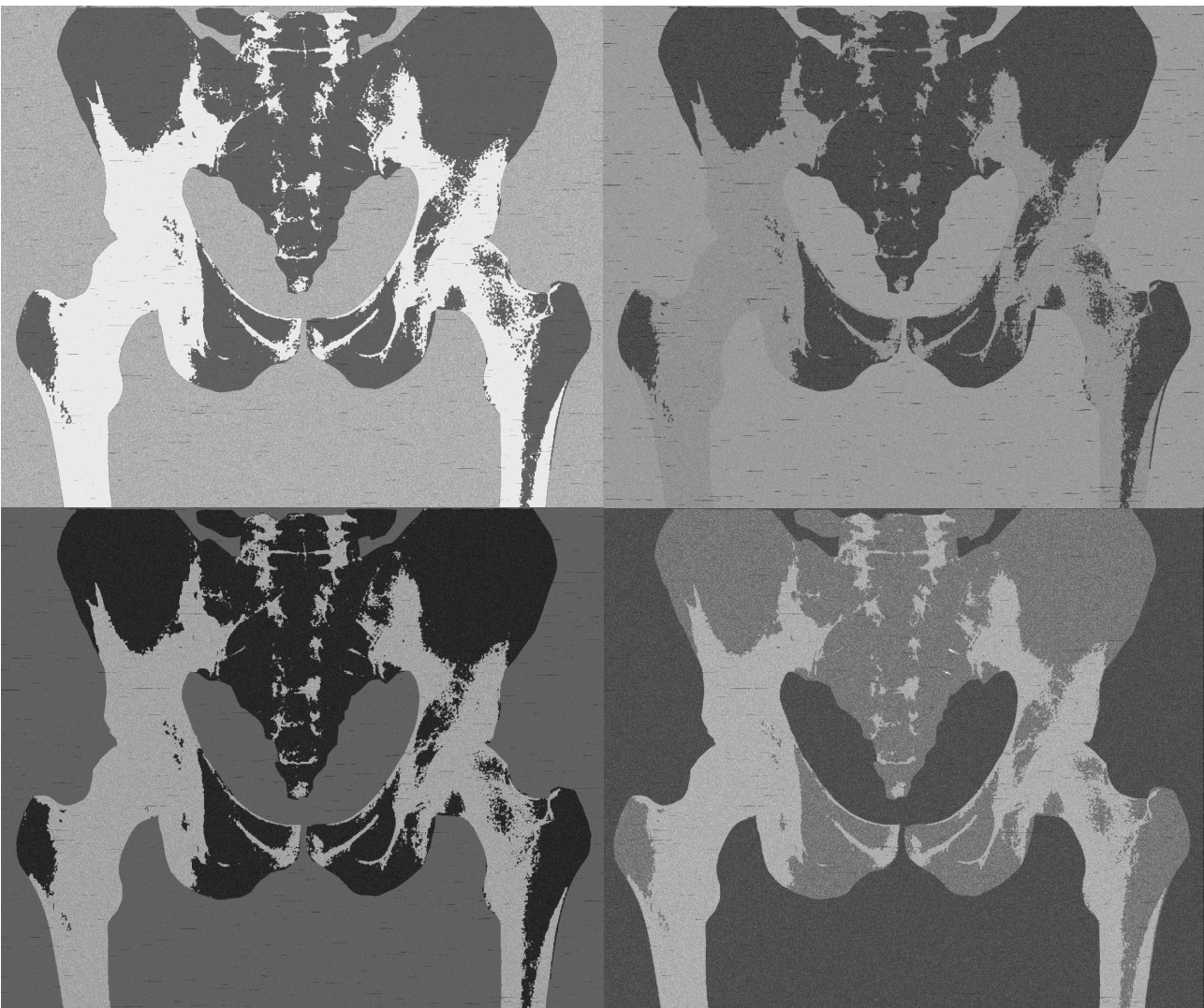
Prof. dr. G. Kloppenburg
Prof. dr. R.B.M. Landewé (Academisch Medisch Centrum, Amsterdam)
Prof. dr. P. Claudepierre (Université Paris Est Créteil/Hôpital Henri-Mondor, Créteil, France)

'Jij bent toch onderzoeker? Wat heb je ontdekt dan?'
(Jaap Andriese, 7 jaar)

CONTENT

General introduction

1



General introduction

The history of Ankylosing Spondylitis

Ankylosing spondylitis (AS; σκληρός= stiff σπονδύλων= vertebrae) is a chronic inflammatory rheumatic disease with a long history, primarily affecting the axial skeleton. For example, evidence of ossification of joints and entheses corresponding with lesions seen in AS was already seen in a 5000-year old Egyptian mummy [1]. Realdo Colombo, an Italian surgeon and professor in anatomy, probably described AS-like characteristics for the first time in 1559, and in 1691 Bernard Connor described pathologic changes of the skeleton possibly associated with AS [2, 3]. The symptoms and complaints of Leonard Trask (1805-1861) - he developed progressive severe thoracic kyphosis after he fell from a horse, resulting in invalidating disabilities - were extensively described as he was considered a medical curiosity. Until his death, his condition remained a medical mystery, however, post-mortem he was diagnosed with AS [4].

In 1893, descriptions of what might have been AS were given by the neurologist and psychologist Vladimir Bekhterev, as well as by the neurologists Adolph Strümpell in 1897 and Pierre Marie in 1898 [4-6]. The disease was long known as morbus Bechterew and Pierre Marie Strümpell disease although it is not certain the cases these authors described were what is now known as AS. Initially, AS was not identified as a separate entity, but considered as a subtype of rheumatoid arthritis ("rheumatoid spondylitis") until the 1960s [5]. Only during the mid-1960s, AS was recognized as a separate entity with well-defined manifestations and radiographic criteria [6, 7].

Clinical manifestations and epidemiology

AS is characterized by inflammation in the sacroiliac joints (SI-joints) and the vertebrae, causing severe pain and stiffness in the back and/or buttock area. In some patients the inflammation ultimately leads to bone formation in the SI-joints and/or spine, thereby deteriorating spinal mobility resulting in impaired daily functioning. Complaints associated with AS usually start in the 2nd and 3rd decade of life, and by the age of 45 years, more than 95% of the patients are symptomatic [8, 9].

The cause of AS is multifactorial, consisting of genetic and environmental factors, but is not completely elucidated yet. Regarding genetic factors, a strong association with the major histocompatibility complex (MHC) class I human leukocyte antigen (HLA)-B27, present in 80-95% of the AS patients, is known [10]. In addition to the prototypical genetic risk factor HLA-B27, HLA-B60 is a modest risk factor for AS [11]. Moreover, several new genetic risk factors outside of the MHC locus, including genetic variants in the ERAP1 and IL-23 receptor gene, have been discovered recently [12].

Worldwide, the prevalence of AS varies depending on the prevalence of HLA-B27. In central Europe, the prevalence of HLA-B27 varies from 6 to 9% and the estimated prevalence of AS ranges between 0.1-0.7% [13-16]. In northern Europe the prevalence of HLA-B27 is higher, around 14%, and the estimated prevalence of AS is accordingly higher as well: 1.1-1.4% [17]. In the USA the prevalence of HLA-B27 is also around 6% and the estimated prevalence of AS is around 0.5% [18, 19], and among Haida Indians the prevalence of HLA-B27 is very high, around 50%, and the prevalence of AS is estimated around 6% [20].

In addition to complaints in the axial skeleton, AS patients may suffer from complaints in peripheral joints (peripheral arthritis, dactylitis and enthesitis) and extra-articular manifestations (uveitis, psoriasis and inflammatory bowel disease (IBD)), often with substantial overlap. The estimated prevalence of peripheral complaints and extra-articular manifestations differ in various studies, due to differences in inclusion criteria and methodological characteristics resulting in different study populations regarding clinical characteristics, as well as due to differences in geographical area. Pooled prevalence

revealed that approximately 25.8% (95% CI: 24.1%-27.6%) of the patients suffer from at least one episode of acute anterior uveitis during their disease course. Approximately 9.3% (95% CI: 8.1%-10.6%) of the patients have psoriasis, and the pooled prevalence of IBD is 6.8% (95% CI: 6.1%-7.7%) [21]. Reported prevalence of peripheral arthritis ranges from 14.4% to 46.6% [22-24], of dactylitis it ranges from 1.9 to 3.1% [23, 24] and the reported prevalence of enthesitis ranges from 9.8% to 49% [22-26].

Classification and diagnostic criteria

Appropriate diagnostic criteria for AS are lacking, but classification criteria are available. According to the modified New York criteria (table 1), the presence of radiographic sacroiliitis (grade at least 2 bilaterally or grade 3-4 unilaterally) in combination with one of the clinical criteria, is mandatory in order to classify a patient as AS [27].

Table 1: the modified New York criteria for Ankylosing Spondylitis.

Definite ankylosing spondylitis	If the radiological criterion is associated with at least 1 clinical criterion
Clinical criteria	Low back pain and stiffness for >3 months which improves with exercise, but is not relieved by rest.
	Limitation of motion of the lumbar spine in both the sagittal and frontal planes.
	Limitation of chest expansion relative to normal values correlated for age and sex.
Radiological criterion	Sacroiliitis grade ≥2 bilaterally or grade 3-4 unilaterally
Grading of radiographic sacroiliitis	
Grade 0	Normal
Grade 1	Suspicious changes
Grade 2	Minimal abnormality – small localized areas with erosion or sclerosis, without alteration in the joint width
Grade 3	Unequivocal abnormality – moderate or advanced sacroiliitis with one or more of: erosions, evidence of sclerosis, widening, narrowing or partikal ankylosis
Grade 4	Severe abnormality – total ankylosis

Adapted from van der Linden et al. A&R 1984;27:361-8 [27].

However, it often takes 6-8 years from the onset of symptoms before radiographic sacroiliitis can be detected on plain radiographs [28-30]. It is thought that radiographic changes (erosions, sclerosis, joint space narrowing/widening, ankylosis) reflect the consequences of inflammation rather than inflammation itself [28-30] (figure 1).
However, the underlying mechanisms of the inflammatory process leading to new bone formation are not fully understood [31, 32]. Moreover, not every patient with symptoms will develop radiographic sacroilitis [28-30]. AS is therefore considered as the prototype disorder of the whole concept of spondyloarthritis (SpA), which is a group of interrelated rheumatic diseases with common features (figure 2).
Furthermore, in some SpA patients the disabling problems are not the back pain and/or stiffness of the back, but predominantly peripheral and/or extra-articular complaints, at least during some periods of the disease course. To be able to classify the whole group of SpA, the Amor and ESSG criteria have been developed in the 1990s. In these classification

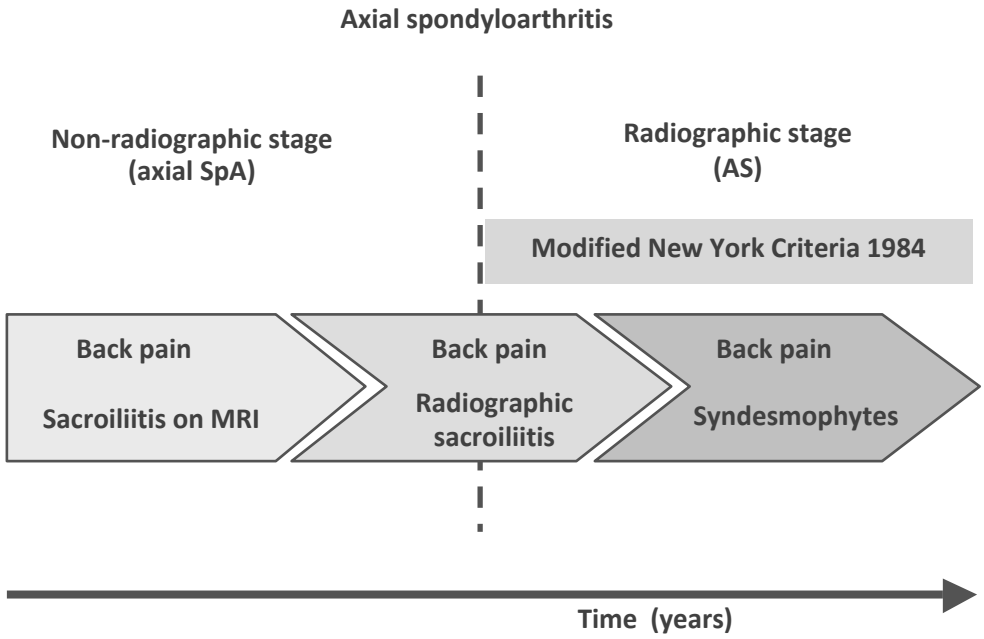


Figure 1: Axial spondyloarthritis. Reprinted from Rudwaleit et al. A&R 2005;52:1000-8 [29].

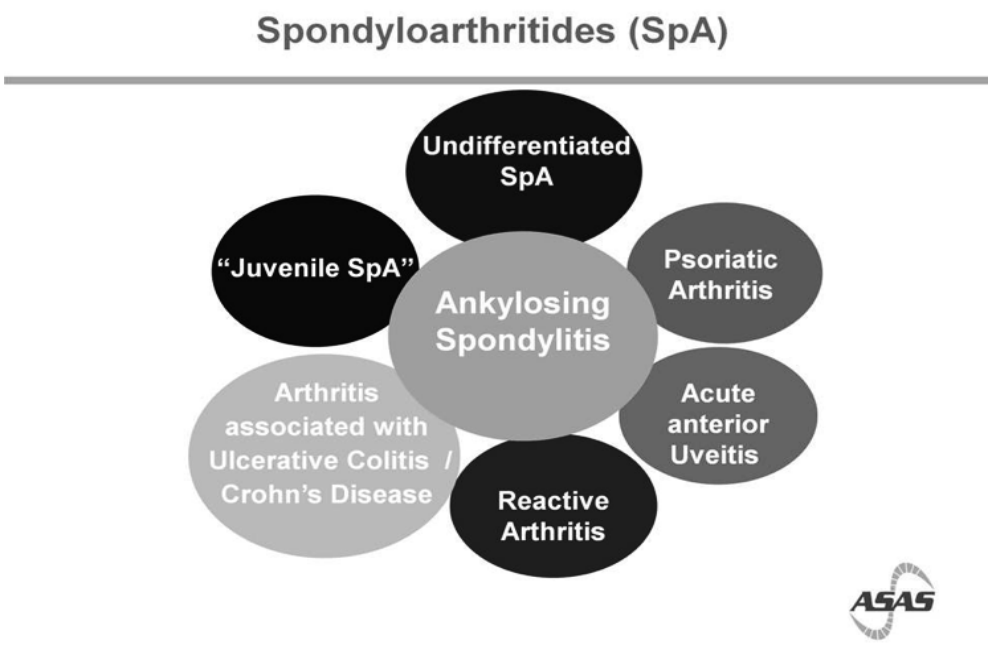


Figure 2: The concept of spondyloarthritis (SpA). Reprinted from the ASAS website [48].

criteria, radiographic sacroiliitis is included as one of the SpA-features, but in contrast to the modified New York criteria, it is not a mandatory criterion [33, 34]. The entry criteria of the ESSG criteria are inflammatory back pain (IBP) and/or peripheral arthritis. According to the ESSG criteria, patients with at least one of the entry criteria in combination with one minor criterion are classified as having SpA. The Amor criteria consist of a list of signs, none of which is required to classify a patient as having SpA [9, 35, 36].

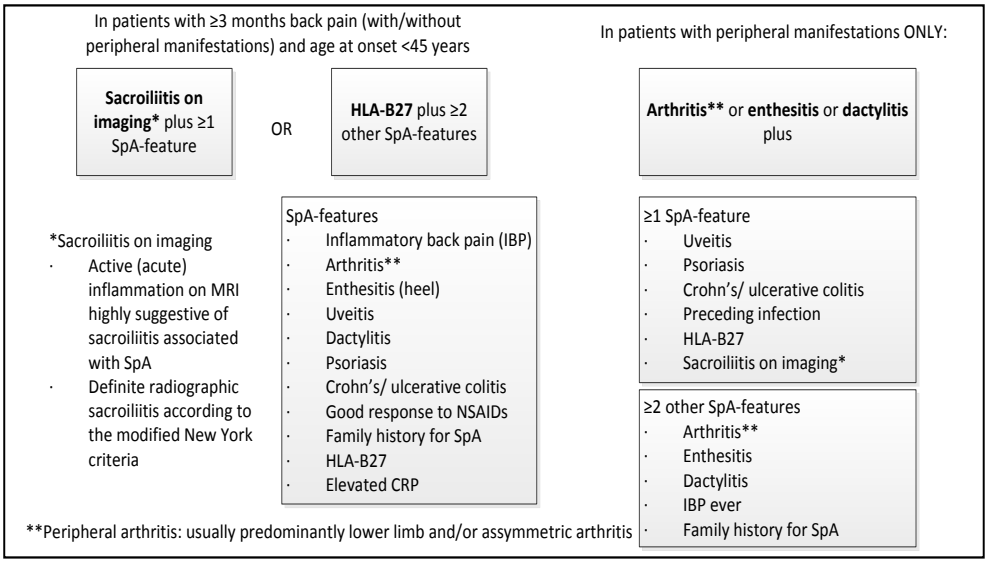


Figure 3: Combined use of the ASAS criteria for axial SpA and ASAS criteria for peripheral SpA in the entire SpA population. Reprinted from Rudwaleit et al. ARD 2011;70:25-31 [23].

Both the ESSG and Amor classification criteria were developed before Magnetic Resonance Imaging (MRI) became widely available. Yet, the introduction of MRI in the field of SpA made it possible to detect inflammation in the SI-joints and spine, which is considered the first step in the development of structural damage as seen in AS [31, 32]. In 2009, the Assessment of SpondyloArthritis international Society (ASAS) proposed two classification criteria sets for SpA (figure 3). One set can be applied in patients with predominantly peripheral manifestations (peripheral SpA) [23] and the other set in patients with predominantly axial manifestations (axial SpA) [30]. For the first time in history, MRI was included in classification criteria for SpA and plays an important role in especially the axial SpA criteria. The axial SpA criteria can only be applied in patients with back pain for more than 3 months and the onset of back pain before the age of 45 years and consists of two arms. In the imaging arm, patients can be classified as axial SpA if one SpA-feature is present in addition to sacroiliitis on MRI or radiographs sacroiliitis [30]. Patients can be classified as axial SpA in the clinical arm if in addition to HLA-B27 positivity two other SpA-features are present [30]. The classification criteria for peripheral SpA can be applied in patients with currently peripheral manifestations only. In order to classify a patient as having peripheral SpA, other SpA-features should be present in addition to peripheral arthritis compatible with SpA (usually asymmetric and/or predominantly involvement of the lower limb), enthesitis or dactylitis. A patient with current peripheral arthritis/enthesitis/dactylitis can fulfill the ASAS peripheral SpA criteria if at least one of the following SpA-feature is present: uveitis, psoriasis, IBD, preceding infection (urethritis/cervicitis or diarrhea within one month before the onset of arthritis/enthesitis/dactylitis), HLA-B27 positivity or sacroiliitis on imaging. A patient with current

peripheral arthritis/enthesitis/dactylitis can also fulfill the ASAS peripheral SpA criteria if at least two of the following SpA-features are present: peripheral arthritis compatible with SpA (present or past), enthesitis (present or past), dactylitis (present or past), IBP ever, or a positive family history for SpA (figure 3) [23]. Definitions of all SpA-features are given in table 2.

Table 2: Definitions of SpA-features applied in the ASAS classification criteria for axial and peripheral spondyloarthritis.

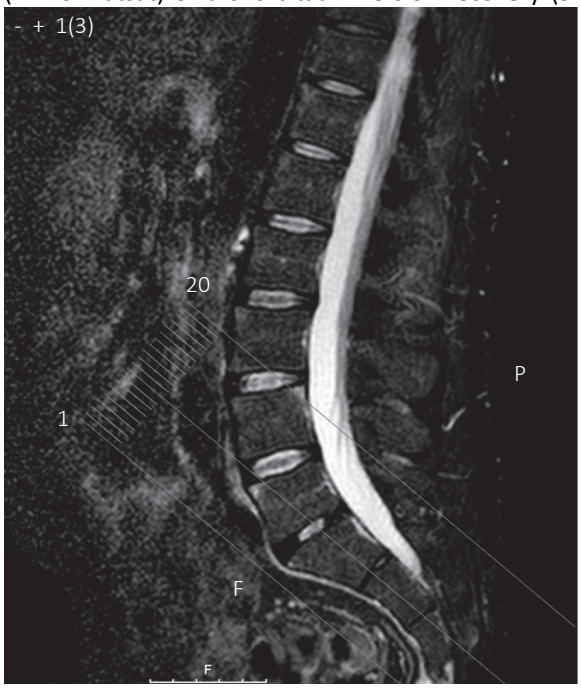
SpA-feature	Definition
Inflammatory back pain (IBP)	IBP according to experts: 4 out of 5 of the following paramters present: 1. Age at onset <40 years 2. Insidious onset 3. No improvement with exercise 4. No improvement with rest 5. Pain at night (with improvement upon getting up)
Arthritis	Past or present active synovitis diagnosed by a physician
Enthesitis (heel)	Heel enthesitis: past or present spontaneous pain or tenderness at examination of the site of the insertion of the Achilles tendon or plantar fascia at the calcaneus
Uveitis	Past or present uveitis anterior, confirmed by an ophthalmologist
Dactylitis	Past or present dactylitis, diagnosed by a physician
Psoriasis	Past or present psoriasis, diagnosed by a physician
Inflammatory bowel disease (IBD)	Past or present Crohn's disease or ulcerative colitis, diagnosed by a physician
Good response to NSAIDs	24-48 hours after a full dose of a non-steroidal anti-inflammatory drug (NSAID) the back pain is not present anymore or is much better
Family history	Presence in first-degree (mother, father, sisters, brothers, children) or second-degree (maternal and paternal grandparents, aunts, uncles, nieces and nephews) relatives of any of the following: 1. Ankylosing Spondylits 2. Psoriasis 3. Uveitis 4. Reactive Arthritis 5. Inflammatory Bowel Disease
Elevated CRP	C-reactive protein (CRP) concentration above upper normal limit in the presence of back pain, after exclusion of other causes for elevated CRP concentration
HLA-B27	Positive testing according to standard laboratory techniques
Sacroiliitis by radiographs	Bilateral grade 2-4 or unilateral grade 3-4 sacroiliitis on plain radiographs, according to the modified New York criteria
Sacroiliitis by MRI	Active inflammatory lesions of sacroiliac joints with definite bone marrow edema/osteitis, suggestive of sacroiliitis associated with spondyloarthritis

Reprinted from Rudwaleit et al. ARD 2009;68:777-83 [30].

Since the clinical presentation of SpA is heterogeneous and because of the lack of diagnostic criteria, diagnosing SpA can be challenging for physicians, especially in the absence of (radiographic) sacroiliitis. Because no single shared distinguishing feature exists, the diagnosis is usually based on the combination of symptoms, the findings of physical examination, imaging and laboratory results [37, 38]. To assist in the diagnostic process, a tool based on Bayes' theorem has been developed by a group of rheumatologists, incorporating all relevant SpA-features [35]. The formula of this tool, which is known as the Berlin algorithm, allows calculation of the disease probability for any individual patient with IBP according to the clinical presentation, and the final post-test probability may help in making the diagnosis of axial SpA. In general, three other SpA-features in addition to the presence of IBP result in a probability of about 90% for axSpA. However, the drawback of this algorithm is that it is developed for patients with IBP, while it is becoming more and more evident that many patients with axial SpA do not have IBP, and vice versa, that many patients with IBP do not have axial SpA [8, 39].

Magnetic Resonance Imaging

To provide a solid basis for the application of MRI of the sacroiliac joints (MRI-SI) in the ASAS axial SpA criteria, ASAS developed recommendations how to perform an optimal MRI-SI and how to define a positive MRI-SI [40, 41]. Inflammatory changes are best visualized by a water-sensitive sequence; a T2-weighted turbo spin-echo sequence with fat-saturation (T2 TSE fatsat) or a short tau inversion recovery (STIR) sequence with a high resolution. The latter has a robust performance and a high sensitivity and is therefore preferred. Structural changes such as erosions and fatty depositions are best visualized using a T1-weighted turbo spin-echo sequence (T1 TSE). Therefore, ASAS recommends to perform T1 TSE sequence and STIR sequence of the SI-joints in a semi-coronal section orientation along the long axis of the sacral bone with slices of 4mm thickness using an MR machine with a field strength of 1.5 Tesla (figure 4) [40].



Alternatively, the administration of a paramagnetic contrast medium (gadolinium-chelate; Gd) in a T1 TSE sequence with fat saturation (T1 TSE fatsat) could be considered as it occasionally gives additional value, especially in depicting enthesitis and capsulitis.

Figure 4: Scout view of the SI-joints in a semi-coronal section orientation along the long axis of the sacral bone.

However, the STIR sequences and the T1 post-Gd sequences give largely overlapping information [42, 43]. Therefore, the administration of Gd is not recommended by ASAS [40].
The clear presence of one BME lesion highly suggestive of SpA visible on a T2 or STIR sequence (or alternatively osteitis (on T1 post-Gd)) located in the typical anatomical areas

(subchondral or periarticular bone marrow) on at least two consecutive slices or the clear presence of several lesions on a single slice allows to define the MRI-SI as positive. The presence of isolated synovitis, enthesitis or capsulitis without the presence of BME (or osteitis) is only occasionally seen [42] and is not sufficient for a positive MRI-SI [41]. Furthermore, the presence of isolated structural lesions without concomitant BME (or osteitis) does not suffice for the definition of a positive MRI-SI either [41].
In addition to the dichotomous evaluation of an MRI-SI according to the ASAS definition, MRI-SIs can be evaluated in a semi-quantitative manner, for example according to the Spondyloarthritis Research Consortium of Canada (SPARCC) method [44]. The SPARCC-score ranges from 0 to 72 points and has a high sensitivity to change. Therefore, the SPARCC-score is of particular value in clinical trials, testing the efficacy of (biological) treatment in terms of changes in inflammation over time.

Treatment

In 2006, ASAS in collaboration with EULAR developed evidence -based recommendations for the management of AS in order to contribute to the improvement of outcomes in patients [45]. Treatment should not only aim on improving signs and symptoms, but also on improving function and socioeconomic factors as well as preventing structural damage, thereby improving the quality of life of patients [46]. The standard treatment of AS patients consists of a combination of non-pharmacological and pharmacological treatment. The non-pharmacological treatment comprises education, exercise, physical therapy, rehabilitation, patient associations and self-help groups. Pharmacological treatment comprises treatment with non-steroidal anti-inflammatory drugs (NSAIDs), including selective cyclo-oxygenase 2 (COX-2) inhibitors, as first line drug. NSAIDs may rapidly improve spinal pain, peripheral joint pain and function. In patients with peripheral complaints, disease modifying antirheumatic drugs (DMARDs), including sulfasalazine and methotrexate, might be considered, but effect on axial complaints is lacking. Moreover, in patients with persistently high disease activity despite conventional treatments according to the ASAS recommendations, treatment with tumor necrosis factor-alpha (TNF-α) inhibitors should be considered [45].

Outline of this thesis

The recognition and treatment of SpA has been improved a lot over the last 10-15 years. These improvements resulted, inter alia, in the development of the ASAS classification criteria. At the same time, the ASAS classification criteria are contributing to further enhancements. Moreover, the treatment armamentarium has been broadened with TNF-α inhibitors.
However, there are still major challenges in recognizing, diagnosing, classifying and treating (early) SpA patients. Those aspects will be addressed in this thesis in three dedicated parts. Part I focuses on the early recognition of SpA and on the evaluation of classification criteria of SpA. The focus of part II is on the role of imaging in the early diagnosis of SpA. The current recommendations for management of AS and axSpA and the evidence as the base for these recommendations is the focus of part III.
The studies described in part I and II are largely performed in the SPondyloArthritis Caught Early (SPACE)-cohort and the DEvenir des Spondylarthropathies Indifférenciées Récentes (the evaluation of outcome of recent onset undifferentiated spondyloarthritis; DESIR)-cohort [25]. Both cohorts include patients with back pain to study - among other research questions - how patients with SpA can best be differentiated from patients without SpA, which factors are predictive for SpA and which factors are predictive for the progression of the disease. Therefore, information about the presence/absence of all SpA-features in all patients is collected in both the SPACE-cohort and the DESIR-cohort. Besides the similarities

between the two cohorts, small differences exist. The DESIR-cohort is a purely French cohort, with 25 participating centers across France, while the SPACE-cohort started as a Dutch single-center cohort in the LUMC. In the meantime, the SPACE-cohort has become an international multi-center cohort with participating centers in Norway, Italy and Sweden in addition to other participating hospitals in the Netherlands as well. Another difference between the DESIR-cohort and the SPACE-cohort is that the DESIR-cohort includes patients with inflammatory back pain (≥3 months, but <3 years) aged 18-50 with a suspicion of SpA while the SPACE-cohort includes patients with chronic back (pain ≥3 months, but ≤2 years) with the onset <45 years. As a consequence, the populations in both cohorts are slightly different.

One of the research questions we have is to assess the performance of existing classification criteria, including the recently developed ASAS axial SpA and peripheral SpA criteria. It is of particular interest to know the performance of the ASAS axial SpA and peripheral SpA criteria as this has not yet been evaluated in another cohort than the validation cohort. Some experts in the field impeach the ASAS axial SpA criteria as they question whether patients fulfilling the clinical arm reflect the same disease as patients fulfilling the imaging arm. Since the inclusion criteria of the SPACE-cohort yield also the inclusion of a control group with similar age, gender and symptom duration as the patients with axial SpA, the SPACE-cohort offers the opportunity to investigate this research question. The results of this investigation, as well as an extensive description of the SPACE-cohort, are presented in **chapter 2**. Some experts in the field who fear that the ASAS peripheral SpA criteria are not specific enough impeach the ASAS peripheral SpA criteria too. Therefore, a very similar analysis on the performance of the various classification criteria as described in **chapter 2**, is performed in the Early Arthritis Clinic (EAC)-cohort in **chapter 3**.

The Leiden EAC-cohort is a population-based prospective cohort, started in 1993 in order to detect and treat inflammatory disorders early in the disease state, especially early rheumatoid arthritis (RA). Patients with suspected arthritis are referred by general practitioners to the LUMC as soon as possible, and are seen within 2 weeks from referral. Patients are included in the EAC-cohort if arthritis of recent onset (<2 years) is confirmed by the rheumatologist [47]. Since it is known that up to 67% of the SpA patients with a symptom duration <2 years report arthritis as the first symptom, SpA and PsA are important parts of the differential diagnosis. Therefore, the EAC-cohort is a suitable cohort to try to answer this research question.

In **chapter 4**, the performance of the Berlin algorithm is evaluated. The inclusion criterion of the Berlin algorithm is IBP, however, the increasing evidence that not all patients with axial SpA have IBP stimulated us to propose two modifications of the algorithm. We test these proposed modifications in the SPACE-cohort since the inclusion criteria of the SPACE-cohort are chronic back pain (and not IBP) thereby offering the possibility of yielding axial SpA patients without IBP. In addition, we tested the proposed modifications in the original validation cohort of the ASAS classification criteria (**chapter 4**).

The focus of part II is on the role of imaging in classifying and diagnosing patients. It is known that it is challenging to reliably judge imaging of the SI-joints, especially plain radiographs. However, the consequences of a different judgment of the same set of imaging by different readers on the classification of patients, and in turn the consequences of a possibly other classification on the access to treatment is not known. As the DESIR-cohort contains judgments of MRIs and X-rays of the SI-joints by local radiologists and/or rheumatologists as well as judgments of the same images by central trained readers, this cohort offered the unique opportunity to study this. In **chapter 5** the evaluation of X-rays of the SI-joints by local readers is compared to the evaluation by central readers in terms of agreement on abnormal versus normal SI-joints permitting to diagnose radiographic sacroiliitis. **Chapter 6** is about the role of differences in judgments by local readers and central readers of all

imaging (combined radiographs and MRI of the SI-joints) and the effect on the ASAS axial SpA criteria.

Nowadays, more and more trials are being performed in patients with non-radiographic axial SpA. Eligibility of patients for these trials is often based on the judgment of local readers on a positive/negative MRI (ASAS definition). Within clinical trials, semi-quantitative scoring methods like the SPARCC-score method are used to measure changes in inflammation over time, evaluated by central trained readers. However, interreader reliability of the SPARCC-score in terms of smallest detectable change (SDC) is not known. Moreover, it is known that inflammation may spontaneously change over time, but it is not sufficiently investigated how many SPARCC-score points these spontaneous changes comprise. Furthermore, in case one needs to link the read for eligibility to the efficacy reading, it is not known what SPARCC-score cut-off value the equivalent is of a positive MRI (ASAS definition). These three questions are addressed in **chapter 7**.

As the field of SpA is rapidly moving, the ASAS in collaboration with EULAR intends to update the current recommendations for both the treatment of AS with TNF-α inhibitors and the recommendations for the management of AS which is the focus of part III. Preambles to these updates, up-to-date overviews regarding the implementation of these recommendations worldwide and regarding clinical trials and publications on AS therapy were needed. First, a comparison of national recommendations on TNF-α inhibitor use is made, with a focus on the similarities and differences compared to the ASAS/EULAR recommendations of 2006 (**chapter 8**). Second, a systematic literature review about the management of AS with non-pharmacological treatment and non-biologic drugs is performed (**chapter 9**). Third, a systematic literature review on biologic treatment of AS is performed (**chapter 10**). The results of these studies are presented to the working group of international experts who met during ASAS workshops to develop the new management recommendations, presented in **chapter 11**.

References

1. Dieppe P. Did Galen describe rheumatoid arthritis? Ann Rheum Dis 1988;47:84-5.
2. Benoist M. Pierre Marie. Pioneer investigator in ankylosing spondylitis. Spine 1995;20:849-52.
3. Blumberg B, Blumberg J. Bernard Connor (1666-1698) and his contribution to the pathology of ankylosing spondylitis. J Hist Med Allied Sci 1958;13:349-66.
4. Jayson M. Leonard Trask: the wonderful invalid: the first American description of ankylosing spondylitis. Arthritis Rheum 2003;48:612-3.
5. Claudepierre P, Wendling D, Breban M, *et al*. Ankylosing spondylitis, spondyloarthropathy, spondyloarthritis, or spondylarthritis: what's in a name? Joint Bone Spine 2012;79:534-5.
6. Kellgren J, Jeffrey M, Ball J. The epidemiology of chronic rheumatism. Oxford, Blackwell Scientific Publications 1963;1:326-7.
7. Bennett P, Burch T. Population studies of the rheumatic diseases. Amsterdam, Excerpta Medica Foundation 1968;456-7.
8. Rudwaleit M. New classification criteria for spondyloarthritis. Int J Adv Rheumatol 2010;8:1-7.
9. Rudwaleit M. New approaches to diagnosis and classification of axial and peripheral spondyloarthritis. Curr Opin Rheumatol 2010;22:375-80.
10. Feldtkeller E, Khan M, van der Heijde D, *et al*. Age at disease onset and diagnosis delay in HLA-B27 negative vs. positive patients with ankylosing spondylitis. Rheumatology International 2003;23:61-6.
11. van Gaalen F, Verduijn W, Roelen D, *et al*. Epistasis between two HLA antigens defines a subset of individuals at a very high risk for ankylosing spondylitis. Ann Rheum Dis 2013;72:974-8.
12. Brown M. Genetics and the pathogenesis of ankylosing spondylitis. Curr Opin Rheumatol 2009;21:318-23.
13. Guillemin F, Saraux A, Guggenbuhl P, *et al*. Prevalence of rheumatoid arthritis in France: 2001. Ann Rheum Dis 2005;64:1427-30.
14. Saraux A, Guillemin F, Guggenbuhl P, *et al*. Prevalence of spondyloarthropathies in France: 2001. Ann Rheum Dis 2005;6:1431-5.
15. Braun J, Listing J, Sieper J. Reply. Arthritis Rheum 2005;52:4049-50.
16. van der Linden S, Valkenburg H, de Jongh B, *et al*. 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.
17. Gran J, Husby G, Hordvik M. Prevalence of ankylosing spondylitis in males and females in a young middle-aged population of Tromso, northern Norway. Ann Rheum Dis 1985;44:359-67.
18. Helmick C, Felson D, Lawrence R, *et al*. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Arthritis Rheum 2008 Jan;58(1):15-25.
19. Reveille J, Hirsch R, Dillon C, *et al*. The prevalence of HLA-B27 in the US: data from the US National Health and Nutrition Examination Survey, 2009. Arthritis Rheum 2012;64:1407-11.
20. Gofton J, Robinson H, Trueman G. Ankylosing spondylitis in a Canadian Indian population. Ann Rheum Dis 1966;25:525-7.
21. Stolwijk C, van Tubergen A, Castillo-Ortiz J, *et al*. Prevalence of extra-articular manifestations in patients with ankylosing spondylitis: a systematic review and meta-analysis. Ann Rheum Dis [2-9-2013] doi: 10.1136/annrheumdis-2013-203582.
22. 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 (Oxford) 2009;48:404-9.
23. Rudwaleit M, van der Heijde D, Landewé R, *et al*. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. Ann Rheum Dis 2011;70:25-31.
24. 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.

25. Dougados M, D'Agostino M, 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.
26. Carron P, Van Praet L, Van den Bosch F. Peripheral manifestations in spondyloarthritis: relevance for diagnosis, classification and follow-up. Curr Opin Rheumatol 2012;24:370-4.
27. van der Linden S, Valkenburg H, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361-8.
28. Bennett A, 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.
29. Rudwaleit M, Khan M, Sieper J. The challenge of diagnosis and classification in early ankylosing spondylitis: do we need new criteria? Arthritis Rheum 2005;52:1000-8.
30. Rudwaleit M, van der Heijde D, Landewé 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.
31. Sieper J, Appel H, Braun J, *et al*. Critical appraisal of assessment of structural damage in ankylosing spondylitis: implications for treatment outcomes. Arthritis Rheum 2008;58:649-56.
32. Appel H, Sieper J. Spondyloarthritis at the crossroads of imaging, pathology, and structural damage in the era of biologics. Curr Rheumatol Rep 2008;10:356-63.
33. Amor B, Dougados M, Mijiyawa M. Classification Criteria of Spondyloarthropathies. Revue du Rhumatisme 1990;57:85-9.
34. 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.
35. Rudwaleit M, van der Heijde D, Khan M, *et al*. How to diagnose axial spondyloarthritis early. Ann Rheum Dis 2004;63:535-43.
36. Rostom S, Dougados M, Gossec L. New tools for diagnosing spondyloarthropathy. Joint Bone Spine 2010;77:108-14.
37. Khan M. Update on spondyloarthropathies. Annals of Internal Medicine 2002;136:896-907.
38. van Tubergen A, Weber U. Diagnosis and classification in spondyloarthritis: identifying a chameleon. Nat Rev Rheumatol 2012;8:253-61.
39. Rudwaleit M, Landewé 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.
40. 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(Suppl2):ii1-44.
41. Rudwaleit M, Jurik A, Hermann K, *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.
42. de Hooge M, van den Berg R, Navarro-Compán V, *et al*. Magnetic resonance imaging (MRI) of the sacroiliac joints in the early detection of spondyloarthritis (SpA): No added value of gadolinium compared to short tau inversion recovery (STIR) sequence. Rheumatology (Oxford) 2013;52:1220-4.
43. Madsen K, Egund N, Jurik A. Grading of inflammatory disease activity in the sacroiliac joints with magnetic resonance imaging: comparison between short-tau inversion recovery and gadolinium contrast-enhanced sequences. J Rheumatol 2010;37:393-400.
44. Maksymowych W, Inman R, Salonen D, *et al*. Spondyloarthritis research Consortium of Canada magnetic resonance imaging index for assessment of sacroiliac joint inflammation in ankylosing spondylitis. Arthritis Rheum 2005;53:703-9.

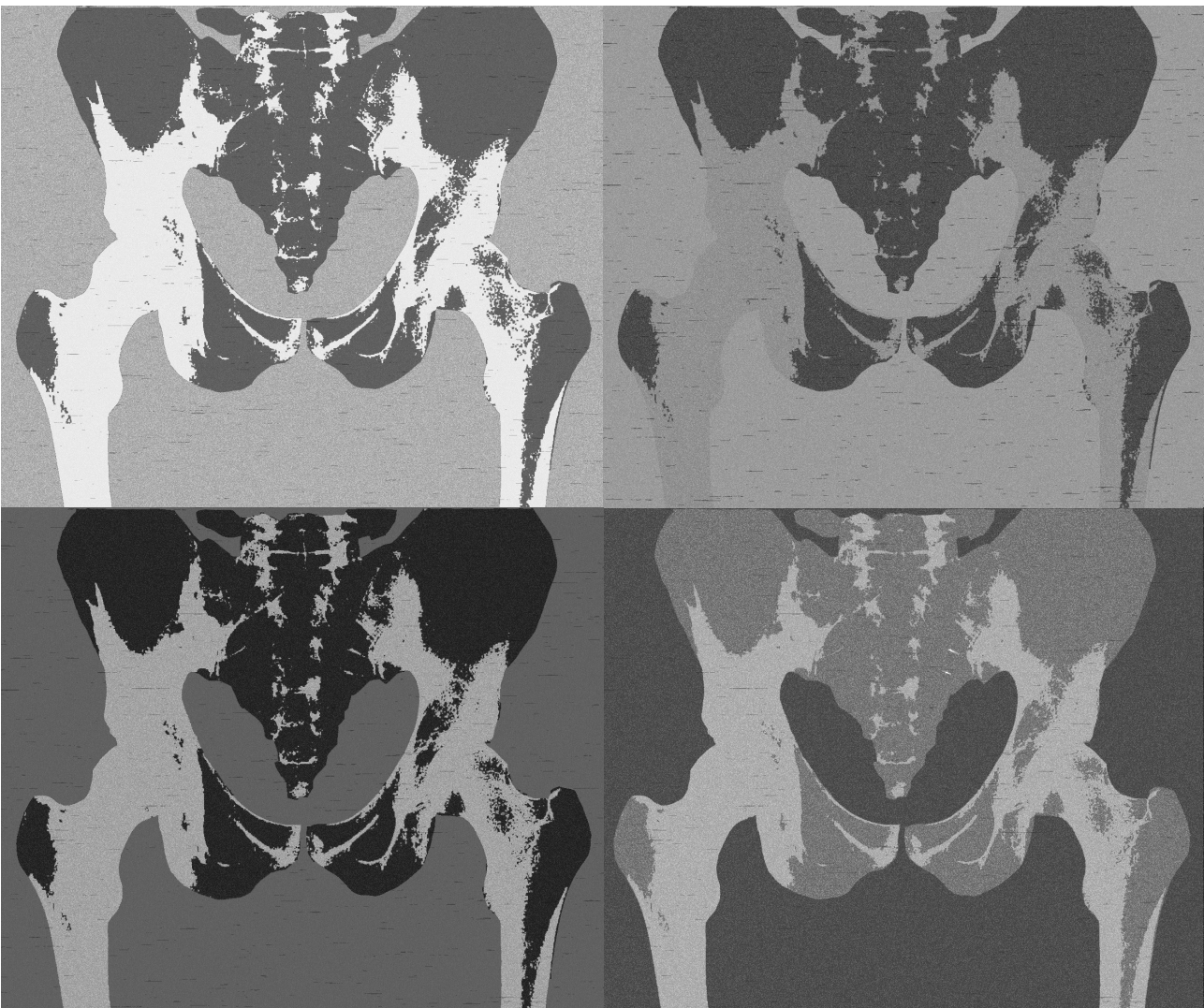
45. Zochling J, van der Heijde D, Burgos-Vargas R, *et al*. ASAS/EULAR recommendations for the management of ankylosing spondylitis. Ann Rheum Dis 2006;65:442-52.
46. http://slides.asas-group.org/app/slides/search?q=&c=17&κ=ALL. (7 May 2014, date last accessed).
47. de Rooy D, van der Linden M, Knevel R, *et al*. Predicting arthritis outcomes - what can be learned from the Leiden Early Arthritis Clinic? Rheumatology (Oxford) 2011;50:93-100.
48. http://slides.asas-group.org/app/slides/search?q=&c=ALL&κ=ALL. (28 May 2014, date last accessed).

2

Percentage of patients with spondyloarthritis in patients referred because of chronic back pain and performance of classification criteria: experience from the SpondyloArthritis Caught Early (SPACE)-cohort

Rheumatology (Oxford) 2013;52:1492-9

R. van den Berg
M. de Hooge
F. van Gaalen
M. Reijnierse
T. Huizinga
D. van der Heijde



Abstract

Objectives

The objectives of the study are to describe the Spondyloarthritis Caught Early (SPACE) cohort, present the performance of various SpA classification criteria and compare patients fulfilling the imaging arm with patients fulfilling the clinical arm of the Assessment of Spondyloarthritis international Society (ASAS) axSpA criteria on demographics, presence of SpA-features and level of disease activity.

Methods

Patients with back pain (≥3 months but ≤2 years, onset <45 years) visiting the rheumatology outpatient clinic of the Leiden University Medical Center were included in the SPACE-cohort. Patients were classified according to the modified New York (mNY), ESSG, Amor and ASAS axSpA criteria. The sensitivity and specificity of criteria were tested against a rheumatologist's diagnosis.

Results

In total, 157 patients were included; 92 patients fulfilled any criteria, 11 fulfilled the mNY (sensitivity 16.9%, specificity 100%), 68 the ESSG (sensitivity 64.6%, specificity 71.7%), 48 the Amor (sensitivity 47.7%, specificity 81.5%) and 60 the ASAS axSpA criteria (sensitivity 84.6%, specificity 94.6%). Of those 60 patients, 30 fulfilled the imaging arm and 30 the clinical arm. Patients in the imaging arm are statistically significantly more often male, have a longer symptom duration and less often a positive family history for SpA than patients fulfilling the clinical arm. Patients in both arms are very similar regarding all other SpA-features and level of disease activity.

Conclusion

The inclusion criteria of the SPACE-cohort yield the same high numbers of SpA patients compared with referral strategies like inflammatory back pain, HLA-B27+ or sacroiliitis, yet are easier to apply. The ASAS axSpA criteria outperformed the other criteria; 38.2% fulfilled the ASAS axSpA criteria. Patients fulfilling the clinical arm of the ASAS axSpA reflect a group of patients similar to those fulfilling the imaging arm.

Introduction

SpA comprises a group of interrelated rheumatic diseases, including AS, PsA and arthritis associated with IBD [1]. The diagnosis is challenging because of the lack of diagnostic criteria for (early) SpA.

Over the years, several criteria sets have been developed to classify patients with SpA. The modified New York (mNY) criteria are available to classify patients with AS [2], however, they are of limited use in early disease or other subtypes of SpA [3]. The ESSG and the Amor criteria are widely used to define the whole concept of SpA [4, 5]. More recently, the Assessment of Spondyloarthritis international Society (ASAS) developed criteria to classify patients with predominantly axial SpA (axSpA) and criteria to classify patients with predominantly peripheral SpA [6, 7]. It is possible to classify patients as having axSpA according to the imaging arm if they have sacroiliitis on radiographs and/or MRI plus at least one additional SpA feature, or according to the clinical arm based on HLA-B27 positivity in combination with at least two other SpA-features [6]. Yet the question arose of whether patients fulfilling the clinical arm reflect a group of patients similar to those fulfilling the imaging arm.

The ASAS axSpA criteria should be applied in patients with back pain (almost daily for ≤3 months, onset <45 years) of unknown origin, which is considered to be the leading symptom of axSpA [8]. However, it is difficult to recognize axSpA in an early stage among the enormous number of patients with back pain, since the clinical presentation of axSpA is very heterogeneous and there is no single shared distinguishing feature [9]. Hence some have stated that not just chronic back pain, but specific inflammatory back pain (IBP) is typical of axSpA [10]. Therefore IBP is often proposed as one of the referral parameters [11, 12]. However, there is increasing evidence that not all patients with axSpA have IBP, and vice versa, which is also evident from the relatively low sensitivity and specificity of IBP criteria (e.g. 79.6% and 72.4%, respectively, for the ASAS IBP criteria) [3, 13-16].

The SpondyloArthritis Caught Early (SPACE) cohort in the Leiden University Medical Center (LUMC) in Leiden, the Netherlands, uses chronic back pain (≥3 months but ≤2 years, onset <45 years) as the only inclusion criteria. These inclusion criteria are, to our knowledge, unique for a SpA cohort. Other early back pain cohorts like ESPAC (the Early SPondyloArthritis Clinic) and DESIR (DEvenir des Spondylarthropathies Indifférenciées Récentes) included only patients with IBP [17, 18].

The goal of this study is to give a description of the characteristics of the patients included in the SPACE-cohort. The percentage of patients fulfilling at least one of the classification criteria sets for SpA is given. Second, the performance of the various classification criteria for SpA is tested. Furthermore, demographics, number of SpA-features and level of disease activity in patients fulfilling the imaging arm and patients fulfilling the clinical arm of the ASAS axSpA criteria are compared.

Patients and methods

Patients

The SpondyloArthritis Caught Early (SPACE) cohort started in January 2009 and is an ongoing project. General practitioners as well as other specialists such as ophthalmologists and gastroenterologists were informed about the start of the SPACE-cohort and about the inclusion criteria. Patients aged 16 years and older with chronic (almost daily) back pain for ≥3 months but ≤2 years with the onset before the age of 45 years referred to the rheumatology outpatient clinic of the LUMC were included after signing informed consent. The SPACE study protocol was approved by the local medical ethics committee of the LUMC. Patients could not be included if other painful conditions not related to SpA could interfere with the evaluation of disease activity or if any reason was present that was likely

to invalidate informed consent or limit the ability of the subject to comply with the protocol requirements.

Assessments and visits

All patients underwent a diagnostic workup at baseline; descriptions of the performed diagnostic workup follow below. Thereafter only patients with definite or possible SpA were included for follow-up visits after 3, 12 and 24 months. Definite axSpA is defined as a patient fulfilling the ASAS axSpA criteria. Possible SpA is defined as the presence of at least one of the following specific SpA-features [high likelihood ratio (LR+) [6, 14]: HLA-B27 positivity, positive family history for SpA, sacroiliitis (MRI or radiographs), acute anterior uveitis] or at least two of the following less-specific SpA-features [lower LR+: IBP (ASAS definition [16]), (heel) enthesitis, peripheral arthritis, psoriasis, IBD, good response to NSAIDs or elevated levels of ESR or CRP], but not fulfilling any of the classification criteria. Annual visits after the first 2 years were scheduled for patients with definite axSpA (ASAS criteria). Unless otherwise specified, all measurements were performed by one of the researchers (RvdB or MdH) during every visit.

Physical examination

In total, 68 joints were examined for tenderness and 66 for swelling. Entheses were examined according to the Maastricht Ankylosing Spondylitis Enthesitis Score (MASES) index [19]. Spinal mobility was assessed by measuring chest expansion, occiput to wall distance, modified Schober test, cervical rotation, lateral spinal flexion and intermalleolar distance as described in the ASAS handbook [20]. The tragus-to-wall distance was derived from the OWD by adding 8 cm to the OWD score. By doing so, the value of zero in the OWD corresponds to a score of zero in the calculation of the BASMI [21]. Based on these measurements, the BASMI was calculated [21].

Patient-reported questionnaires

Patients completed the BASDAI [22] and BASFI [23].

Other parameters

Overall assessment of disease activity was done by the physician on an 11-point numerical rating scale (NRS), 0 representing inactive disease and 10 extremely active disease. The presence (past or current) of extra-spinal and extra-articular manifestations (acute anterior uveitis, urethritis, balanitis, cervicitis, IBD and psoriasis, enthesitis) and a positive family history of SpA (AS, reactive arthritis, psoriasis, IBD, uveitis) all according to the definition of the ASAS criteria [6] was recorded. Treatment with NSAIDs, DMARDs and biologic therapies was recorded. NSAID intake is recorded according to the ASAS recommendations [24]. A good response of back pain to a full dose of NSAID was defined as not present anymore or much better [6]. Furthermore, the Ankylosing Spondylitis Disease Activity Score (ASDAS) was calculated [25]. More information about performed measurements during the visits can be found in the supplementary data, available at Rheumatology Online.

Laboratory assessment

The laboratory assessment during each visit consisted of measurements of ESR (Westergren method in mm/1 h) and CRP (ELISA in mg/l). HLA-B27 was only typed at baseline.

Imaging assessment

MR imaging was performed on a 1.5T (Philips Medical Systems, Best, Netherlands) T1 weighted turbo spin echo (T1TSE) (TR 550/TE 10) and short tau inversion recovery (STIR) (TR 2500/TE 60) sequences were acquired, coronal oblique of the SI joints (MRI-SI). The slice thickness was 4 mm. Radiographs of the pelvis (anterior-posterior view) were performed at baseline, after 1 and 2 years, and thereafter every second year.
SI joints, both on MRI and on radiograph, were independently scored by two trained readers (MdH and RvdB). MRI-SIs were scored on the presence of bone marrow edema (BME) according to the ASAS/OMERACT definition [26], according to the Spondyloarthritis Research Consortium of Canada (SPARCC) score [27] and on the presence of capsulitis/enthesitis. All radiographs of the SI joints (X-SIs) were scored according to the modified mNY criteria [2]. In case the first two readers disagreed on an image [MRI (ASAS/OMERACT definition) or radiograph], a third trained reader (VN) served as adjudicator. If two of three readers scored positive, the image was marked positive. Moreover, all positive X-SIs were checked by a senior rheumatologist (DvdH) who gave a final judgement about the X-SI. All readers were blinded for clinical and laboratory data as well as the results of the other imaging modality.

Diagnosing the patients

A rheumatologist experienced in the field of SpA diagnosed all patients as predominantly axSpA, both axSpA and peripheral SpA, or no SpA based on all collected information, including imaging and HLA-B27 status. For this analysis, patients with only axSpA were used. In the case of no SpA, the rheumatologists filled out another suitable diagnosis. Furthermore, the rheumatologist marked the level of confidence about the diagnosis, either SpA or no SpA, on an 11-point NRS from 0 (not confident at all) to 10 (very confident).

Classification of patients

All patients were classified according to the Amor, ESSG, mNY and ASAS axSpA criteria [2, 4-6]. In addition, both the ESSG and AMOR criteria were modified by judging active sacroiliitis on MRI similarly to radiographic sacroiliitis.

Data analysis

For the present analysis, only data of the baseline visit were used. First, it was investigated how many patients fulfilled at least one of the classification criteria sets for SpA, shown in Venn diagrams.
Next, the number of patients diagnosed as axSpA according to the rheumatologist was described. The diagnosis of the rheumatologist served as external standard to test the performance of the various classification criteria. The performance was determined by calculating sensitivity, specificity, positive likelihood ratio (LR+) and negative likelihood ratio (LR-). For further analyses, the ASAS axSpA criteria set was selected to differentiate between SpA and no SpA patients. Characteristics of the patients were described using t-tests and χ^2-tests. In a following step, the ASAS axSpA criteria were studied in more detail. Patients fulfilling the clinical arm and patients fulfilling the imaging arm were compared on demographics, the presence of SpA-features and level of disease activity. Furthermore, within the imaging arm, patients with sacroiliitis on radiograph were compared with patients with sacroiliitis on MRI only, also by t-tests and χ^2-tests.
Missing values for the presence of SpA-features were interpreted as being absent. All analyses were performed using SPSS version 17. P-values <0.05 were considered significant.

Results

Performance of classification criteria

In total, 157 patients were included in the SPACE-cohort. The mean age at inclusion was 31.2 (SD 12.6) years, the mean symptom duration was 13.5 (SD 7.2) months and 33.1% were male. Of the 157 patients, 92 (58.6%) fulfilled any classification criteria set at baseline. Sixty (38.2%) patients fulfilled the ASAS axSpA criteria; this percentage has been similar over the years the SPACE-cohort has been running (40.4% in 2009, 36.2% in 2010, 38.9% in 2011 and 34.1% in 2012). Thirty-nine of these 60 patients fulfilled at least one other criteria set as well. Sixty-eight (43.3%) patients fulfilled the ESSG criteria; 53/68 fulfilled at least one other criteria set as well. Forty-eight (30.6%) patients fulfilled the Amor criteria; the majority of the patients (45/48) also fulfilled another criteria set. Eleven (7.0%) patients fulfilled the mNY criteria; all fulfilled at least one other classification criteria as well. Nine patients fulfilled all four criteria sets, 15 patients fulfilled three criteria sets (14 the combination of ASAS axSpA, ESSG and Amor and 1 the combination of ASAS axSpA, Amor and mNY) and 38 patients fulfilled two criteria sets (16 both ASAS axSpA and ESSG, 7 both ASAS axSpA and Amor, 14 both ESSG and Amor and 1 both ASAS axSpA and mNY) (figure 1).

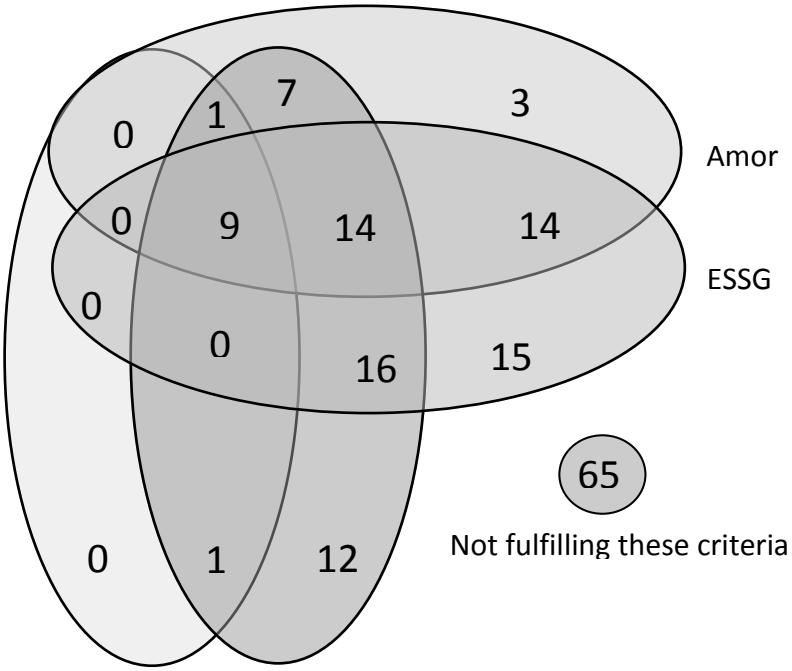


Figure 1: Venn-diagram representing the overlap between the various classification criteria for axial SpA.

To calculate the performance of the various classification criteria, the diagnosis of the rheumatologist was used as external standard. The rheumatologist diagnosed 65 patients (41.4%) as axSpA and 92 patients as no SpA. The mean level of confidence about the diagnosis is similar for patients fulfilling the ESSG, Amor and ASAS axSpA criteria (6.2-6.7 out of 10), but higher for patients fulfilling the mNY criteria (7.8 out of 10) (table 1). The mNY criteria showed the lowest sensitivity (16.9%) but highest specificity (100%). The Amor criteria

showed a sensitivity of 47.7%, which increased to 67.7% in the modified version, without a decrease in specificity (71.7%). The ESSG criteria showed a sensitivity of 64.6%, which increased to 75.4% in the modified version without a decrease in specificity (81.5%). The ASAS axSpA criteria outperformed all other classification criteria, including the modified Amor and modified ESSG criteria, in terms of sensitivity (84.6%), specificity (94.6%), LR+ (15.6) and LR- (0.16) (table 1). For all further analyses we used the ASAS axSpA criteria for the definition if a patient fits into the category axSpA or no SpA. This criterion is exactly defined and reproducible for readers, while the diagnosis by the rheumatologist is not.

Table 1: Performance of the various classification criteria for axial spondyloarthritis with the diagnosis and the level of confidence about the diagnosis of axSpA of rheumatologist as external standard for axSpA versus no SpA.

axSpA patients versus no axSpA patients	axSpA patients (n=65), N positive (sensitivity)	no axSpA patients (n=92), N negative (specificity)	LR+	LR
ASAS axSpA	55 (84.6)	87 (94.6)	15.6	0.16
mNY	11 (16.9)	92 (100)	15.6	0.99
ESSG	42 (64.6)	66 (71.7)	2.3	0.49
Amor	31 (47.7)	75 (81.5)	2.6	0.64
Modified ESSG (with MRI)	49 (75.4)	66 (71.7)	2.7	0.34
Modified Amor (with MRI)	44 (67.7)	75 (81.5)	3.7	0.40

ESSG, European Spondylarthropathy Study Group; ASAS, Assessment of SpondyloArthritis international Society (ASAS); mNY, modified New York; LR+, positive likelihood ratio; LR-, negative likelihood ratio. Level of confidence about the diagnosis SpA on an 11-point NRS from 0 (not confident at all) to 10 (very confident).

Patient characteristics

The majority of the patients referred to the SPACE-cohort were from the Leiden area; over the years, 17.0%, 7.3%, 10.2% and 17.7% of the referrals in 2009, 2010, 2011 and 2012, respectively, were from outside the Leiden area.

Thirty-three patients were not included for follow-up because of the lack of specific SpA-features; 13 patients did not have any SpA-features and the remaining 20 patients had only one less specific SpA feature (1 patient with peripheral arthritis only, 1 patient with heel enthesitis only, 6 patients with a good response to NSAIDs only, 12 patients with IBP only). Of the patients included for follow-up, 64 had possible SpA and the remaining 60 patients fulfilled the ASAS axSpA criteria.

Patients classified as axSpA according to the ASAS axSpA criteria were compared with the group of noaxSpA patients including possible SpA patients and patients excluded for follow-up, revealing some statistically significant differences. AxSpA patients are more frequently male (p=0.001), more often have a positive family history for SpA (p=0.001), IBP (p=0.001), a good response to NSAIDs (p=0.004) and sacroiliitis on radiograph (p<0.001) and MRI (p<0.001), and are more often HLA-B27 positive (p<0.001) compared with no axSpA patients. Furthermore, there was a trend that axSpA patients more often have uveitis (p=0.07) and higher levels of ESR (p=0.08) (table 2).

Table 2: Baseline characteristics of axSpA patients versus no axSpA patients, according to the ASAS axSpA criteria.

	axSpA patients, n=60	no axSpA patients, n=97	P-values axSpA versus no axSpA patients
Age (years) at inclusion, mean ± SD	29.5 ± 8.7	32.3 ± 14.4	0.17
Male, n (%)	29 (48.3)	23 (23.7)	0.001
Duration of back pain (months), mean ± SD	13.4 ± 7.7	13.6 ± 6.9	0.88
HLA-B27 positive, n (%)	47 (79.7)	6 (6.2)	<0.001
Pos. Fam. History SpA, n (%)	31 (51.7)	25 (25.8)	0.001
IBP, n (%)	50 (83.3)	55 (56.7)	0.001
Psoriasis, n (%)	8 (13.3)	8 (8.2)	0.31
Dactylitis, n (%)	3 (5.0)	3 (3.1)	0.55
Enthesitis, n (%)	8 (13.3)	17 (17.5)	0.49
Uveitis, n (%)	9 (15.0)	6 (6.2)	0.07
IBD, n(%)	3 (5.0)	6 (6.2)	0.76
Preceding infection, n (%)	1 (1.7)	1 (1.0)	0.73
CRP (mg/l), mean ± SD	8.4 ± 11.9	5.8 ± 6.9	0.12
ESR (mm/h), mean ± SD	14.4 ± 16.7	10.1 ± 10.6	0.08
Alternating buttock pain, n (%)	16 (26.7)	17 (17.5)	0.17
Good response to NSAIDs, n (%)	29 (48.3)	25 (25.8)	0.004
Elevated CRP/ESR, n (%)	16 (26.7)	15 (15.5)	0.09
Asymmetric lower limb arthritis, n (%)	8 (13.3)	15 (15.5)	0.71
Sacroiliitis radiograph, n (%)	11 (18.3)	1 (1.1)	<0.001
Sacroiliitis MRI, n (%)	25 (41.7)	2 (2.1)	<0.001

IBP, Inflammatory Back Pain; IBD, Inflammatory Bowel Disease; age, age at baseline; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; HLA-B27, Human Leukocyte Antigen; preceding infection can be balinitis, urethritis, cervicitis and/or acute diarrhea.

ASAS imaging arm versus clinical arm

The comparison of patients fulfilling the imaging arm with patients fulfilling the clinical arm revealed that patients in the imaging arm are more often male (p=0.02), have a longer symptom duration (p=0.04) and less often have a positive family history for SpA (p=0.001) than patients fulfilling the clinical arm. However, patients fulfilling the clinical arm reflect a group of patients similar to those fulfilling the imaging arm with respect to the presence of other SpA-features and level of disease activity (table 3). Nevertheless, the mean level of confidence about the diagnosis axSpA in patients fulfilling the clinical arm of the ASAS axSpA criteria (4.9 ± 1.5) is lower in comparison to the level of confidence about the diagnosis in patients fulfilling the imaging arm (7.7 ± 0.8). Within the imaging arm, patients with and without sacroiliitis on radiographs were compared. Remarkably, there was no difference in symptom duration (table 3).

Table 3: Characteristics of patients in the clinical arm compared to patients in the imaging arm of the ASAS axSpA criteria.

	Imaging arm, n=30			Clinical-arm, n=30	P-value imaging arm vs clinical arm
	mNY+, n=11	mNY-, n=19	Total, n=30		
Age (years) at inclusion, mean ± SD	28.6 ± 9.6	32.9 ± 8.7	31.2 ± 9.0	28.2 ± 8.4	0.14
Male, n (%)	8 (72.7)	11 (57.9)	19 (63.3)	10 (33.3)	0.02
Duration of back pain (months), mean ± SD	15.6 ± 8.5	16.0 ± 6.9	15.5 ± 7.6	11.4 ± 7.3	0.04
HLA-B27 positive, n (%)	6 (54.5)	11 (61.1)	17 (58.6)	30 (100)	<0.001
Pos. Fam. History SpA, n (%)	4 (36.4)	5 (26.3)	9 (30.0)	22 (73.3)	0.001
IBP, n (%)	9 (81.8)	14 (73.7)	23 (76.7)	27 (90.0)	0.17
Psoriasis, n (%)	2 (18.2)	2 (10.5)	4 (13.3)	4 (13.3)	1
Dactylitis, n (%)	0 (0.0)	2 (10.5)	2 (6.7)	1 (3.3)	0.55
Enthesitis, n (%)	2 (18.2)	2 (10.5)	4 (13.3)	4 (13.3)	1
Uveitis, n (%)	1 (9.1)	1 (5.3)	2 (6.7)	7 (23.3)	0.07
IBD, n(%)	2 (18.2)	1 (5.3)	3 (10.0)	0 (0.0)	0.08
Preceding infection, n (%)	0 (0.0)	0 (0.0)	0 (0.0)	1 (3.3)	0.31
CRP (mg/l), mean ± SD	6.9 ± 7.2	7.6 ± 8.6	7.3 ± 8.0	15.6 ± 18.9	0.58
ESR (mm/h), mean ± SD	11.4 ± 13.9	14.2 ± 14.8	13.2 ± 14.3	9.4 ± 14.9	0.50
Alternating buttock pain, n (%)	6 (54.5)	5 (26.3)	11 (36.7)	5 (16.7)	0.08
Good response to NSAIDs, n (%)	6 (54.5)	10 (52.6)	16 (53.3)	13 (43.3)	0.44
Elevated CRP/ESR, n (%)	4 (36.4)	5 (26.3)	9 (30.0)	7 (23.3)	0.56
Asymmetric lower limb arthritis, n (%)	0 (0.0)	4 (21.1)	4 (13.3)	4 (13.3)	1
Sacroiliitis radiograph, n (%)	11 (100)	-	11 (36.7)	-	-
Sacroiliitis MRI, n (%)	6 (54.5)[†]	19 (100)[†]	25 (86.2)	-	-
BASDAI	3.7 ± 1.8	4.0 ±2.5	3.9 ± 2.3	3.9 ± 1.9	0.97
ASDAS	2.4 ± 0.7	2.5 ± 0.9	2.4 ± 0.8	2.4 ± 0.9	0.94
BASFI	3.3 ± 1.9	2.4 ± 2.2	2.7 ± 2.1	2.3 ± 2.2	0.50
BASMI	1.9 ± 0.7	1.6 ± 0.5	1.7 ± 0.6	1.6 ± 0.8	0.51
NSAID use, n (%)	9 (81.8)	15 (78.9)	24 (80.0)	22 (73.3)	0.54
DMARD use, n (%)	1 (9.1)	1 (5.3)	2 (6.7)	1 (3.3)	0.55
Biological use, n (%)	0 (0.0)	1 (5.3)	1 (3.3)	0 (0.0)	0.31
Confidence diagnosis axSpA, mean ± SD	7.8 ± 1.1	7.5 ± 0.6	7.7 ± 0.8	4.9 ± 1.5	<0.001

† Statistically significant difference between patients fulfilling the modified New York criteria and patients not fulfilling the modified New York criteria within the total imaging arm. IBP, Inflammatory Back Pain; IBD, Inflammatory Bowel Disease; age, age at baseline; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; HLA-B27, Human Leukocyte Antigen; preceding infection can be balinitis, urethritis, cervicitis and/or acute diarrhea; mNY, modified New York criteria; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; ASDAS, Ankylosing Spondylitis Disease Activity Score; NSAID, Non-Steroidal Anti-Inflammatory Drug; DMARD, Disease Modifying AntiRheumatic Drug. Level of confidence about the diagnosis SpA on an 11-point NRS from 0 (not confident at all) to 10 (very confident).

Discussion

The SPACE-cohort consists of patients with chronic back pain (≥3 months, but ≤2 years, onset <45 years). The only available numbers about the prevalence of chronic back pain (≥3 months duration) in the Netherlands stem from the mid-90s and show a prevalence of 20.8% [28]. The majority of these patients (90%) have non-specific back pain [29]. Hence Dutch rheumatologists in general, and likewise rheumatologists in our department, feared that outpatient clinics would be overloaded by patients with non-specific back pain by using the above-described criteria as the sole referral symptom, although we showed that this fear is unfounded in at least the setting of a tertiary hospital, since ~60% of the patients in the SPACE-cohort fulfill one or more axSpA classification criteria at baseline and 41.4% of patients are directly diagnosed as SpA by the rheumatologist. Moreover, in the light of these results, the value of the numbers about prevalence of chronic back pain from the mid-90s is questionable, thereby indicating that more up-to-date numbers are needed. Furthermore, this percentage of SpA is similar to the percentage of 41.8% found by a muticenter study using a referral strategy consisting of the presence of either IBP or HLA-B27 or sacroiliitis on imaging (MRI and/or radiograph) [11] and the 35.1% found in a study using IBP or a good response to NSAIDs as referral symptom [12]. Although the test result for the presence of HLA-B27 is not difficult to interpret, it is challenging for referring physicians to interpret back pain as inflammatory or not and to detect sacroiliitis, as demonstrated by the low agreement between general practitioners and rheumatologists [11].

It could be argued that our observed prevalence of axSpA is influenced by referral bias; e.g. that due to increased awareness among referring physicians about the SPACE-cohort over time, patients from areas other than the Leiden area are referred to the LUMC or that only patients with a high suspicion of axSpA are referred. However, the percentage of axSpA among all referred patients over the years was similar, and the percentage of referrals from outside the Leiden area was also similar over time. Moreover, 33 of the 157 patients (21.0%) included at baseline had none or only one less specific SpA feature. This indicates, but does not prove, that there is no referral bias, thereby suggesting that the observed prevalence of axSpA could be generalized to primary care. In addition, other studies should investigate the prevalence of SpA among patients with chronic back pain >2 years previously not recognized as SpA.

Around 80% of the axSpA patients in the SPACE-cohort have IBP, thereby confirming that IBP is not present in all SpA patients [13]. Moreover, IBP is frequently (56.7%) present in no SpA patients in the SPACE-cohort, which is consistent with the 45.1% found in another study [11]. These results show that IBP is not a strong discriminating feature and that if IBP was used as an inclusion criterion instead of chronic back pain, 20% of the SpA patients would have been missed.

Depending on the presence and type of SpA-features, patients fulfill various classification criteria. The performance of the Amor, ESSG and ASAS axSpA criteria was better than the mNY criteria at the time of presentation of patients to rheumatologists. This can be explained by the fact that it takes several years before patients develop radiographic sacroiliitis [30].

Moreover, the ASAS axSpA criteria outperformed the Amor and ESSG criteria, even after adding active sacroiliitis (MRI) to the list of SpA-features. These results are in contrast with the results found in a more established cohort [the Cochin Spondyloarthritis (COSPA) cohort] where the ASAS axSpA criteria (fulfilled by 90% of the patients) did not have additional value in comparison to the Amor (fulfilled by 96% of the patients) and ESSG criteria (fulfilled by 83% of the patients) [31]. A possible explanation for these contrasting results is that the longer the symptom duration, the more chance that (extra-articular) features develop. To fulfill the Amor criteria, a patient needs to have at least 6 points representing three to four items. This is quite difficult to reach, especially for patients early in the disease, as in the SPACE-cohort, reflected by the fact that only 31% of these patients fulfilled the Amor criteria. Patients in the COSPA cohort, however, had a mean symptom duration of 16 years (range 8-27 years) and therefore fulfill the Amor criteria more easily.

To fulfill the ESSG criteria, a patient needs to have either IBP or synovitis (asymmetric or predominantly in the lower limbs) and at least one additional feature. The focus of the SPACE-cohort is towards axSpA and not peripheral SpA, and therefore the number of patients with peripheral complaints (synovitis) is low. Furthermore, IBP is only present in about 80% of the axSpA patients in the SPACE-cohort. Therefore it is not possible for some patients to fulfill the ESSG criteria.

It could be argued that the good performance of the ASAS axSpA criteria might be biased by the fact that patients are diagnosed by only one rheumatologist accustomed to work with the ASAS axSpA. However, this bias is unlikely when looking at the level of confidence about the diagnosis, which is similar for patients fulfilling the ESSG, Amor and ASAS axSpA criteria, and when looking at the small numbers of misclassifications by the ASAS axSpA criteria compared with the diagnoses yielded by the modified Berlin algorithm, which is a diagnostic tool [32]. The ASAS axSpA criteria yield 3.8-6.1% of wrongly diagnosed patients as SpA and 7.6-10.2% of missed diagnoses compared with the modified Berlin algorithm. It might even support the rationale to use the ASAS axSpA criteria as diagnostic criteria in this type of setting with referrals to rheumatologists based on chronic back pain starting before the age of 45.

Within the ASAS axSpA criteria, it was questioned whether patients fulfilling the clinical arm of the ASAS axSpA criteria reflect the same disease as patients fulfilling the imaging arm. We found that patients in the SPACE-cohort fulfilling the clinical arm were remarkably similar to patients fulfilling the imaging arm with respect to the presence of most SpA-features and level of disease activity. Another study (ABILITY I trial) found the same results [33]. However, the difference in level of confidence about the diagnosis indicates that the judgement by the rheumatologist is heavily weighted by positive imaging. Furthermore, within the imaging arm of the ASAS axSpA criteria, patients with sacroiliitis on radiographs have the same level of disease activity and symptom duration as patients with sacroiliitis on MRI only.

In conclusion, the inclusion criteria used for the SPACE-cohort, almost daily chronic back pain of short duration (≤2 years) starting before the age of 45 years (in accordance with the entry criteria for the ASAS axSpA criteria), yield the same high number of patients with SpA compared with other referral strategies such as IBP, HLA-B27+ or sacroiliitis, yet are easier to apply. Furthermore, the ASAS axSpA criteria outperformed the other classification criteria; almost 40% fulfilled the ASAS axSpA criteria. Patients fulfilling the clinical arm of the ASAS axSpA reflect a group of patients similar to those fulfilling the imaging arm.

Supplementary data

Supplementary data are available at Rheumatology Online.

References

1. Khan M. Update on spondyloarthropathies. Ann Intern Med 2002;136:896-907.
2. van der Linden S, Valkenburg H, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361-8.
3. Rudwaleit M, Landewé R, van der Heijde D, *et al.* The development of Assessment of SpondyloArthritis international Society classification criteria for axialspondyloarthritis (part I): classification of paper patients by expert opinion including uncertainty appraisal. Ann Rheum Dis 2009;68:770-6.
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. Amor B, Dougados M, Mijiyawa M. Classification criteria of spondyloarthropathies. Rev Rhum Mal Osteoartic 1990; 57:85-9.
6. Rudwaleit M, van der Heijde D, Landewé 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.
7. Rudwaleit M, van der Heijde D, Landewé R, *et al.* The Assessment of SpondyloArthritis international Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. Ann Rheum Dis 2011;70:25-31.
8. Rudwaleit M, Khan M, Sieper J. The challenge of diagnosis and classification in early ankylosing spondylitis: do we need new criteria? Arthritis Rheum 2005;52:1000-8.
9. van Tubergen A, Weber U. Diagnosis and classification in spondyloarthritis: identifying a chameleon. Nat Rev Rheumatol 2012;8:253-61.
10. Braun J, Sieper J. Ankylosing spondylitis. Lancet 2007;369:1379-90.
11. Poddubnyy D, Vahldiek J, Spiller I, *et al.* Evaluation of 2 screening strategies for early identification of patients with axial spondyloarthritis in primary care. J Rheumatol 2011;38:2452-60.
12. Braun A, Saracbasi E, Grifka J, *et al.* Identifying patients with axial spondyloarthritis in primary care: how useful are items indicative of inflammatory back pain? Ann Rheum Dis 2011;70:1782-7.
13. 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.
14. Rudwaleit M, van der Heijde D, Khan M, *et al.* How to diagnose axial spondyloarthritis early. Ann Rheum Dis 2004;63:535-43.
15. Rudwaleit M. New approaches to diagnosis and classification of axial and peripheral spondyloarthritis. Curr Opin Rheumatol 2010;22:375-80.
16. Sieper J, van der Heijde D, Landewé R, *et al.* New criteria for inflammatory back pain in patients with chronic back pain: a real patient exercise by experts from the Assessment of SpondyloArthritis international Society (ASAS). Ann Rheum Dis 2009;68:784-8.
17. Heuft-Dorenbosch L, Landewé R, Weijers R, *et al.* Performance of various criteria sets in patients with inflammatory back pain of short duration; the Maastricht early spondyloarthritis clinic. Ann Rheum Dis 2007;66:92-8.
18. Dougados M, D'Agostino M, 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.
19. Heuft-Dorenbosch L, Spoorenberg A, van Tubergen A, *et al.* Assessment of enthesitis in ankylosing spondylitis. Ann Rheum Dis 2003;62:127-32.
20. 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(Suppl2):ii1-44.
21. van der Heijde D, Landewé 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.

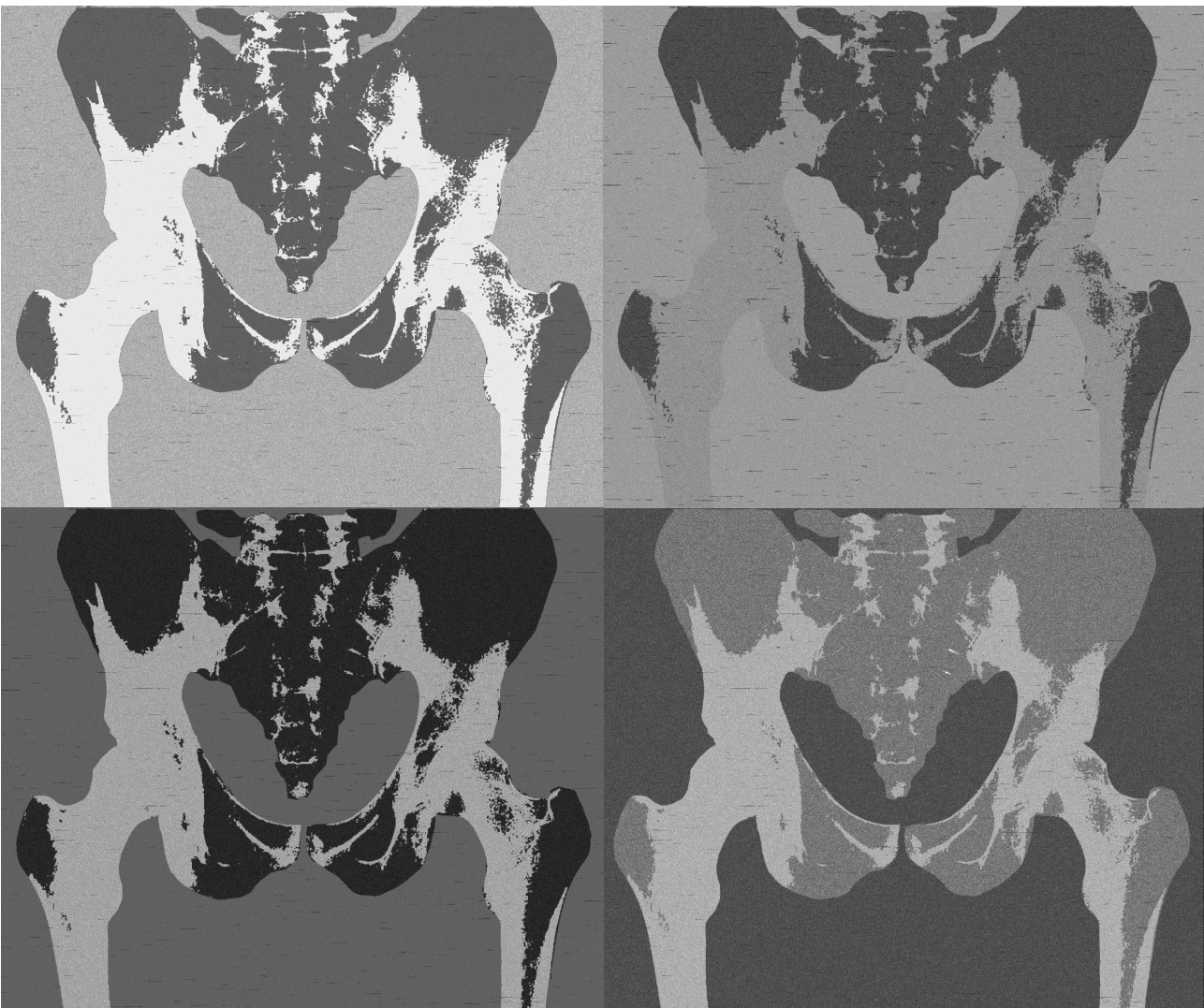
22. Garrett S, Jenkinson T, Kennedy L, *et al.* A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol 1994;21:2286-91.
23. Calin A, Garrett S, Whitelock H, *et al.* A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. J Rheumatol 1994;21:2281-5.
24. Dougados M, Simon P, Braun J, *et al.* ASAS recommendations for collecting, analysing and reporting NSAID intake in clinical trials/ epidemiological studies in axial spondyloarthritis. Ann Rheum Dis 2011;70:249-51.
25. Lukas C, Landewé R, Sieper J, *et al.* Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68:18-24.
26. Rudwaleit M, Jurik A, Hermann K, *et al.* Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification bof axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. Ann Rheum Dis 2009;68:1520-7.
27. Maksymowych W, Inman R, Salonen D, *et al.* Spondyloarthritis Research Consortium of Canada magnetic resonance imaging index for assessment of sacroiliac joint inflammation in ankylosing spondylitis. Arthritis Rheum 2005;53:703-9.
28. Picavet H, Schouten J. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC(3)-study. Pain 2003;102:167-78.
29. Deyo R, Rainville J, Kent D. What can the history and physical examination tell us about low back pain? JAMA 1992;268:760-5.
30. Feldtkeller E, Khan M, van der Heijde D, *et al.* Age at disease onset and diagnosis delay in HLA-B27 negative vs. positive patients with ankylosing spondylitis. Rheumatol Int 2003;23:61-6.
31. Cheung P, Paternotte 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 with criteria sets of amor and the European Spondylarthropathy Study Group. J Rheumatol 2012;39:816-21.
32. van den Berg R, de Hooge M, Rudwaleit M, *et al.* ASAS modification of the Berlin algorithm for diagnosing axial spondyloarthritis: results from the SPondyloArthritis Caught Early (SPACE)-cohort and from the Assessment of SpondyloArthritis international Society (ASAS)-cohort. Ann Rheum Dis 2013;72:1646-53.
33. Sieper J, van der Heijde D, Dougados M, *et al.* Comparable levels of disease activity between axial spondyloarthritis patients classified through the MRI or HLA-B27 pathway of the ASAS criteria. Ann Rheum Dis 2011;70(Suppl3):332.

3

ASAS modification of the Berlin algorithm for diagnosing axial spondyloarthritis: results from the SPondyloArthritis Caught Early (SPACE)-cohort and from the Assessment of SpondyloArthritis international Society (ASAS)-cohort

Ann Rheum Dis 2013;72:1645-53

R. van den Berg
M. de Hooge
M. Rudwaleit
J. Sieper
F. van Gaalen
M. Reijnierse
R. Landewé
T. Huizinga
D. van der Heijde



Abstract

Objective

To compare the original Berlin algorithm for diagnosing axial Spondyloarthritis (axSpA) with two modifications in the SPondyloArthritis Caught Early (SPACE)- cohort and the Assessment of SpondyloArthritis international Society (ASAS) axSpA criteria validation (ASAS)-cohort.

Methods

Patients in the SPACE-cohort (back pain ≥3 months, ≤2 years, onset <45 years) and the ASAS-cohort (undiagnosed chronic back pain) were diagnosed according to three algorithms: original (inflammatory back pain (IBP) mandatory), modification 1 (IBP defined by ≥3/5 IBP-features instead of ≥4/5) and modification 2 (IBP deleted as obligatory entry criterion, added as SpA-feature). Diagnosis by rheumatologist, ASAS axSpA criteria and likelihood ratio product were used as external standards to test the performance of the algorithms.

Results

SPACE-cohort: Compared to the diagnosis by rheumatologist (either axSpA or no axSpA), the original algorithm agreed in 120 patients (76.4%). Agreement decreased using modification 1 (119 patients; 75.8%), increased using modification 2 (125 patients; 79.6%). Sensitivity increased from 66.2% (original) to 72.3% (modification 1) and 78.5% (modification 2). Specificity decreased more using modification 1 (83.7% to 78.3%) than when using modification 2 (83.7% to 79.6%).
ASAS-cohort: Compared to the diagnosis by rheumatologist (either axSpA or no axSpA), the original algorithm agreed in 484 patients (70.7%). Agreement increased using modification 1 (520 patients; 75.9%) and modification 2 (548 patients; 80.0%). Sensitivity increased from 65.3% (original) to 77.9% (modification 1) and 79.6% (modification 2). Specificity decreased more using modification 1 (79.2% to 72.2%) than when using modification 2 (79.2% to 75.6%).

Conclusions

ASAS accepted a modified algorithm for diagnosing axSpA in which IBP is excluded as obligatory entry criterion and added as SpA-feature.

Introduction

Spondyloarthritis (SpA) consists of a heterogeneous group of inter-related rheumatic diseases, divided into categories according to the predominant site of involvement: axial SpA (axSpA) or peripheral SpA. AxSpA is the overall umbrella term for both patients with damage visible on radiographs of the sacroiliac joints (X-SI) and nonradiographic axSpA. The heterogeneity of SpA makes early detection challenging [1]. A helpful tool in the early diagnosis of axSpA is the Berlin diagnostic algorithm; a decision tree applicable to patients with inflammatory back pain (IBP).

The algorithm is fully based on data from the literature on the sensitivity and specificity of characteristic SpA-features. The likelihood ratio (LR)-product of (past or current) SpA-features is calculated for each patient as they follow the algorithm taking into account the a priori probability of SpA, thereby avoiding unnecessary diagnostic tests. The algorithm consists of several diagnostic steps, of which assessment of IBP is the first critical step. Patients may follow the algorithm in various ways depending on whether they have sacroiliitis on x-ray, the number of (past or current) SpA-features, human leukocyte antigen (HLA)-B27 positivity and sacroiliitis on MRI.

Since only 70–80% of patients with axial SpA have typical IBP symptoms, IBP as an obligatory entry criterion in the algorithm has some limitations because patients with axSpA but without IBP will not be captured [2–5]. To circumvent this limitation, it was proposed in 2004 that in back pain patients without IBP other causes of back pain should be considered in general, unless SpA is suspected because of the presence of other SpA-features. This recommendation, however, was not further specified in the original algorithm.

This has stimulated us to test two modifications of the algorithm in two independent cohorts; an observational inception cohort including patients with chronic back pain (the SPondyloArthritis Caught Early (SPACE)-cohort) and a larger, international cohort created for the validation of the new Assessment of SpondyloArthritis international Society (ASAS) axSpA criteria (the ASAS-cohort).

Methods

SPACE-cohort

Patients with chronic (almost daily) back pain for ≥3 months but ≤2 years, with the onset <45 years University Medical Center, were included in the SPACE-cohort since January 2009. At baseline, patients underwent a diagnostic work-up consisting of physical examination, MRI and X-rays of the SI-joints (MRI-SI and X-SI) and laboratory assessments including HLA-B27 testing (online supplementary text 1). Furthermore, the presence of SpA-features is recorded (online supplementary table S1) [2]. After that, a rheumatologist experienced in SpA diagnosed all patients as having SpA or no SpA.

All MRI-SIs and X-SIs were independently scored by two trained readers (MdH and RvdB) according to the ASAS/ OMERACT definition (MRI-SI) [6], and the modified New York (mNY) criteria (X-SI) [7]. A third trained reader (VNC) served as adjudicator and scored only the images in which the first two readers disagreed. If two/three readers scored positive, the image was scored accordingly. All readers were blinded for clinical and laboratory data, and for the results of the other imaging method.

ASAS-cohort

The ASAS-cohort was compiled for the validation of the new classification criteria for axSpA. Patients with chronic back pain of ≥3 months with onset <45 years and with a suspicion of SpA but without a definite diagnosis were included and assessed according to a fixed protocol by rheumatologists who are experts in the field of SpA.

Complete and detailed data collection of the ASAS-cohort has been described before [8]. This included assessment of (past or current) SpA-features [2], C-reactive protein, and HLA-B27 typing. Plain radiographs of the pelvis were taken in all patients. The local rheumatologist and/or radiologist assessed sacroiliitis on X-SI (mNY criteria) [9], and the presence or absence of typical signs of active inflammation on MRI-SI [8].

Diagnosis of patients according to the Berlin algorithm

According to the original algorithm (figures 1A and 2A), patients were diagnosed as having axSpA if they had IBP and ≥3 SpA-features, or if patients had IBP with 1-2 SpA-features and were HLA-B27 positive. Patients with no other SpA-features besides IBP could only be diagnosed as having axSpA if both HLA-B27 and active sacroiliitis (MRI-SI) were present.
In the original algorithm, IBP was defined according to the Calin criteria [10]. In the SPACE-cohort and ASAS-cohort, however, IBP was defined according to the 'ASAS expert criteria', which are slightly more specific [5].
Subsequently, two modifications of the algorithm were constructed. In modification 1, fulfilllment of the ASAS IBP criteria [11], was adapted (figures 1B and 2B). The IBP criteria are: onset of back pain before the age of 40, insidious onset, improvement of back pain with exercise, no improvement of back pain with rest and pain at night with improvement upon getting up [5]. Patients who fulfilled ≥3 IBP criteria instead of ≥4 out of 5 criteria could now be diagnosed as having IBP. During validation of these ASAS IBP criteria sensitivity (79.6%) and specificity (72.4%) were found to be best when patients fulfilled ≥4/5 criteria, a higher sensitivity (95.1%) was reached at the cost of specificity (47.5%) if ≥3/5 criteria for IBP were considered sufficient [5, 12].
Modification 2 slightly changed the structure and the set of SpA-features by deleting IBP as obligatory entry criterion, and adding it as SpA-feature. This resulted in three entry groups based on the requirement of ≥4, 2-3 and 0-1 SpA-features (figures 1C and 2C). All patients were diagnosed according to the three algorithms.

Statistical methods

The disease probability in each patient was calculated by multiplying the individual likelihood ratios (LRs) of all identified SpA-features. An LR-product of 79 results in a positive predictive value of 80% in patients with chronic back pain with an assumed disease prevalence of axSpA of 5% [2]. Missing values for the presence of SpA-features were interpreted as being absent and were included in the following analyses with the missing values set as 'negative'. Because of the lack of a true gold standard, the fulfilllment of the ASAS axSpA criteria [7], the disease probability based on the likelihood ratio (LR)-product [13], and the diagnosis by the rheumatologist were used as external standards to test the performance of the algorithms. The performance was assessed by calculating the sensitivity, specificity, percentage of agreement on the diagnosis as well as the percentage of patients erroneously diagnosed as axSpA and/or diagnosis of axSpA missed by the algorithm.

Results

Baseline characteristics

SPACE-cohort

In total, 157 patients were included in the analyses of the SPACE-cohort. The rheumatologist diagnosed axSpA in 65/157 (41.4%) of the patients. Characteristics are presented in table 1.

Table 1: Baseline characteristics of patients in the SPACE-cohort and the ASAS-cohort; SpA versus no SpA based on the diagnosis of the rheumatologist.

	SPACE-cohort*			ASAS-cohort*		
	axSpA (n=65)	**no SpA (n=92)**	**P-value**	**axSpA (n=421)**	**no SpA (n=264)**	**P-value**
Age (years) at inclusion, mean ± SD	31.5 ± 16.6	31.1 ± 8.8	0.86	31.0 ± 10.8	35.8 ± 10.5	0.839
Male, n (%)	29 (44.6)	23 (25.0)	0.01	225 (53.4)	87 (33.0)	<0.001
Duration of back pain, mean ± SD	13.4 ± 7.4 (months)	13.7 ± 7.1 (months)	0.79	6.3 ± 7.8 (years)	9.3 ± 10.7 (years)	0.792
HLA-B27 positive, n (%)	44 (67.7)	9 (9.8)	<0.001	270 (64.1)	73 (27.7)	<0.001
Pos. fam. history SpA, n (%)	31 (47.7)	25 (27.2)	0.01	106 (25.2)	52 (19.7)	0.097
IBP, n (%)	52 (80.0)	53 (57.6)	0.003	324 (77.0)	125 (47.3)	<0.001
Psoriasis, n (%)	10 (15.4)	6 (6.5)	0.07	36 (8.6)	13 (4.9)	0.073
Dactylitis, n (%)	4 (6.2)	2 (2.2)	0.20	28 (6.7)	5 (1.9)	0.005
Enthesitis, n (%)	10 (15.4)	15 (16.3)	0.88	86 (20.4)	38 (14.4)	0.046
Uveitis, n (%)	10 (15.4)	5 (5.4)	0.04	43 (10.2)	21 (8.0)	0.323
IBD, n(%)	4 (6.2)	5 (5.4)	0.85	14 (3.3)	4 (1.5)	0.149
Preceding infection, n (%)	2 (3.1)	0 (0.0)	0.09	12 (0.17)	5 (0.14)	0.434
CRP (mg/l), mean ± SD	8.3 ± 11.6	5.7 ± 6.9	0.11	7.1 ± 14.9	2.4 ± 4.4	<0.001
ESR (mm/h), mean ± SD	13.6 ± 16.3	10.4 ± 10.7	0.17	#	#	
Alternating buttock pain, n (%)	15 (23.1)	18 (19.6)	0.60	174 (41.3)	65 (24.6)	<0.001
Good response to NSAIDs, n (%)	27 (41.5)	27 (29.3)	0.11	259 (61.5)	73 (27.7)	<0.001
Elevated CRP/ESR, n (%)	15 (23.1)	16 (17.4)	0.38	170 (40.4)	43 (16.3)	<0.001
Arthritis, n (%)	13 (20.0)	10 (10.9)	0.11	171(40.6)	59 (22.3)	<0.001
Sacroiliitis X-ray, n (%)	11 (16.9)	1 (1.1)	<0.001	123 (29.2)	9 (3.4)	<0.001
Sacroiliitis MRI, n (%)	27 (41.5)	0 (0.0)	<0.001	202 (48)	8 (3)	<0.001

* Diagnosis according to rheumatologist. # Not estimated in ASAS-cohort. P-values <0.05 are defined statistically significant. HLA-B27, Human Leukocyte Antigen; IBP, Inflammatory Back Pain; preceding infection can be balinitis, urethritis, cervicitis and/or acute diarrhea; CRP, C-reactive protein; IBD, Inflammatory Bowel Disease; ESR, erythrocyte sedimentation rate; MRI, magnetic resonance imaging.

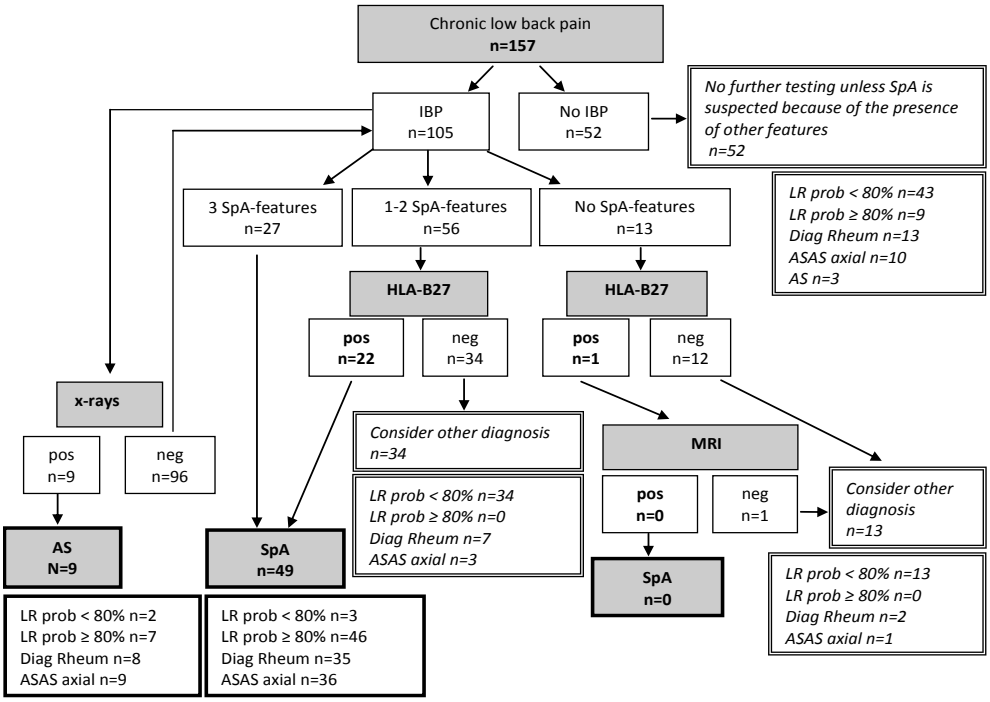


Figure 1a: Original Berlin algorithm (SPACE-cohort).

ASAS-cohort

From the 685 patients of the ASAS-cohort used in this study, 421 (61.5%) were diagnosed as axSpA by the rheumatologist. Characteristics are presented in table 1.

Diagnosis by the algorithms

SPACE-cohort

According to the original algorithm, 58 patients were diagnosed as having axSpA. Nine of them were diagnosed as ankylosing spondylitis (AS), based on the presence of radiographic sacroiliitis, 27 patients had axSpA based on clinical grounds (≥3 SpA-features present), the remaining 22 patients were HLA-B27 positive with 1-2 SpA-features present (figure 1A).
According to modification 1, 22 patients immediately leave the algorithm. A total of 56 patients are diagnosed as having axSpA: 11 patients are directly diagnosed as AS, 29 patients had ≥3 SpA-features and 27 HLA-B27 positive patients had 1-2 SpA-features (figure 1B).
In modification 2, 69 patients were diagnosed as having axSpA: 12 patients were diagnosed as AS, 27 patients had ≥4 SpA-features and 29 patients were HLA-B27 positive and had 2-3 SpA-features. In addition, there was one patient with 0-1 SpA-features, HLA-B27 positivity and a positive MRI-SI who was diagnosed as having SpA (figure 1C).

Table 2: Sensitivity, specificity, percentage of axSpA diagnosis missed and erroneously diagnoses of axSpA by the algorithm in the SPACE-cohort and ASAS-cohort according to the three external standards ASAS axial SpA criteria, LR-product probability ≥ 80% and diagnosis rheumatologist.

	Sensitivity (%)	Specificity (%)	Correct classified (%)	False-negatives (%)	False-positives (%)
SPACE					
ASAS axial SpA					
Original	72.6	86.3	80.9	10.8	8.3
Modification 1	81.7	81.4	81.5	7.0	11.5
Modification 2	89.8	83.7	86.0	3.8	10.2
LR-product probability ≥80%					
Original	85.5	94.7	91.1	5.7	3.2
Modification 1	92.4	93.4	93.0	3.2	3.8
Modification 2	100	92.6	95.5	0.0	4.5
Diagnosis rheumatologist					
Original	66.2	83.7	76.4	14.0	9.6
Modification 1	72.3	78.3	75.8	11.5	12.7
Modification 2	78.5	80.4	79.6	8.9	11.5
ASAS					
ASAS axial SpA					
Original	72.6	84.1	77.5	15.6	6.9
Modification 1	86.7	78.3	83.1	7.6	9.3
Modification 2	89.4	83.0	88.8	6.1	7.6
LR-product probability ≥80%					
Original	83.5	99.3	90.2	9.5	0.3
Modification 1	96.0	85.6	91.2	2.2	6.6
Modification 2	97.2	90.7	96.6	1.6	4.2
Diagnosis rheumatologist					
Original	65.3	79.2	70.7	21.3	8.0
Modification 1	77.9	72.2	75.9	13.6	10.8
Modification 2	79.6	75.6	80.0	12.7	9.8

Modification 1: IBP 3/5 instead of 4/5. Modification 2: IBP as additional SpA-feature instead of entry criterion. LR-product: Likelihood Ratio-product.

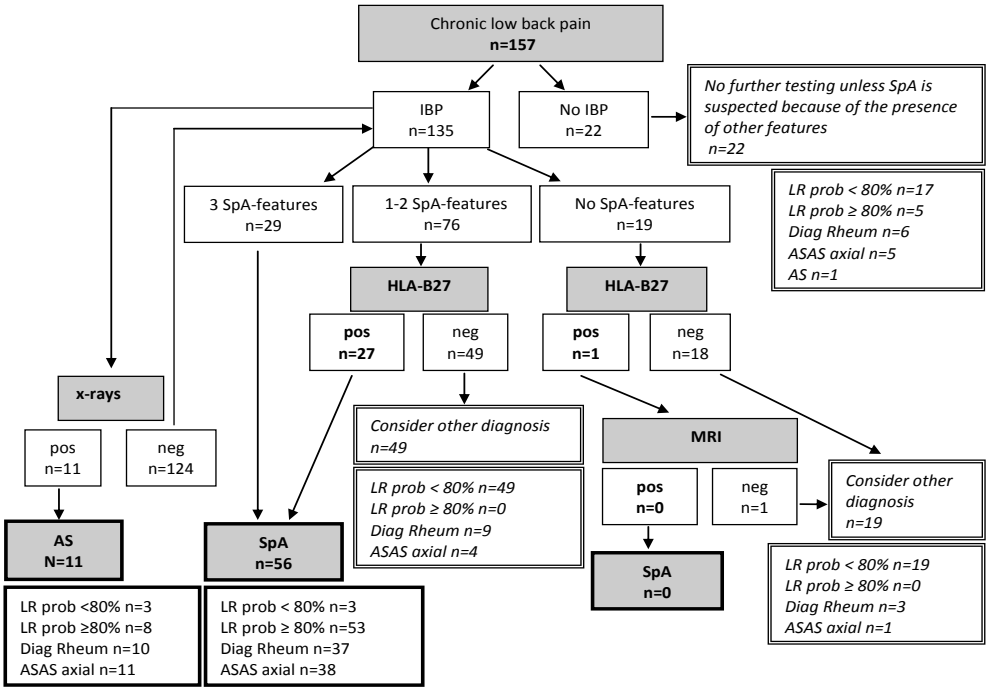


Figure 1b: Modification 1 of the Berlin algorithm; IBP defined when 3 out of 5 criteria are fulfilled instead of 4 out of 5 criteria (SPACE-cohort).

In 120 patients (76.4%) the diagnosis of the rheumatologist and the original algorithm agreed. Modification 1 (IBP 3/5) diagnosed nine more patients as having axSpA, and modification 2 (IBP excluded as obligatory entry criterion) diagnosed 11 more patients as having axSpA, resulting in agreement with the diagnosis of the rheumatologist in 119 (75.8%) and 125 patients (79.6%) respectively (table 2). Compared to the diagnosis of the rheumatologist as external standard, sensitivity was 66.2% using the original algorithm. Sensitivity was higher, 77.9% (+11.7% compared to the original algorithm) using modification 1 and increased more using modification 2, 79.6% (+13.4%). Yet specificity slightly decreased from 83.7% using the original algorithm to 78.3% (−5.4%) using modification 1 and to 80.4% (−3.3%) using modification 2. The same trend was observed compared to the other external standards. The best balance between sensitivity and specificity is present in modification 2 (table 2).

ASAS-cohort

In the original algorithm (figure 2A), 236 patients immediately leave the algorithm. Out of the 449 patients that continue in the algorithm, 330 were diagnosed as having axSpA: 102 fulfilled the mNY criteria for AS, 138 patients with ≥3 SpA-features and another 86 HLA-B27 positive patients with 1-2 SpA-features are diagnosed as having axSpA. In addition, four HLA-B27 positive patients with active sacroiliitis (MRI-SI), but without other SpA-features are diagnosed as having axSpA. In all HLA-B27 negative patients with 1-2 SpA-features (n=93), patients without SpA-features (n=19) and patients without sacroiliitis (MRI-SI) (n=7), the algorithm suggests another diagnosis than axSpA.

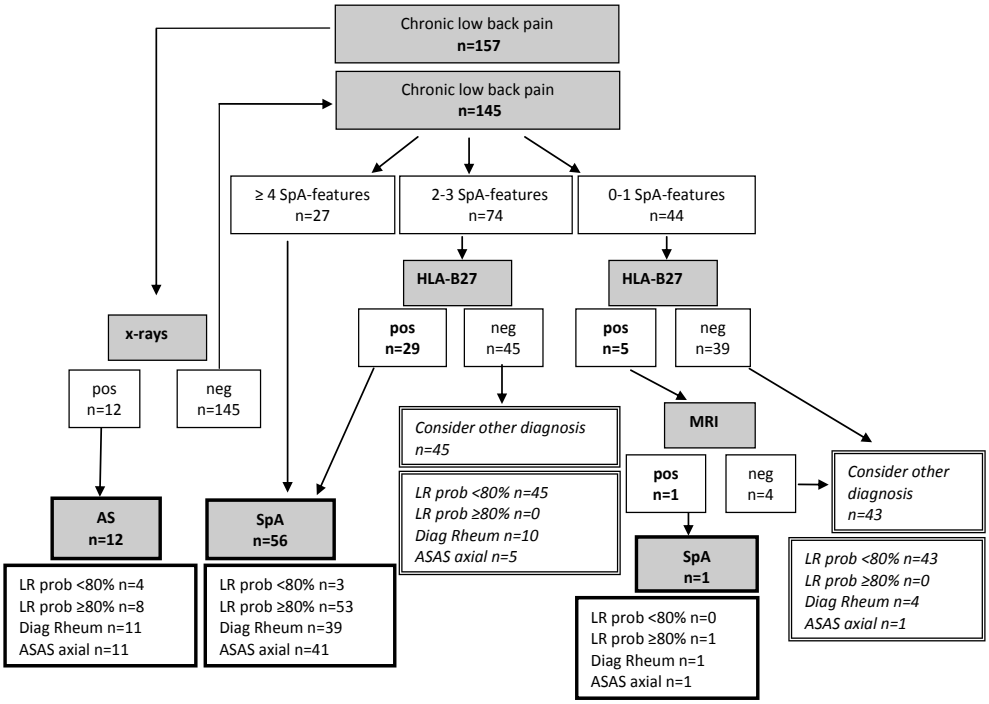


Figure 1c: Modification 2 of the Berlin algorithm; IBP is deleted as entry criterion and implemented as an additional SpA-feature (SPACE-cohort).

In modification 1 (figure 2B), 113 patients immediately leave the algorithm and 402 patients are diagnosed as having axSpA: 122 patients are directly diagnosed as AS, 164 patients with ≥3 SpA-features, 111 HLA-B27 positive patients with 1-2 SpA-features and five HLA-B27 positive patients with a positive MRI-SI but without SpA-features. In 150 HLA-B27 negative patients and in 10 HLA-B27 positive patients with a negative MRI-SI, the algorithm suggested another diagnosis than axSpA.
In modification 2 (figure 2C), the number of patients immediately leaving the algorithm is reduced to 17 patients. In total, 407 patients are diagnosed as having axSpA. Of those, 132 patients are directly diagnosed as AS, 148 patients with ≥4 SpA-features are diagnosed as having axSpA, as were 115 HLA-B27 positive patients with 2-3 SpA-features and 12 HLA-B27 positive patients with a positive MRI-SI and 0-1 SpA-features. For the remaining 278 patients another diagnosis than axSpA should be considered.
The rheumatologist diagnosis and the original algorithm agreed in 70.7% of the patients. Modification 1 showed agreement with the diagnosis of the rheumatologist in 75.9% of the patients (+5.2% compared to the original algorithm). Modification 2 showed a similar trend; 80% (+9.3% compared to the original algorithm) agreement. Sensitivity increased from 65.3% in the original algorithm to 77.9% (+12.6%) in modification 1 and 79.6% (+14.3%) in modification 2, when using the diagnosis of the rheumatologist as external standard. Specificity decreased from 79.2% in the original algorithm to 72.2% (−7.0%) in modification 1 and to 75.6% (−3.6%) in modification 2 (table 2). The performance of the three algorithms with the ASAS axSpA criteria and the LR-product as external standard are also presented in table 2 and show similar results.

Asymmetrical arthritis

In additional calculations on the performance, we replaced the SpA-feature 'peripheral arthritis' with 'asymmetrical arthritis preferentially of the lower limbs' (only performed in the ASAS-cohort). When doing so, sensitivity decreased while specificity increased in all three algorithms (online supplementary text 2).

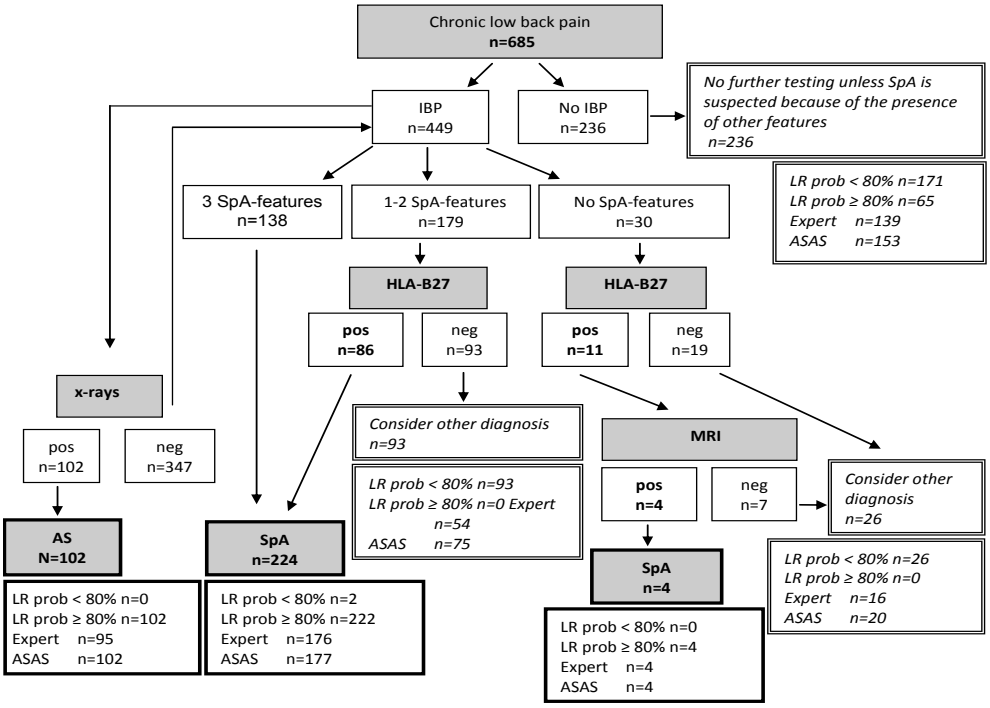


Figure 2a: Original Berlin algorithm (ASAS-cohort).

Reasons for misdiagnoses

SPACE-cohort

Compared to the diagnosis of the rheumatologist as external standard, 15 patients (9.6%) were erroneously diagnosed as having axSpA by the original algorithm and in 22 patients (14.0%) the diagnosis axSpA was missed by the algorithm, especially in the group of patients without IBP. In both modifications, a few more patients, n=20 by modification 1 and n=18 by modification 2, were erroneously diagnosed as having axSpA (12.7% (+3.2% compared to the original algorithm) by modification 1 and 11.5% (+1.9%) by modification 2) but the number of patients in which the diagnosis axSpA was missed dropped to 18 (11.5% by modification 1 (−3.8% compared to the original algorithm)) and 14 patients (8.9% by modification 2 (−7%)) (table 2).

Table 3: Misdiagnoses by the three versions of the Berlin algorithm in the SPACE-cohort (diagnosis of rheumatologist is used as external standard).

Sacroiliitis (MRI and/or X-ray)	HLA-B27 pos. or neg.	No. SpA-features	Missed diagnoses axSpA (%)	Erroneous diagnoses axSpA (%)
Original algorithm*			**n=22**	**n=15**
Imaging+	HLA-B27+	≥3	-	-
		1-2	3	-
		0	2	-
	HLA-B27-	≥3	1	1
		1-2	5	-
		0	3	-
Imaging-	HLA-B27+	≥3	1	-
		1-2	1	3
		0	1	-
	HLA-B27-	≥3	-	11
		1-2	5	-
		0	-	-
Modification 1 (IBP 3/5 instead of 4/5)*			**n=18**	**n=20**
Imaging+	HLA-B27+	≥3	-	-
		1-2	2	-
		0	-	-
	HLA-B27-	≥3	1	1
		1-2	5	-
		0	3	-
Imaging-	HLA-B27+	≥3	1	-
		1-2	-	6
		0	1	-
	HLA-B27-	≥3	-	13
		1-2	5	-
		0	-	-
Modification 2 (IBP excluded as entry criterion)		**No. SpA-features, including IBP**	**n=14**	**n=18**
Imaging+	HLA-B27+	≥4	-	-
		2-3	-	-
		0-1	-	-
	HLA-B27-	≥4	-	1
		2-3	6	-
		0-1	2	-
Imaging-	HLA-B27+	≥4	-	-
		2-3	-	6
		0-1	1	-
	HLA-B27-	≥4	-	11
		2-3	4	-
		0-1	1	-

*Patients following the original cohort and modification 1 have IBP in addition to the other SpA-features, otherwise they did not enter the algorithm, except the patients that are excluded because they had no IBP. Imaging +: sacroiliitis present on MRI and/or X-rays. Imaging -: no sacroiliitis present on MRI and/or X-rays. HLA-B27: Human Leukocyte Antigen. A list of SpA-features is given in table S1 (online supplementary material).

Table 4: Misdiagnoses by the three versions of the Berlin algorithm in the ASAS-cohort (expert opinion is used as external standard).

Sacroiliitis (MRI and/or X-ray)	HLA-B27 positive or negative	No. SpA-features	Missed diagnoses of axSpA (%)	Erroneous diagnoses of axSpA (%)
Original algorithm*			**N=146**	**n=55**
Imaging+	HLA-B27+	≥3	12	1
		1-2	17	-
		0	2	-
	HLA-B27-	≥3	7	-
		1-2	25	1
		0	8	-
Imaging-	HLA-B27+	≥3	10	-
		1-2	13	24
		0	3	-
	HLA-B27-	≥3	9	16
		1-2	37	2
		0	3	-
Modification 1 (IBP 3/5 instead of 4/5)*			**N=93**	**n=74**
Imaging+	HLA-B27+	≥3	4	2
		1-2	4	1
		0	1	-
	HLA-B27-	≥3	-	-
		1-2	24	1
		0	8	-
Imaging-	HLA-B27+	≥3	2	13
		1-2	4	34
		0	3	-
	HLA-B27-	≥3	3	21
		1-2	37	2
		0	3	-
Modification 2 (IBP excluded as entry criterion)		**No. SpA-features, including IBP**	**N=87**	**n=67**
Imaging+	HLA-B27+	≥4	-	2
		2-3	-	1
		0-1	-	-
	HLA-B27-	≥4	-	-
		2-3	25	1
		0-1	11	-
Imaging-	HLA-B27+	≥4	1	11
		2-3	1	32
		0-1	5	-
	HLA-B27-	≥4	1	17
		2-3	33	2
		0-1	10	1

*Patients following the original cohort and modification 1 have IBP in addition to the other SpA-features, otherwise they did not enter the algorithm, except the patients that are excluded because they had no IBP. Imaging +: sacroiliitis present on MRI and/or X-rays. Imaging -: no sacroiliitis present on MRI and/or X-rays. HLA-B27: Human Leukocyte Antigen. A list of SpA-features is given in table S1 (online supplementary material).

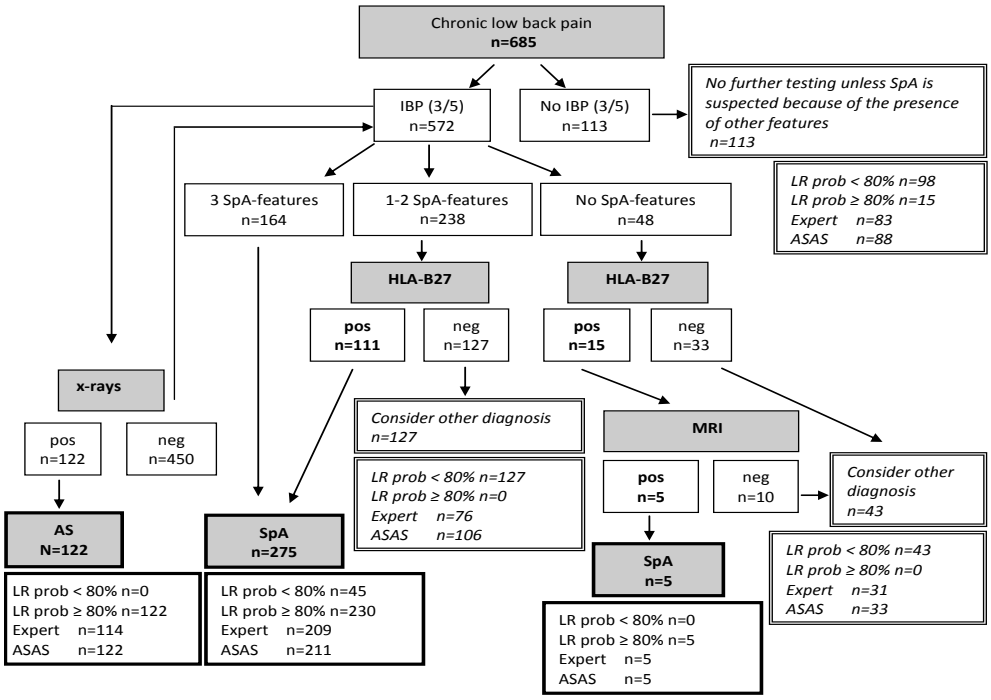


Figure 2b: Modification 1 of the Berlin algorithm; IBP defined when 3 out of 5 criteria are fulfilled instead of 4 out of 5 criteria (ASAS-cohort).

An extensive description of all misdiagnoses is given in table 3. Most patients who were erroneously diagnosed with axSpA by the algorithm have ≥4 SpA-features (including IBP) and were therefore diagnosed as SpA according to the algorithm, but are HLA-B27 negative and do not have sacroiliitis (X-SI or MRI-SI) and are not considered as having axSpA according to the rheumatologist. This pattern was seen in all three algorithms. Most patients in whom the diagnosis axSpA was missed by the algorithm are HLA-B27 negative, and have ≤3 SpA-features (including IBP) and were therefore diagnosed as no axSpA according to the algorithm. However, those patients do have sacroiliitis (MRI-SI), which is missed by the algorithm since the patients were excluded before the MRI-step. Again, this pattern was seen in all three algorithms.

ASAS-cohort

Table 2 also shows the misdiagnoses of the algorithms in the ASAS-cohort.
Using the rheumatologist diagnosis as external standard, 8.0% of the patients were erroneously diagnosed as axSpA and in 21.3% of the patients the diagnosis axSpA was missed by the original algorithm. Modification 1 showed in 10.8% (+2.8% compared to the original algorithm) of the patients an erroneous diagnosis of axSpA and in 13.6% (−7.7%) of the patients the diagnosis of axSpA was missed. Modification 2 showed a similar trend; 9.8% (+1.8%) of the patients were erroneously diagnosed as axSpA and in 12.7% (−8.6%) of the patients the diagnosis axSpAwas missed by the algorithm.
As shown in table 4, the majority of the patients (n=53) erroneously diagnosed as axSpA, have a negative MRI-SI. Two third of these patients (n=35) are HLA-B27 positive with one or more SpA-features present. This trend is seen in all three algorithms.

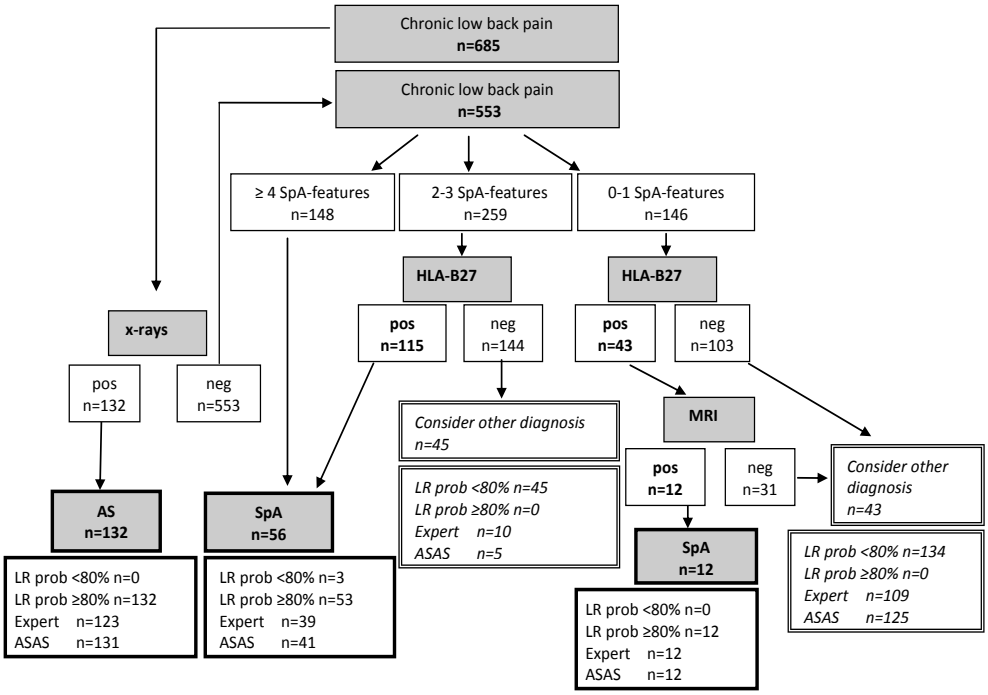


Figure 2c: Modification 2 of the Berlin algorithm; IBP is deleted as entry criterion and implemented as an additional SpA-feature (ASAS-cohort).

Most of the patients in whom the diagnosis of axSpA was missed by the algorithm are HLA-B27 negative (n=89). Almost half of these misdiagnosed patients (n=40) have a positive MRI-SI. Again, this trend is also seen using the two modifications.

Discussion

In this study we investigated the performance of the original Berlin algorithm and two modifications in the SPACE-cohort and the ASAS-cohort.

In modification 1, the ASAS criteria for IBP were defined less stringent (≥3/5 instead of ≥4/5 IBP criteria). In both cohorts this resulted in a major increase in sensitivity, while specificity only slightly decreased. Modification 2 (IBP excluded as obligatory entry criterion) resulted in an even further decrease of missed axSpA diagnoses by the algorithm. Modification 2 showed the best combination of sensitivity and specificity in both cohorts.

Our findings show that IBP as obligatory entry criterion induces too many misdiagnoses, thereby confirming the results found before of a percentage of axSpA-patients without IBP up to 30%[4, 5]. Moreover, this is also the reason that the ASAS axSpA criteria are formed without IBP as entry criterion[4, 12]. However, IBP is suitable for screening for axSpA in primary care as several studies have shown[14–16], hence also a good (albeit non-mandatory) SpA-feature, as modification 2 suggests. Also for general practitioners it is important to realise that absence of IBP does not exclude axSpA. Furthermore, a relatively young age at onset of chronic back pain is a strong signal that the back pain might be a symptom of SpA. This is one of the factors explaining the difference of the 5% of SpA in the general population at the general practitioner level, and the 61% in this age-selected population seen by rheumatologists with a special interest in SpA. It should be noted that the algorithm is intended for use by the

rheumatologist, in this specific age-defined patient population, and not in an unselected population of patients with chronic back pain.
According to all versions of the algorithm, MRI-SI is not performed in HLA-B27 negative patients with 2-3 other SpA-features; those patients leave the algorithm as no axSpA patients. In order to further decrease these missed axSpA diagnoses, it could be considered to perform MRI-SI in HLA-B27 negative (especially male) patients with 2-3 other SpA-features [17]. There are suggestions that an MRI-SI should be classified as positive on the basis of inflammatory lesions and structural changes to increase sensitivity of the MRI-SI. Moreover, there are data showing that spinal changes on MRI-spine might be present in absence of inflammation on MRI of the SI-joints, yet this accounts for no more than 5% of patients [12]. The importance of these findings in the process of diagnosis is unclear at the moment.
It was not possible to decrease the number of patients who were erroneously diagnosed as axSpA by the proposed modifications. This might be caused by the fact that this mostly concerns patients with an (atypical) presentation of ≥3 SpA-features, but who are HLA-B27 negative and do not have sacroiliitis (X-SI and/or MRI-SI). Those patients are considered by rheumatologists as no axSpA, suggesting that rheumatologists base their diagnosis, besides the total presentation, to a large extent on MRI-SI and HLA-B27 findings. For the same reasons, those patients could never be classified according to the imaging arm, nor the clinical arm of the ASAS axSpA criteria. However, missed axSpA diagnoses in 3.8% to 6.1% and erroneously diagnosed axSpA patients in 7.6% to 10.2% of the cases (table 2), is surprisingly good. This also favours using the ASAS axSpA classification criteria in a diagnostic approach.
The use of both cohorts has strengths and limitations. A limitation of the use of the ASAS-cohort is that the ASAS axSpA criteria have been validated in this cohort while the ASAS axSpA criteria are used as one of the three external standards to test the performance of the algorithms. However, this is obviated since similar results are found in the SPACE-cohort, which is independent of the validation of the ASAS axSpA criteria. A downside of the SPACE-cohort is that the diagnosis of patients was based on the judgment of a single rheumatologist, what in turn is a strong point of the ASAS-cohort where the diagnosis was made by several ASAS-rheumatologists. For both cohorts the lack of follow-up data, which reduces the certainty on the diagnosis, is a limitation.
The results of both cohorts on the performance of the three diagnostic algorithms were presented to the ASAS-members during the January 2012 meeting in Amsterdam. The membership voted for modification 2 as the diagnostic algorithm of their choice.
In conclusion, ASAS accepted a modified algorithm in which IBP is excluded as obligatory entry criterion and is added as additional SpA-feature. We have added an online figure without the data on the cohorts that can be used in daily practice (online supplementary figure S1). This modification yields a higher agreement on the diagnoses in accordance with the diagnosis by the rheumatologist, the ASAS axSpA criteria and the LR-product probability ≥80%, mainly as a result of the reduction of missed axSpA diagnoses by the algorithm. This modified algorithm might be a useful tool for rheumatologists in daily practice.

Supplementary data

Additional data are published online only. To view these files please visit the journal online (http://dx.doi.org/ 10.1136/annrheumdis-2012- 201884)

References

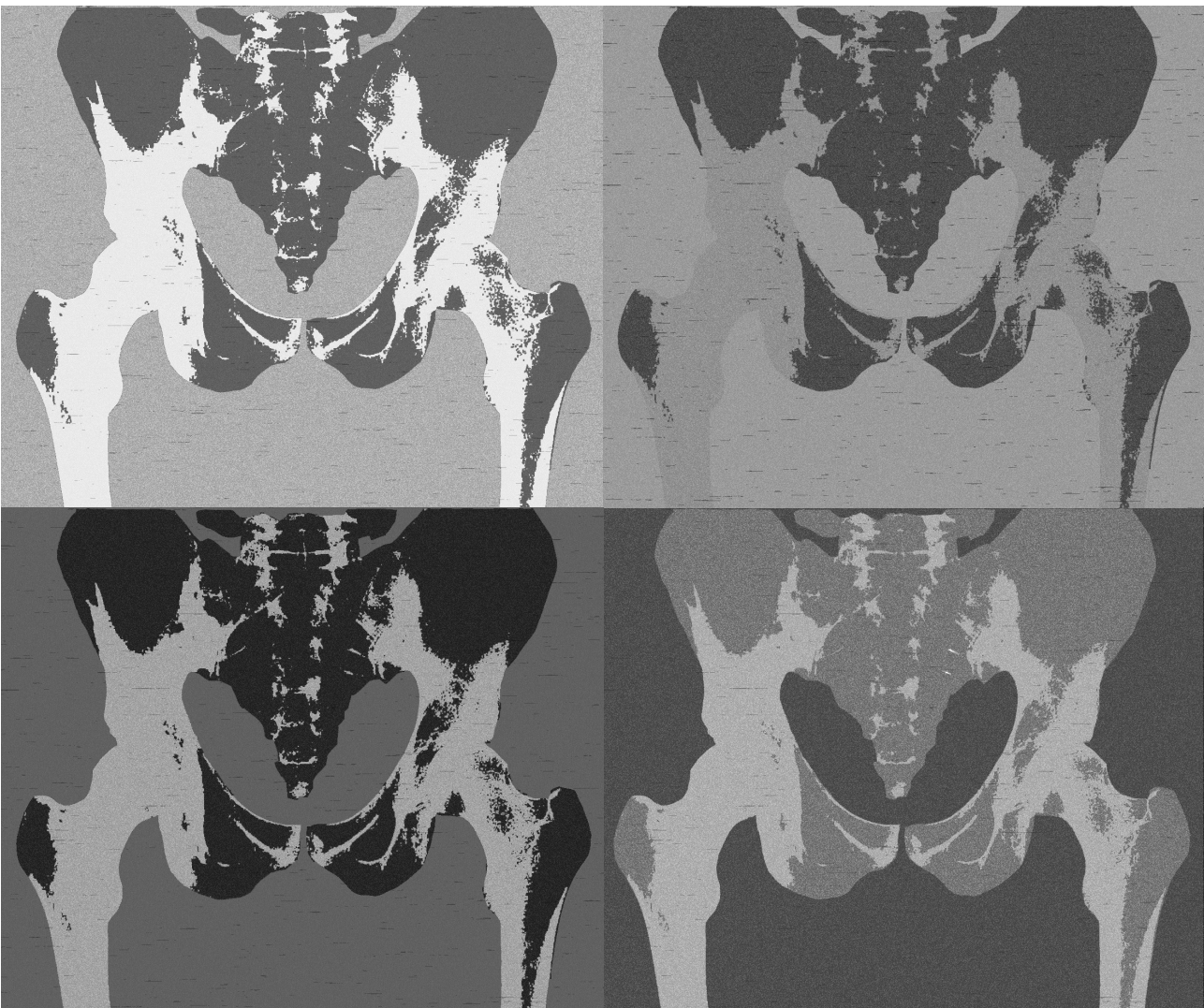
1. Khan M. Update on spondyloarthropathies. Ann Intern Med 2002;136:896-907.
2. Rudwaleit M, van der Heijde D, Khan M, *et al.* How to diagnose axial spondyloarthritis early. Ann Rheum Dis 2004;63:535-43.
3. Rojas-Vargas M, Munoz-Gomariz E, Escudero A, *et al.* First signs and symptoms of spondyloarthritis–data from an inception cohort with a disease course of two years or less (REGISPONSER-Early). Rheumatology (Oxford) 2009;48:404-9.
4. Rudwaleit M, Landewé 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.
5. Sieper J, van der Heijde D, Landewé R, *et al.* New criteria for inflammatory back pain in patients with chronic back pain: a real patient exercise by experts from the Assessment of SpondyloArthritis international Society (ASAS). Ann Rheum Dis 2009;68:784-8.
6. Rudwaleit M, Jurik A, Hermann K, *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.
7. van der Linden S, Valkenburg H, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361-8.
8. Braun J, Sieper J. Ankylosing spondylitis. Lancet 2007;369:1379-90.
9. Braun J, Inman R. Clinical significance of inflammatory back pain for diagnosis and screening of patients with axial spondyloarthritis. Ann Rheum Dis 2010;69:1264-8.
10. Calin A, Porta J, Fries J, *et al.* Clinical history as a screening test for ankylosing spondylitis. JAMA 1977;237:2613-4.
11. Rudwaleit M, Khan M, Sieper J. The challenge of diagnosis and classification in early ankylosing spondylitis: do we need new criteria? Arthritis Rheum 2005;52:1000-8.
12. Rudwaleit M, van der Heijde D, Landewé 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.
13. Rudwaleit M. New classification criteria for spondyloarthritis. Int J Adv Rheumatol 2010;8:1-7.
14. Poddubnyy D, Vahldiek J, Spiller I, *et al.* Evaluation of 2 screening strategies for early identification of patients with axial spondyloarthritis in primary care. J Rheumatol 2011;38:2452-60.
15. Braun A, Saracbasi E, Grifka J, *et al.* Identifying patients with axial spondyloarthritis in primary care: how useful are items indicative of inflammatory back pain? Ann Rheum Dis 2011;70:1782-7.
16. Brandt H, Spiller I, Song I, *et al.* Performance of referral recommendations in patients with chronic back pain and suspected axial spondyloarthritis. Ann Rheum Dis 2007;66:1479-84.
17. van Onna M, Jurik A, van der Heijde D, *et al.* HLA-B27 and gender independently determine the likelihood of a positive MRI of the sacroiliac joints in patients with early inflammatory back pain: a 2-year MRI follow-up study. Ann Rheum Dis 2011;70:1981-5.

4

Performance of classification criteria for peripheral spondyloarthritis and psoriatic arthritis in the Leiden Early Arthritis Cohort

Ann Rheum Dis 2012;71:1366–1369

R. van den Berg
F. van Gaalen
A. van der Helm-van Mil
T. Huizinga
D. van der Heijde



Abstract

Objectives

The performance of spondyloarthritis (SpA) classification criteria is not well-established in general early arthritis cohorts. Therefore, the authors tested their performance in the Leiden Early Arthritis Clinic (EAC)-cohort and assessed whether these criteria can assist rheumatologists in diagnosing patients.

Methods

The authors identified all SpA and psoriatic arthritis (PsA) patients in the EAC-cohort according to the diagnosis of the treating rheumatologist. A control group consisting of arthritis patients with other diagnoses was matched to the SpA and PsA patients on gender, age and symptom duration. The authors assessed the fulfilllment of SpA criteria in all three groups.

Results

Of the patients in the EAC-cohort (n=2011), 7.5% was diagnosed with PsA and 3.8% with SpA. In the PsA group, the ClASsification criteria for Psoratic Arthritis (CASPAR) criteria had the highest sensitivity (88.7%). In the SpA group, the Assessment of SpondyloArthritis international Society (ASAS) peripheral SpA and European Spondylarthropathy Study Group (ESSG) criteria had the highest sensitivity (both 48.7%). Specificity of all criteria sets was good: ranging from 88.5% (ESSG) to 100% (Amor).

Conclusions

In early arthritis, sensitivity of SpA classification criteria is modest except for the CASPAR criteria in PsA. However, specificity of classification criteria, including the new ASAS peripheral SpA criteria, is high.

Introduction

Early recognition of spondyloarthritis (SpA) is challenging since the concept of SpA comprises a heterogeneous group of diseases [1]. Over the years, several classification criteria have been developed. The European Spondylarthropathy Study Group (ESSG) criteria and the Amor criteria were developed to classify patients with all subtypes of SpA [2, 3]. Recently, two new sets were developed by the Assessment of SpondyloArthritis International Society (ASAS) to distinguish between patients with predominantly axial SpA (axSpA) and with predominantly peripheral SpA (pSpA) [4, 5]. Furthermore, the ClASsification criteria for Psoriatic ARthritis (CASPAR), is a classification set especially for psoriatic arthritis (PsA) [6].

The performance of all classification criteria was good when tested in the original validation population, frequently with longstanding symptoms [2, 3, 6–8]. The performance was less known in general early arthritis cohorts like the Leiden Early Arthritis Clinic (EAC)-cohort. It is known that 67% of the SpA-patients with a disease duration <2 years report arthritis as the first symptom [9]. So, in an EAC, SpA and PsA are important parts of the differential diagnosis. Therefore, it is important to test the performance of classification criteria in early disease cohorts.

First, we described the prevalence of SpA and PsA among patients presenting with peripheral arthritis in the Leiden EAC-cohort. Thereafter, we tested the performance of the described classification criteria and investigated whether these criteria sets can assist rheumatologists in diagnosing patients with peripheral arthritis.

Methods

Patients

Data from the Leiden EAC-cohort were used; a population-based prospective cohort including patients with recent-onset arthritis. Since 1993, general practitioners in the Leiden area referred patients with suspected arthritis as quickly as possible to the rheumatology department of the Leiden University Medical Center to detect and treat inflammatory disorders early. Patients with an objective evidence of arthritis, with a symptom duration <2 years and a signed informed consent, were included [10].

A database was built consisting of, among others: medical history, physical examination and laboratory tests according to the EAC protocol. Besides these parameters, individual patients' charts were reviewed for additional extra-articular SpA-features (past and/or present), necessary to apply the criteria sets (online supplementary table S1). Furthermore, in all patients, human leucocyte antigen (HLA)-B27 typing was performed if possible. All collected data (baseline to 1 year) was used for analysis. The diagnosis of the treating rheumatologist recorded from a list of proposed diagnoses including PsA and SpA after 1 year served as the gold standard.

Between 1 February 1993 and 1 February 2009, 2011 patients with early arthritis were included in the EAC-cohort. All PsA and SpA patients, according to the treating rheumatologist (n=226) at 1-year follow-up visit, were included in the present analysis. In the SpA group, 13 patients dropped out after 3 months; in the PsA group, 8 patients; and in the control group 36 patients. Of these patients, we used all available data which is until the third visit after 3 months.

Furthermore, a control group (n=226) was selected from the EAC-cohort, matched to the combined SpA-PsA group on gender, age and symptom duration (p=0.978, p=0.637 and p=0.03, respectively). Thereafter, the combined SpA-PsA group was split into the SpA group and PsA group. The control group included patients with the following diagnoses at 1 year: 82 with rheumatoid arthritis (1987 ACR criteria); 60 with undifferentiated arthritis; 13 with post-streptococcal reactive arthritis; 12 with osteoarthritis; 8 with gout; 15 with sarcoidosis;

and 26 patients with other diagnoses like palindromic arthritis and post-traumatic arthritis.

Data analysis

Baseline characteristics of the three groups (SpA, PsA and control group) were analysed using t-tests and χ^2-tests.
For each patient with SpA or PsA, we assessed the fulfilllment of the criteria sets and compared this with the fulfilllment of the criteria in patients in the control group using cross-table analysis. Missing values for the presence of SpA-features were interpreted as being absent. The checked diagnosis after 1 year was used as the gold standard.
Furthermore, the performance of the various criteria sets was determined by calculating sensitivity, specificity, positive likelihood ratio (LR+), and negative likelihood ratio (LR−). All analyses were performed using SPSS V. 17.0; p values <0.05 were considered significant.

Results

Patient characteristics during the first year of inclusion

Of the 2011 patients included in the EAC-cohort, 150 (7.5%) were diagnosed with PsA and 76 (3.8%) with SpA.
The control group was matched to the combined group of patients with SpA and PsA on gender, age and symptom duration. After splitting the patients into the SpA group and PsA group, differences with the control group were observed. The mean age of the SpA group was significantly lower (p<0.001), and of the PsA group significantly higher (p=0.03) compared with the control group. In the SpA group, 55 (72.3%) patients, in the PsA group, 64 (42.6%) patients, and in the control group, 121 (53.5%) patients were diagnosed before the age of 45 years.
Moreover, patients within the SpA group with a preceding infection had a significantly shorter self-reported symptom duration (6.4 (SD 9.4) versus 22.1 (SD 27.9) weeks; p<0.001), while patients without a preceding infection in the SpA group and PsA patients had a longer duration than patients in the control group (30.3 (SD 42.6) and 35.5 (SD 58.0) weeks; p=0.10 and p=0.01, respectively).
In the SpA group, all patients had at least one other SpA feature in addition to arthritis. In the control group, 66.4% of the patients had 1 or 2, and 22.1% had 3 or more SpA-features in addition to the arthritis, while in the PsA group and SpA group, respectively, 16.0% and 21.0% had 1 or 2 SpA-features, and 82.7% and 79.0% of the patients had 3 or more SpA-features (see online supplementary figure S1).
The most frequently reported SpA-features in the PsA group were psoriasis (94.0%), positive family history (89.2%) and dactylitis (36.7%). The PsA group differed significantly from the control group on these SpA-features, and on a higher prevalence of enthesitis, HLA-B27 positivity and rheumatoid factor (RF) negativity. By contrast, inflammatory back pain was significantly less frequent, and C-reactive protein levels significantly lower in the PsA group than in the control group. In the SpA group, HLA-B27 positivity (47.5%), positive family history (47.4%) and preceding infection (36.8%) were the most frequent SpA-features. SpA patients differed significantly from the control group on these features and on the presence of inflammatory back pain, enthesitis, uveitis, inflammatory bowel disease, negative RF and elevated erythrocyte sedimentation rate levels (table 1).

Table 1: Baseline characteristics.

	PsA (n=150)	SpA (n=76)	Control-group (n=226)	P-value PsA-group vs control-group	P-value SpA-group vs control-group
Age (years) at inclusion, mean ± SD	47.0 ± 13.8 n=150	37.0 ± 15.0 n=76	43.6 ± 15.6 n=226	0.03	<0.001
Male, n (%)	86 (57.3) n=150	37 (48.7) n=76	118 (52.2) n=226	0.33	0.59
Symptom duration* (weeks) at first visit, mean ± SD	35.5 ± 58.0 n=134	22.8 ± 37.3 (n=70) 6.4 ± 9.4 (n=22)** 30.3 ± 42.6 (n=48)***	22.1 ± 27.9 n=203	0.01	0.87
HLA-B27 positive, n (%)	19 (15.2) n=125	29 (47.5) n=61	14 (7.7) n=181	0.04	<0.001
Pos. fam. history SpA, n (%)	133 (88.7) n=150	36 (47.4) n=76	18 (8.0) n=226	<0.001	<0.001
IBP, n (%)	13 (8.7) n=150	22 (28.9) n=76	39 (17.3) n=226	0.02	0.03
Psoriasis, n (%)	141 (94.0) n=150	4 (5.3) n=76	13 (5.8) n=226	<0.001	0.87
Dactylitis, n (%)	55 (36.7) n=150	5 (6.6) n=49	5 (2.2) n=200	<0.001	0.07
Enthesitis, n (%)	17 (11.3) n=150	13 (17.1) n=76	11 (4.9) n=226	0.02	<0.001
Uveitis, n (%)	1 (0.7) n=150	7 (9.2) n=76	1 (0.4) n=226	0.83	<0.001
IBD, n (%)	1 (0.7) n=150	11 (14.5) n=76	4 (1.8) n=226	0.36	<0.001
Preceding infection, n (%)	10 (6.7) n=150	28 (36.8) n=76	11 (4.9) n=226	0.46	<0.001
RF negative, n (%)	133(91.7) n=145	67 (91.8) n=73	155 (72.4) n=214	<0.001	0.001
CRP (mg/l), mean ± SD	19.2 ± 24.6 n=133	40.7 ± 49.5 n=69	29.2 ± 43.9 n=211	0.01	0.07
ESR (mm/h), mean ± SD	29.2 ± 27.3 n=143	41.4 ± 32.5 n=75	31.6 ± 27.6 n=215	0.41	0.01
Good response to NSAIDs, n (%)	0 (0.0) n=150	7 (9.2) n=76	20 (8.8) n=226	<0.001	0.92
Asymmetric lower limb arthritis, n (%)	32 (21.3) n=150	32 (42.1) n=76	41 (18.1) n=226	0.44	<0.001
Sacroiliitis X-ray, x (%)	3 (9.4) n=32	9 (34.6) n=26	0 (0.0) n=35	0.24	<0.001
Juxta-articular new bone formations, n (%)	19 (12.7) n=150	0 (0.0) n=76	0 (0.0) n=226	<0.001	-

*Patient reported. **Patients with preceding infection. ***Patients without preceding infection
IBP, Inflammatory Back Pain; IBD, Inflammatory Bowel Disease; RF, Rheumatoid Factor; age, age at baseline; delay, duration between first complaints and first visit outpatient clinic Rheumatology; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; HLA-B27, Human Leukocyte Antigen; preceding infection can be balinitis, urethritis, cervicitis and/or acute diarrhea.

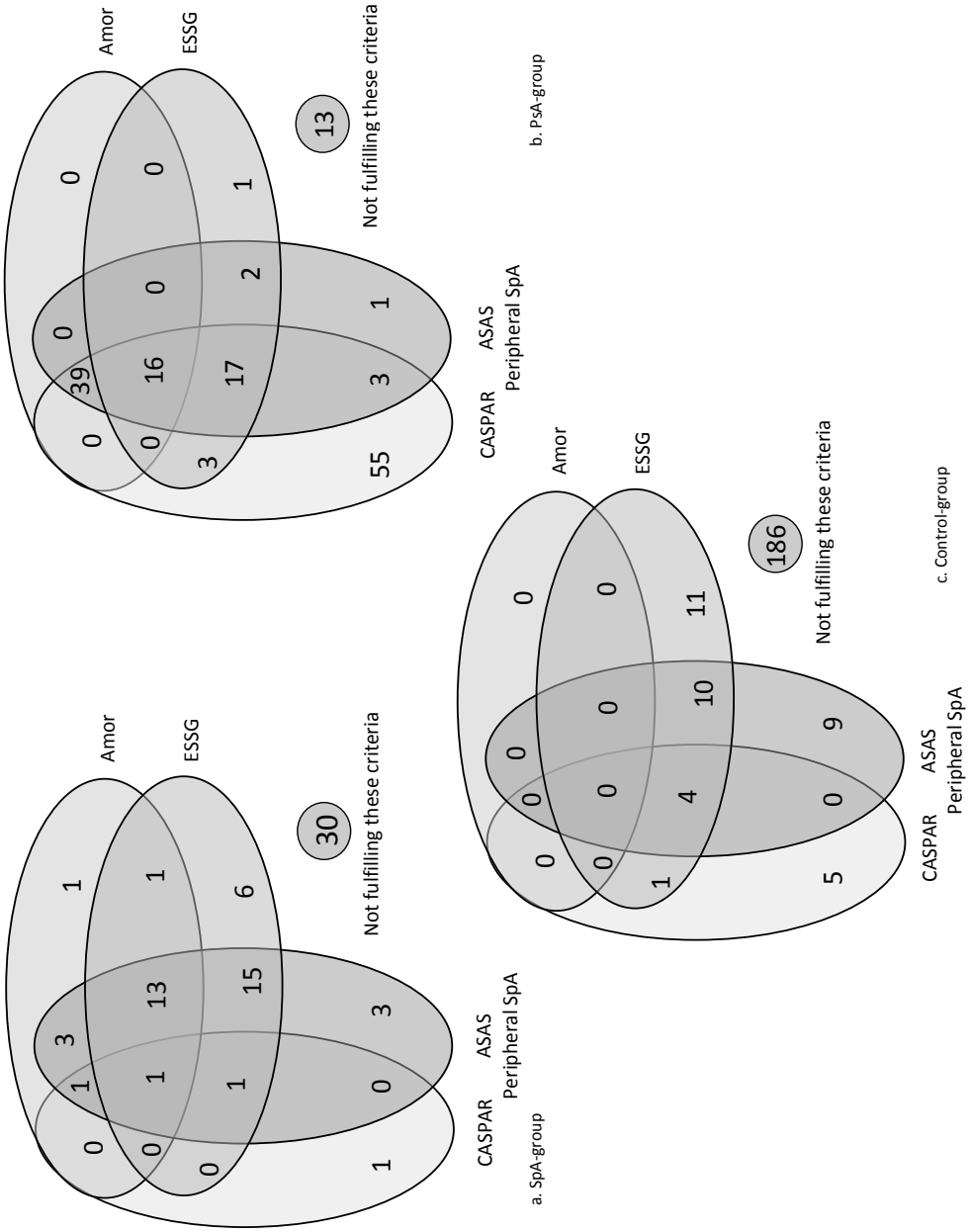


Figure 1: Venn-diagram representing the number of patients from the Leiden Early Arthritis Clinic meeting different criteria for spondyloarthritis. ESSG, European Spondylarthropathy Study Group; ASAS, Assessment of SpondyloArthritis International Society; CASPAR, ClASsification criteria for Psoratic ARthritis in the SpA-group (a), PsA-group (b), control-group (c).

Performance of classification criteria

In the PsA group, 133 patients fulfilled the CASPAR, and 78 the ASAS-pSpA criteria, (sensitivity of 88.7% and 52.0%, respectively). In the SpA group, 37 patients fulfilled the ASAS-pSpA and ESSG criteria each, (sensitivity both 48.7%). Specificity of all criteria sets was good, ranging from 88.5% (ESSG criteria) to 100% (Amor criteria). In the PsA group, LR+ and LR− of the CASPAR criteria were the best (20.04 and 0.12). In the SpA group, the best LR+ was identified for the ASAS-aSpA criteria, and the best LR− was identified for the ASAS-pSpA criteria (7.3 and 0.57, respectively) (table 2).

In the control group, 186 patients (82.3%) did not fulfill any of the four classification criteria, and 13 (8.7%) in the PsA group and 30 patients (39.5%) in the SpA group. The overlap of the criteria is presented in figure 1. In the PsA group, 16 patients fulfilled all four criteria sets and 39 fulfilled the combination of Amor, ASAS-pSpA and CASPAR criteria. Also in the SpA group many patients fulfilled at least two criteria sets (1 patient fulfilled all four criteria, 13 the combination of ASAS-pSpA, ESSG and Amor, and 15 the combination of ASAS-pSpA and ESSG criteria). In contrast in the control group, very few patients fulfilled more than one criteria set, and none of them all four criteria.

Furthermore, the concordance between the CASPAR and ASAS-pSpA criteria was calculated. In the PsA group, 75 patients fulfilled both the CASPAR and the ASAS-pSpA criteria, 58 patients fulfilled the CASPAR criteria and 3 fulfilled the ASAS-pSpA criteria only (59.3% agreement). Only 3 patients in the SpA group fulfilled both the CASPAR and the ASAS-pSpA criteria, 34 fulfilled the ASAS-pSpA criteria only, and one patient fulfilled the CASPAR criteria only (53.9% agreement).

Table 2: Number of patients fulfilling the various criteria sets and performance of the various criteria sets.

PsA vs control	**ASAS peripheral**	**ESSG**	**Amor**	**CASPAR**
PsA N pos. (sensitivity)	78 (52.0)*	39 (26.0)*	55 (36.7)*	133 (88.7)*
LR+	5.11	2.26	-	20.04
LR-	0.53	0.84	0.63	0.12
SpA vs control	**ASAS peripheral**	**ESSG**	**Amor**	**CASPAR**
SpA N pos. (sensitivity)	37 (48.7)*	37 (48.7)*	20 (26.3)*	4 (5.3)
LR+	4.78	4.2	-	1.19
LR-	0.57	0.58	0.74	0.99
Control N neg. (specificity)	203 (89.8)	200 (88.5)	226 (100)	216 (95.6)

* Marks a p-value <0.05, PsA-group or SpA-group is significantly different than the control-group on fulfilling the criteria ESSG, European Spondylarthropathy Study Group; ASAS, Assessment of SpondyloArthritis international Society (ASAS); CASPAR, Classification of Psoriatic Arthritis; LR+, positive likelihood ratio; LR-, negative likelihood ratio; PPV, positive predictive value; NPV, negative predictive value.

Discussion

More than 10% of the patients in the EAC-cohort were diagnosed with PsA or SpA. Unquestionably, this discovered prevalence is not representative for the whole concept of SpA among patients referred to rheumatologists since this cohort does not include patients with dactylitis or enthesitis, or patients with predominantly axSpA (back pain).

The sensitivities found in this report are lower for all criteria sets than the reported sensitivities as tested in the original cohorts with established disease patients [2–4, 6, 8].

A different patient selection in the EAC-cohort compared with the original cohorts can explain this difference. The EAC-cohort consists of patients with early arthritis, while the original cohorts included patients with various presenting features, not only arthritis, and a longer symptom duration. The longer the symptom duration, the more chance that (extra-articular) features develop.
Except for these issues related to symptom duration there are also factors related to more specific characteristics of the criteria sets. The sensitivity of the ESSG criteria may be limited by the lack of HLA-B27 in the list of SpA-features. The poor sensitivity of the Amor criteria in the EAC-cohort may also partially be explained by the strict definition of peripheral arthritis as oligoarthritis (2–4 joints). Besides, to fulfill the Amor criteria, at least six points are necessary corresponding to 3–4 items. Since some of these items are more seen in axSpA patients, this seems quite difficult to reach for patients with peripheral manifestations only [5]. Although it was expected that the sensitivity was lower than reported in the original cohorts it may also indicate that it is difficult to cover all pSpA-patients according to rheumatologists by existing criteria in recent onset disease, which confirms that these criteria are classifications and not diagnostic criteria sets [5].
On the other hand, it is very reassuring that the specificity of all sets were in accordance with the reported specificities, even in this cohort of early arthritis [2–4, 6, 8]. This is especially of note for the ASAS-pSpA criteria as these are quite new, and there was a fear that they might not be specific enough.
In conclusion, the various criteria sets are very good in classifying patients, but are limited in assisting rheumatologists in diagnosing patients.

Supplementary data

An additional supplementary material is published online only. To view this file please visit the journal online (http://ard.bmj.com/content/early/recent).

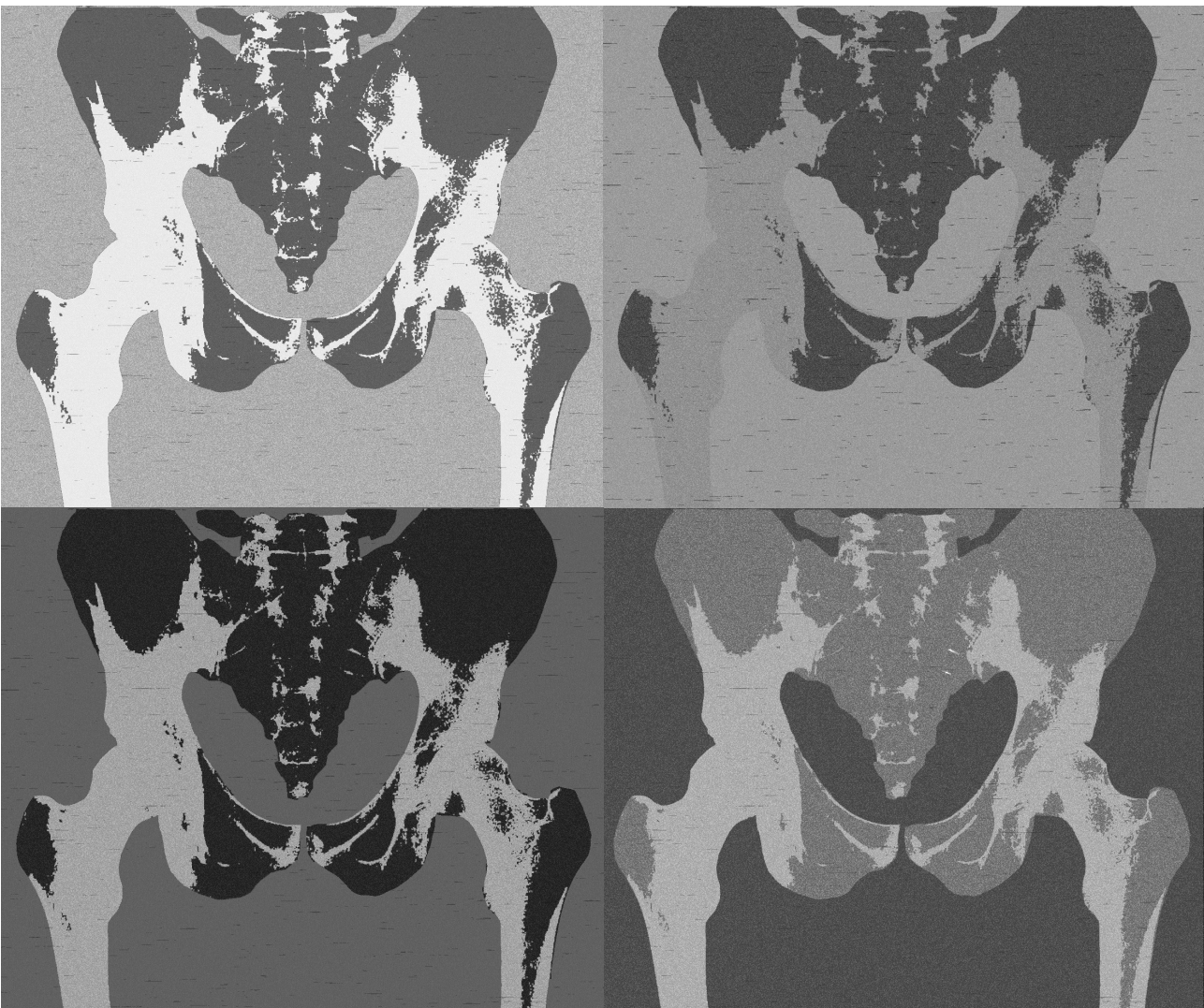
References

1. Khan M. Update on spondyloarthropathies. Ann Intern Med 2002;136:896-907.
2. Amor B, Dougados M, Mijiyawa M. [Criteria of the classification of spondylarthropathies]. Rev Rhum Mal Osteoartic 1990;57:85-9.
3. 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.
4. Rudwaleit M, van der Heijde D, Landewé 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.
5. Rudwaleit M. New approaches to diagnosis and classification of axial and peripheral spondyloarthritis. Curr Opin Rheumatol 2010;22:375-80.
6. Taylor W, Gladman D, Helliwell P, *et al*. Classification criteria for psoriatic arthritis: development of new criteria from a large international study. Arthritis Rheum 2006;54:2665-73.
7. Rudwaleit M, Landewé 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.
8. Rudwaleit M, van der Heijde D, Landewé R, *et al*. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. Ann Rheum Dis 2011;70:25-31.
9. Rojas-Vargas M, Muñoz-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 (Oxford) 2009;48:404-9.
10. de Rooy D, van der Linden M, Knevel R, *et al*. Predicting arthritis outcomes - what can be learned from the Leiden Early Arthritis Clinic? Rheumatology (Oxford) 2011;50:93-100.

Reading of sacroiliac joints on plain pelvic radiographs: agreement between clinical practice and trained central reading. Results of the DESIR-cohort

Arthritis Rheumatol Published Online First: [6 June 2014]
doi: 10.1002/art.38738

R. van den Berg
G. Lenczner
A. Feydy
D. van der Heijde
M. Reijnierse
A. Saraux
A. Rahmouni
M. Dougados
P. Claudepierre



Abstract

Objective

Investigating agreement on presence/absence of radiographic sacroiliitis between local rheumatologists/radiologists and central trained readers (external standard).

Method

Inflammatory back pain patients (IBP; ≥3 months, <3 years) suggestive of axial spondyloarthritis (axSpA) were included in the DESIR-cohort. Baseline sacroiliac-joint-radiographs were read by two central readers (modified New York), adjudicated by a third reader in case of disagreement, yielding a positive or a negative result (central reading). The same radiographs were also read by local radiologists/rheumatologists rated 'normal', 'doubtful sacroiliitis', 'obvious sacroiliitis' or 'SI-joint fusion' (local reading); positivity defined as: 1) at least unilateral 'obvious sacroiliitis' 2) 'bilateral 'obvious sacroiliitis' or at least unilateral 'fusion''. Agreement and misclassifications between central readers and central reading versus local reading were calculated (kappas).

Results

Interreader agreement between the central readers was moderate (κ=0.54); 108/688 radiographs (15.7%) were adjudicated. According to local reading (at least unilateral 'obvious sacroiliitis'), 183/688 patients (26.6%) had sacroiliitis; according to central reading, 145/688 patients (21.1%). Agreement between local reading and central reading was also moderate (κ=0.55); 76/183 patients (41.5%) with at least unilateral 'obvious sacroiliitis' (positive by local reading) and 32/109 patients (29.4%) with 'bilateral 'obvious sacroiliitis' or at least unilateral 'fusion'' (positive by local reading) were rated 'negative' by central reading; 38/505 patients (7.5%) and 68/579 (11.7%) respectively without sacroiliitis (negative by local reading) were read 'positive' by central reading.

Conclusion

In recent onset IBP-patients, both trained readers and local rheumatologists/radiologists agree only moderately in recognizing radiographic sacroiliitis. A significant proportion of locally recognized ankylosing spondylitis patients is not confirmed by central reading (false-positive), while a small minority is false-negative, indicating the necessity of re-evaluating the role of radiographic sacroiliitis as diagnostic criterion for axSpA.

Introduction

Sacroiliitis, detected on plain radiographs, is considered as the hallmark of ankylosing spondylitis (AS) and is mandatory for the classification of AS according to the modified New York criteria [1]. However, it is known that a major delay between symptom onset and the development of radiographic sacroiliitis exists [2]. Recently, the Assessment of SpondyloArthritis international Society (ASAS) published classification criteria to be able to classify patients with non-radiographic axial SpA (nr-axSpA) in addition to AS patients [3]. The classification of nr-axSpA, which was described for the first time already 29 years ago [4], is based on the absence of radiographic sacroiliitis but presence of SpA-features such as uveitis and dactylitis complementary to the presence of HLA-B27 and/or sacroiliitis visible on MRI. Although diagnostic criteria do not exist, radiographic sacroiliitis is also playing an important role in the diagnostic process of patients suspected of having axial SpA [5]. Axial SpA comprises the entire spectrum of patients with radiographic and non-radiographic disease, sacroiliitis on the radiograph being in fact the only discriminating feature. Though, the recognition of radiographic sacroiliitis is considered difficult because of the complex anatomy of the SI-joints, and the undulating articular surface makes the SI-joints hard to image on conventional radiographs, resulting in misinterpretations [6, 7]. A study including 100 rheumatologists and radiologists showed major discrepancies in grading of the SI-joints, especially in grades 1 and 2. Unfortunately, extensive training by workshops and self-education, did not enhance the performance of diagnosing sacroiliitis [7].

The distinction between AS and nr-axSpA based on the presence/absence of radiographic sacroiliitis becomes even more evident by the fact that in many countries TNF-inhibitors (TNFi) are currently approved for patients with established AS but not for nr-axSpA patients [8]. Moreover, the European Medical Agency has approved TNFi for nr-axSpA patients only if additional signs of objective inflammation such as a positive MRI and/or an elevated CRP are present, while this is not required for patients with radiographic axSpA. So there are major consequences depending on the judgement of a pelvis radiograph.

In daily practice, the diagnosis of AS is based on the judgement of the SI-joints on plain radiographs by the local radiologist and/or rheumatologist, frequently with knowledge of the clinical signs and symptoms. In cohorts and clinical trials on the other hand, the quantitative scoring of structural damage on radiographs of the SI-joints is usually done by one or more trained readers blinded for clinical information. In the DESIR (DEvenir des Spondylarthropathies Indifférenciées Récentes)-cohort, radiographs of the SI-joints at inclusion are scored by the local rheumatologist or radiologist and also by two trained central readers, including a third reader in case of discrepancy. Therefore, this cohort offers the unique opportunity to compare the evaluation of the local reading to the centralized reading, as external standard, in terms of agreement on abnormal versus normal SI-joints permitting to diagnose radiographic sacroiliitis.

Methods

Patients

For this analysis, baseline data from the DESIR-cohort were used. The DESIR-cohort is described extensively before 9. In short, consecutive patients aged 18-50 from 25 centers in France with inflammatory back pain (IBP) in the thoracic spine, lumbar spine and/or buttock area (≥3 months, but <3 years) based on either the Calin (4/5 items) or the Berlin (2/4 items) criteria [10, 11], suggestive of axSpA according to the rheumatologist with a score of ≥5 on a scale of 0 to 10 (where 0 was not suggestive of axSpA and 10 was very suggestive of axSpA), were included in this prospective longitudinal cohort to study the natural course and prognosis of axSpA starting at symptom onset. Between December 2007 and 29th of

April 2010, 708 patients were included.
The study fulfilled Good Clinical Practice Guidelines and was approved by the appropriate medical ethical committees. Participants gave written informed consent before they were included in the study. A detailed description of the study protocol is available at the website (http://www.lacohortedesir.fr/desir-in-english/). The research proposal for this particular analysis was approved by the scientific committee of the DESIR-cohort.

Data collection

A database was built by the use of standardized Case Record Form (CRF) on which the following, among others, needed to be filled out: physical examination, on-going treatment, co-morbidities, laboratory tests and questionnaires, according to the DESIR protocol [9]. The database used for this analysis was locked on October 30th 2012.

Images and scoring methods

Two central readers (RvdB and GL), both familiar with scoring SI-joints on plain radiographs (X-SI) in the anteroposterior view according to the modified New York (mNY) method, participated in a calibration session. A grade 0 is given for a normal SI-joint; 1 for suspicious changes; 2 for minimal abnormality - small localised areas with erosions or sclerosis without alteration in joint width; 3 for unequivocal abnormality - with one or more erosions, evidence of sclerosis, joint space narrowing or widening or partial ankylosis, and grade 4 for severe abnormality - a complete ankylosis of the SI-joint. Sacroiliitis is defined as grade ≥2 bilaterally or grade 3-4 unilaterally [1]. The calibration session was a systematic conducted exercise, executed by two senior radiologists (MR and AF) and two senior rheumatologists (PC and MD), who already did such calibration sessions before. The whole process was supervised by an expert in AS and imaging scoring (DvdH). During the first step of the calibration process, definitions of lesions, examples and pitfalls were discussed. The second step of the calibration session consisted of independently reading of training cases by the two readers, under the supervision of those radiologists and rheumatologists. The results of the readings were discussed plenary by the senior radiologists and rheumatologists, focussing on disagreement regarding specific lesions between the two readers in order to achieve agreement. In the third step of the calibration process, 30 X-SIs were read independently by the two central readers. The fourth step consisted of a consensus meeting in which the same four senior radiologists/rheumatologists participated as well. Again, during a plenary presentation the disagreements between the two readers were discussed by one of the senior radiologists/rheumatologist in order to achieve agreement. Next, in a fifth step a second set consisting of 20 X-SIs were read independently, again followed by the last step consisting of a consensus meeting with the same senior radiologists/rheumatologists executed in the same manner. At that time, interreader agreement largely improved even though we recognized that the kappas were still moderate (kappa=0.55). However, considering the results of the study by van Tubergen *et al*. showing that training did not improve performance [7], we didn't expect significant further improvements in agreement. Moreover the kappas were in the same range as what have been found in other studies as well (kappas ranging between 0.12 and 0.69) [6, 12], thereby justifying the decision that the two readers could start reading the DESIR-cohort.
Baseline X-SIs were acquired according to a standardized method, provided in the DESIR protocol. All available digital baseline X-SIs of the DESIR-cohort (n=688) were read independently by the two trained central readers according to the mNY criteria, blinded for all clinical and laboratory data, as well as for the results of the local reads for the X-SIs. Agreement on fulfilllment of the mNY criteria for radiographic sacroiliitis at the patient level between the two readers was calculated and in case the readers disagreed, a radiologist

experienced in the field of SpA (MR) served as adjudicator. A radiograph of the SI-joints was marked positive for sacroiliitis if 2/3 readers agreed on the fulfilllment of mNY criteria (ie: at least unilateral grade 3 sacroiliitis, or at least bilateral grade 2 sacroiliitis), hereafter called 'central reading'. Furthermore, the two central readers marked the type of lesions they recognized (erosions, sclerosis, joint space widening/narrowing and (partial) ankylosis).
Local radiologists or rheumatologists who might have access to all clinical information and lab test results at each study center, read all available baseline radiograph of the SI-joint in their own center, hereafter called 'local reading'. Since the local readers, who are working in regular clinical practice, were not trained experts it was considered more appropriate to use a scoring system that better resembles common clinical practice than the mNY criteria do: local readers were asked to rate each SI-joint either as 'normal' or as 'doubtful sacroiliitis' or as 'obvious sacroiliitis' or as 'SI-joint fusion' [13]. No specification of the type of lesions was provided by the local readers. In this scoring method, at least a unilateral rating of 'obvious sacroiliitis' was considered sufficient to fulfill the imaging criterion of sacroiliitis. This was our primary analysis with regard to the comparison with the fulfilllment of the mNY by central reading. In a second analysis, we have further compared a rating of 'bilateral 'obvious sacroiliitis' or unilateral 'fusion'' with the fulfilllment of the mNY criteria by central reading. Finally, we have also used a stricter definition of the mNY criteria of the central reading: at least bilateral grade 3 or unilateral grade 4.
To compare the grading of the *individual* SI-joints of the central readers to the scoring of the *individual* SI-joints by the local readers, central mNY grades 2 and 3 were combined and compared to local 'obvious sacroiliitis' and mNY grade 4 was compared to local 'SI-joint fusion' for each central reader separately. Moreover, central mNY grades 3 and 4 were combined and compared to the local read 'fusion' again for each central reader separately.

Statistical analysis

Agreement was calculated using cross-tabulation expressed in Cohen's Kappa (κ) or linear weighted kappa (κ_w) as appropriate, agreement on the positive cases (positive agreement) and agreement on the negative cases (negative agreement) for the following comparisons [14,15]: interreader agreement on the presence/absence of radiographic sacroiliitis between the two central readers, between local reading and central reading using the various definitions for sacroiliitis explained above, and interreader agreement on the type of lesion (erosions, sclerosis, joint space narrowing/widening and ankylosis) between the two central readers and on the grading of the SI-joints between the two central readers and between local reading and central reading as explained above. Sensitivity, specificity, positive predictive values (PPV), negative predictive values (NPV) as well as the number of false-positive and false-negative classifications by local reading versus central reading as external standard were calculated.
Among the patients with a positive X-SI according to local reading, it was investigated which type of lesion was most frequently scored by the central readers separately.
All kappas were interpreted according to the standards proposed by Landis and Koch; values <0 as indicating no agreement and 0-0.20 as slight, 0.21-0.40 as fair, 0.41-0.60 as moderate, 0.61-0.80 as substantial, and 0.81-1 as almost perfect agreement [16]. SPSS software version 20.0 was used for the statistical analysis.

Results

The mean age of the 688 included patients was 31.6 (SD 8.6) years, the mean symptom duration was 17.7 (SD 10.5) months, 368 (53.5%) patients were men and 405 (58.9%) patients were HLA-B27 positive. In 648/688 patients imaging data was complete; in 40 additional patients data on MRI-SI was missing. In 582/648 patients with available imaging

data the ASAS axSpA criteria could be applied. The remaining 66 patients had an onset of IBP >45 years of age. Based on local scoring of imaging for sacroiliitis, 408/582 patients (70.1%) fulfilled the ASAS axSpA criteria (84.1% HLA-B27 positive). One hundred sixty-one fulfilled the ASAS axSpA criteria based on the presence of radiographic sacroiliitis (at least unilateral 'obvious sacroiliitis'; 75.8% HLA-B27 positive); 83/408 based on inflammatory sacroiliitis on MRI as judged by the local reading (67.5% HLA-B27 positive) and the remaining 164/408 patients fulfilled the ASAS axSpA criteria based on HLA-B27 positivity.

Table 1: Interreader agreement between central reader 1 and central reader 2 in SI-joints radiographs reading (n=688 patients).

		Central reader 2	
		modified New York +	**modified New York –**
Central reader 1	**modified New York +**	96	58
	modified New York –	50	484
	κ (95% CI): 0.54 (0.46-0.62)	Positive agreement: 64.0%	Negative agreement: 90.0%
		Erosions +	**Erosions –**
	Erosions +	132	131
	Erosions –	133	980
	κ (95% CI): 0.38 (0.32-0.44)	Positive agreement: 50.0%	Negative agreement: 88.1%
		Sclerosis +	**Sclerosis –**
	Sclerosis +	182	144
	Sclerosis –	133	917
	κ (95% CI): 0.44 (0.38-0.49)	Positive agreement: 56.8%	Negative agreement: 86.9%
		Joint space widening +	**Joint space widening –**
	Joint space widening +	13	137
	Joint space widening –	10	1216
	κ (95% CI): 0.13 (0.06-0.19)	Positive agreement: 15.0%	Negative agreement: 94.3%
		Joint space narrowing +	**Joint space narrowing –**
	Joint space narrowing +	14	127
	Joint space narrowing –	21	1214
	κ (95% CI): 0.12 (0.05-0.20)	Positive agreement: 15.9%	Negative agreement: 94.3%
		Ankylosis +	**Ankylosis –**
	Ankylosis +	27	63
	Ankylosis –	67	1219
	κ (95% CI): 0.24 (0.15-0.33)	Positive agreement: 29.3%	Negative agreement: 94.9%

Not possible to evaluate: n=27 for right SI-joint and n=29 for left SI-joint. Positive agreement is the agreement on positive cases. Negative agreement is the agreement on negative cases.

Agreement between the two central readers

Agreement between the two central readers regarding absence/presence of radiographic sacroiliitis (mNY) is moderate (κ=0.54; table 1). The adjudicator needed to read 108/688 X-rays (15.7%) because of disagreement between the 2 central readers.

Agreement regarding the grading of the SI-joints is also moderate (κ_W=0.56 for both left and right SI-joints; table 2). Most disagreement is seen in grade 0 versus grade 1, followed by grade 0 versus grade 2, and followed by grade 1 versus grade 2. Similar numbers of disagreement are seen for grade 2 versus grade 3 as for grade 1 versus grade 2. Depending on the grade of the other SI-joint, this could cause a different classification of a patient. The remaining numbers of disagreement are seen in grade 0 versus grade 3 and, and in grade 1 versus grade 3. With these types of disagreement, a patient is classified differently regardless of the grade of the other SI-joint of the patient.

Table 2: Interreader agreement of the grading of the SI-joints (mNY) between central reader 1 and central reader 2.

		Central reader 2				
Central reader 1	**Right SI-joint**	**Grade 0**	**Grade 1**	**Grade 2**	**Grade 3**	**Grade 4**
	Grade 0	345	51	35	4	0
	Grade 1	28	14	17	5	0
	Grade 2	26	16	40	15	0
	Grade 3	7	6	19	45	0
	Grade 4	0	0	0	5	2
	κ_W (95% CI): 0.56 (0.50-0.61)					
		Central reader 2				
Central reader 1	**Left SI-joint**	**Grade 0**	**Grade 1**	**Grade 2**	**Grade 3**	**Grade 4**
	Grade 0	369	48	22	6	0
	Grade 1	15	13	14	10	0
	Grade 2	24	26	33	20	0
	Grade 3	8	7	23	34	0
	Grade 4	0	0	0	1	5
	κ_W (95% CI): 0.56 (0.51-0.62)					

Not possible to evaluate: n=27 for right SI-joint and n=29 for left SI-joint.

Agreement between local reading and central reading

According to local reading (at least unilateral 'obvious sacroiliitis') 183/688 patients (26.6%) had radiographic sacroiliitis and according to central reading 145/688 patients (21.1%) had radiographic sacroiliitis. Agreement between local reading and central reading was very similar to the interreader agreement between the two central readers (κ=0.55). Comparing local reading to the scores of the individual central readers revealed very similar levels of agreement (table 3).

In the comparison of local reading to central reading, 76 out of the 183 patients (41.5%) with a positive radiograph of the SI-joints according to local reading were read negative by central reading; 38 patients out of 505 (7.5%) negative radiographs of the SI-joints according to local reading were read positive by central reading (table 3). In daily practice, where local readers judge the X-SIs, this would mean that 41.5% of the AS patients are overclassified compared to central reading as external standard (45.9% compared to reader 1 and 42.1% compared to reader 2) and that AS (according to central reading) is not recognized in 7.5% of the patients compared to central reading as external standard (10.9% compared to reader 1 and 7.9% compared to reader 2).

Seventy-six of the 183 patients with 'obvious sacroiliitis' according to local reading had unilateral 'obvious sacroiliitis'; thirty of these were rated as 'negative' by central reading. If only patients with at least bilateral 'obvious sacroiliitis' or at least unilateral 'fusion' were considered 'positive' by local reading, still 32/109 patients (29.4%) were rated as 'negative' by central reading (according to the original mNY; κ=0.52). The proportion of negative ratings according to this definition of local reading that were rated 'positive' by central reading increased to 11.7% (68/579).

Thereafter we compared this more stringent definition of positivity by local reading with a more stringent definition of positivity by central reading (at least grade 3 bilateral or grade 4 unilateral involvement). This resulted in reduced agreement (κ=0.44 for reader 1 and κ=0.43 for reader 2), showing even a higher number of patients being overclassified as AS according to local reading compared to the central readers (62.4% and 65.1% reader 1 and 2 respectively). The interreader agreement between the two central readers for this stricter mNY definition of sacroiliitis was again only moderate (κ=0.56), but better than the agreement between local reading and the individual central readers. Even with the most stringent definition of sacroiliitis ('positivity' defined as bilateral fusion in the local reading and bilateral grade 4 in the central reading) still huge disagreement between local reading and the individual central readers was seen (table 3).

Table 4 shows the comparison of the individual ratings of local reading and the mNY gradings per SI-joint of central reading on a individual reader basis. For the determination of agreement, we have both combined mNY grades 2 and 3 (for 'positivity') as well as mNY grades 3 and 4, and compared these with 'obvious sacroiliitis' from the local read. The kappas were only moderate (range 0.36-0.51) with regard to both readers and both definitions.

Table 3: Agreement between local reading and central reading of presence/absence of sacroiliitis, for the various definitions of sacroiliitis.

		Central reader 1	
Local reading (at least unilateral 'obvious sacroiliitis')		**Sacroiliitis +**	**Sacroiliitis -**
	Sacroiliitis +	99	84
	Sacroiliitis -	55	450
	κ (95% CI): 0.46 (0.38-0.53)	Pos. agreem.: 58.8%	Neg. agreem.: 86.6%
		Central reader 2	
		Sacroiliitis +	**Sacroiliitis -**
	Sacroiliitis +	106	77
	Sacroiliitis -	40	465
	κ (95% CI): 0.53 (0.46-0.61)	Pos. agreem.: 64.4%	Neg. agreem.: 88.8%
		Central reading (2/3 readers)	
		Sacroiliitis +	**Sacroiliitis -**
	Sacroiliitis +	107	76
	Sacroiliitis -	38	467
	κ (95% CI): 0.55 (0.47-0.62)	Pos. agreem.: 65.2%	Neg. agreem.: 89.1%
Sensitivity: 73.8%	Specificity: 86.0%	NPV: 0.92	PPV: 0.58

Table 3: Continued

		Central reader 1	
Local reading (at least bilateral 'obvious sacroiliitis' or unilateral 'fusion')		**Sacroiliitis +**	**Sacroiliitis -**
	Sacroiliitis +	74	35
	Sacroiliitis -	80	499
	κ (95% CI): 0.46 (0.38-0.55)	Pos. agreem.: 56.3%	Neg. agreem.: 89.7%
		Central reader 2	
		Sacroiliitis +	**Sacroiliitis -**
	Sacroiliitis +	76	33
	Sacroiliitis -	70	509
	κ (95% CI): 0.51 (0.42-0.59)	Pos. agreem.: 59.6%	Neg. agreem.: 90.8%
		Central reading (2/3 readers)	
		Sacroiliitis +	**Sacroiliitis -**
	Sacroiliitis +	77	32
	Sacroiliitis -	68	511
	κ (95% CI): 0.52 (0.44-0.60)	Pos. agreem.: 60.6%	Neg. agreem.: 91.1%
		Reader 1, strict definition (≥3 bilaterally, or 4 unilaterally)	
Local reading (at least bilateral 'obvious sacroiliitis' or unilateral 'fusion')		**Sacroiliitis +**	**Sacroiliitis -**
	Sacroiliitis +	41	68
	Sacroiliitis -	15	564
	κ (95% CI): 0.44 (0.34-0.53)	Pos. agreem.: 49.7%	Neg. agreem.: 93.1%
		Reader 2, strict definition (≥3 bilaterally, or 4 unilaterally)	
		Sacroiliitis +	**Sacroiliitis -**
	Sacroiliitis +	38	71
	Sacroiliitis -	10	569
	κ (95% CI): 0.43 (0.33-0.53)	Pos. agreem.: 48.4%	Neg. agreem.: 93.4%
		Reader 1, strict definition (≥4 bilaterally)	
Local reading (bilateral 'fusion')		**Sacroiliitis +**	**Sacroiliitis -**
	Sacroiliitis +	2	3
	Sacroiliitis -	3	680
	κ (95% CI): 0.40 (0.01-0.78)	Pos. agreem.: 40.0%	Neg. agreem.: 99.6%
		Reader 2, strict definition (≥4 bilaterally)	
		Sacroiliitis +	**Sacroiliitis -**
	Sacroiliitis +	0	5
	Sacroiliitis -	1	682
	κ (95% CI): 0.00 (0.00-0.00)	Pos. agreem.: 0.00%	Neg. agreem.: 99.6%

Pos. agreem. is the agreement on positive cases. Neg. agreem. is the agreement on negative cases.

Sensitivity, specificity, NPV (negative predictive value) and PPV (positive predictive value) applies to the comparison of the local reading (at least unilateral 'obvious sacroiliitis') to the central reading (≥2 bilaterally, or ≥3 unilaterally, 2/3 readers).

Table 4: Agreement of the grading of the SI-joints between local reading and the individual central readers

		Central reader 1				
Local read	**Right SI-joint**	**Grade 0**	**Grade 1**	**Grade 2***	**Grade 3*^**	**Grade 4^**
	Normal	321	36	37	15	0
	Doubtful sacroiliitis	80	16	28	10	1
	Obvious sacroiliitis	34	10	31	49	3
	Fusion	0	2	1	3	3
	* Grade 2 and 3 of central reader 1 combined: κ_w (95% CI): 0.40 (0.34-0.46) ^ Grade 3 and 4 of central reader 1 combined: κ_w (95% CI): 0.36 (0.30-0.41)					
		Central reader 1				
Local read	**Left SI-joint**	**Grade 0**	**Grade 1**	**Grade 2***	**Grade 3*^**	**Grade 4^**
	Normal	336	21	36	15	0
	Doubtful sacroiliitis	77	12	24	9	1
	Obvious sacroiliitis	31	18	42	44	1
	Fusion	0	1	2	4	4
	* Grade 2 and 3 of central reader 1 combined: κ_w (95% CI): 0.45 (0.38-0.51) ^ Grade 3 and 4 of central reader 1 combined: κ_w (95% CI): 0.41 (0.35-0.46)					
		Central reader 2				
Local read	**Right SI-joint**	**Grade 0**	**Grade 1**	**Grade 2***	**Grade 3*^**	**Grade 4^**
	Normal	324	46	32	10	0
	Doubtful sacroiliitis	64	27	33	10	1
	Obvious sacroiliitis	21	14	45	50	0
	Fusion	1	1	1	5	1
	* Grade 2 and 3 of central reader 2 combined: κ_w (95% CI): 0.48 (0.42-0.54) ^ Grade 3 and 4 of central reader 2 combined: κ_w (95% CI): 0.47 (0.42-0.53)					
		Central reader 2				
Local read	**Left SI-joint**	**Grade 0**	**Grade 1**	**Grade 2***	**Grade 3*^**	**Grade 4^**
	Normal	330	50	23	7	0
	Doubtful sacroiliitis	60	29	25	12	1
	Obvious sacroiliitis	28	17	45	48	1
	Fusion	0	0	0	8	3
	* Grade 2 and 3 of central reader 2 combined: κ_w (95% CI): 0.51 (0.45-0.57) ^ Grade 3 and 4 of central reader 2 combined: κ_w (95% CI): 0.44 (0.39-0.49)					

Not possible to evaluate: n=27 for right SI-joint and n=29 for left SI-joint. Positive agreement is the agreement on positive cases. Negative agreement is the agreement on negative cases.

Types of lesions

Regarding the type of lesion, agreement between the two central readers varied from κ=0.12 for joint space narrowing to κ=0.44 for sclerosis (table 5). The prevalence of joint space alterations and ankylosis is low in this cohort of recent onset IBP patients; among the SI-joints graded as at least grade 3 the prevalence of joint space widening was 8.7% (reader 2) and 24.6% (reader 1) and the prevalence of ankylosis was 13.7% (reader 1) and 27.3% (reader 2). The frequency of erosions and sclerosis in SI-joints graded as 2 was very similar to the frequency of erosions and sclerosis in SI-joints graded as 3, for both readers.
Among the patients with a positive radiograph of the SI-joints according to local reading (at least unilateral 'obvious sacroiliitis'), sclerosis was the most frequently reported lesion by the central readers (56.8% according to reader 1 and 72.7% according to reader 2), followed by erosions (50.3% according to reader 1 and 61.2% according to reader 2) (table 5).

Table 5: Frequency of lesions read by the central reader 1 and central reader 2 among patients with a positive radiograph of the SI-joints according to local reading (at least unilateral 'obvious sacroiliitis') (n=183).

Type of lesion (%)	Reader 1	Reader 2
Erosions	50.3	61.2
Sclerosis	56.8	72.7
Joint space widening	24.6	8.7
Joint space narrowing	24.0	11.5
Ankylosis	13.7	27.3

Discussion

In the DESIR-cohort, the interreader agreement between two trained central readers and between local reading and central reading of X-SIs is moderate at best. The two central trained readers showed only moderate agreement with regard to presence/absence of radiographic sacroiliitis, grading of the SI-joints and about type of lesion, yet comparable to levels of agreement reported in previous studies where central readers were also trained (κ=0.12-0.69 for absence/presence of sacroiliitis; κ=0.22 for erosions; κ=0.26 for sclerosis and κ=0.19 for joint space alterations) [6, 12, 17]. Furthermore, interreader agreement was at a similar level as the agreement between local reading and central reading about presence/absence of radiographic sacroiliitis and grading of the SI-joints.
As it was not specified which X-SIs were judged by local radiologists and which by local rheumatologists, it was not possible to compare the readings of radiologists and rheumatologists separately to the central reading. However, we did not expect a difference in number of misclassified patients between radiologists and rheumatologists based on the findings of van Tubergen *et al.*, and based on our own findings regarding presence/absence of sacroiliitis on MRI in the DESIR-cohort [7, 18]. Moreover, the results of an unpublished ASAS survey pointed out that more than 55% of the rheumatologists rely on both the judgement of the radiologist as well as their own judgement in assessing sacroiliitis on radiographs (M. Rudwaleit, personal communication, January 18 2014, unpublished data from an ASAS survey).
Misclassifications could have major implications for a patient, as the presence/absence of sacroiliitis is the only difference in the classification of AS (mNY) versus nr-axSpA or even no SpA. This is indeed what the results show; 41.5-45.9% of the patients classified as AS by the local readers are falsely classified with the central read as external standard and 7.5-10.9% of the AS patients according to the central read is not recognized in daily practice.

The percentages of misclassified patients are somewhat higher than reported in another study where 11.4% of AS patients according to local readers were reclassified as nr-axSpA by central trained readers, and 15.5% of nr-axSpA patients according to local readers were reclassified as AS by central readers [19]. However, these lower percentages can probably be explained by the fact that this study also included patients with longstanding AS (>10 years) showing more severe lesions which are easier to recognize, the fact that the presence/absence of syndesmophytes in the spine was taken into account as well, and by the fact that both readers had to agree on the absence/presence of sacroiliitis thereby not looking at a possible reclassification of patients in whom the two central readers disagreed regarding the presence/absence of sacroiliitis [19].

As our aim was to compare the diagnostic performance of readers in daily clinical practice to the performance of trained expert readers using the mNY grading system for the classification of patients in studies, we have applied a "daily practice definition" of sacroiliitis: in daily clinical practice, a rheumatologist will consider a diagnosis of axial SpA when there is 'obvious sacroiliitis' at least unilaterally, but obviously the rheumatologist is more convinced of a diagnosis of axial SpA in case of bilateral involvement. Here we have examined both definitions of 'positivity': one more lenient and one more stringent definition. If 'at least unilateral 'obvious sacroiliitis'' in the local read was required for positivity, 41.5% of the patients with radiographic sacroiliitis could not be confirmed by central reading (using the original mNY definition). However, even if 'bilateral 'obvious sacroiliitis' or at least unilateral 'fusion'' by the local reading was required for positivity, still 29.4% of the patients with a local diagnosis of AS could not be confirmed by central reading. This percentage of misclassification increased to 62.4% (reader 1) and 65.1% (reader 2) if we compared the local rating of either 'bilateral 'obvious sacroiliitis' or unilateral 'fusion'' with a similarly strict definition of sacroiliitis by central reading (at least grade 3 bilaterally or unilateral grade 4).

As the local and central reading are not identical, the true percentage of misclassified patients must be between 29.4% and 65.1%, and likely around 40-45% as the use of 'at least unilateral 'obvious sacroiliitis'' by the local reader is what is required in clinical practice to classify a patient as having AS. While this study has been performed in the DESIR centers in France, and generalizability is formally restricted, there is no valid reason to assume that clinical rheumatologists in other countries apply different diagnostic reasoning.

As the local readers did not mark a specification of the type of lesions, the reads of the central readers were used to gain insight in which type of lesion was best recognized by the local readers. However, because of the low prevalence of joint space alterations and ankylosis it is difficult to investigate the agreement on recognizing this type of lesion. Yet, if joint space alterations are present, the two central readers recognized ankylosis more easily than widening or narrowing of the joint space. The prevalence of sclerosis and erosions, on the other hand, is higher and both types of lesions are more easily recognized than joint space alterations as shown by the higher kappas.

Although the question whether training improves recognition of radiographic sacroiliitis was not addressed directly in this study, the fact that the agreement between local reading and central reading is so similar to the interreader agreement between the two trained central readers seems to confirm the findings of van Tubergen *et al.* [7] that training does not improve recognition of radiographic sacroiliitis. This arises the question whether it is necessary for DESIR and similar cohorts to have a central reading of the radiographs of the SI-joints instead of a local reading. It could be argued based on the discovered levels of agreement that there is no preference for central readers over local readers since both trained central readers and local readers can only poorly recognize radiographic sacroiliitis. Nonetheless, central reading consisting of a judgement of 2 out 3 agreeing readers suggests being more robust than a local reading based on the opinion of a single reader. Depending on the research question, a choice could be made which reads to use: either the reads that would have been

used in clinical practice as well, or the reads of the central reading based on a majority read of 2 out of 3, which are closer to the truth.
For both local and trained central readers, the recognition of radiographic sacroiliitis remains challenging. Nevertheless, the only difference between AS and nr-axSpA is the presence of radiographic sacroiliitis. A patient can be classified as AS if only IBP is present in addition to radiographic sacroiliitis while in the absence of radiographic sacroiliiitis and in the absence of a positive MRI (ASAS definition) a minimum of 3 other SpA-features must be present in order to classify the patient as axSpA [3]. The fact that a patient is classified differently, based on a different read of the same radiograph of the SI-joints - which is shown to happen frequently - arises the question how 'gold' this distinction between AS and nr-axSpA is. It is worrisome that such a small factor can have major consequences for a patient, not only in terms of diagnosis but also in terms of treatment as based on the presence/absence of radiographic sacroiliitis TNFi can be administered or not. Moreover, inclusion of patients in clinical trials is based on the presence/absence of radiographic sacroiliitis as judged by either local or central readers. Rereading the radiographs by different readers or even blinded rereading of the same radiographs by the same readers could lead to significant change in classification of the patients [6].
A limitation of this study is the lack of a gold standard as CT, to confirm the presence/absence of sacroiliitis [20]. Furthermore, this study focussed on sacroiliitis on conventional radiographs only. The role of MRI in the diagnosis and classification of axSpA should be investigated in more detail as well as its correlation to conventional radiographs.
In conclusion, in patients with recent onset IBP, individually trained central readers disagree as much as clinical practice local rheumatologists/radiologists in recognizing radiographic sacroiliitis. While the two central readers disagree with each other in a balanced manner (disagreement in both directions, reflecting measurement error), the local readers primarily overrate sacroiliitis in comparison with central readers, which results in an unacceptably high percentage of false-positive diagnoses of AS. A small minority of patients with a classification of AS according to central reading is not recognized in daily clinical practice. Independently of the precise definition of sacroiliitis, the disagreement regarding the presence/absence of sacroiliitis is so significant that the role of radiographic sacroiliitis as a diagnostic criterion for axSpA should be re-evaluated.

References

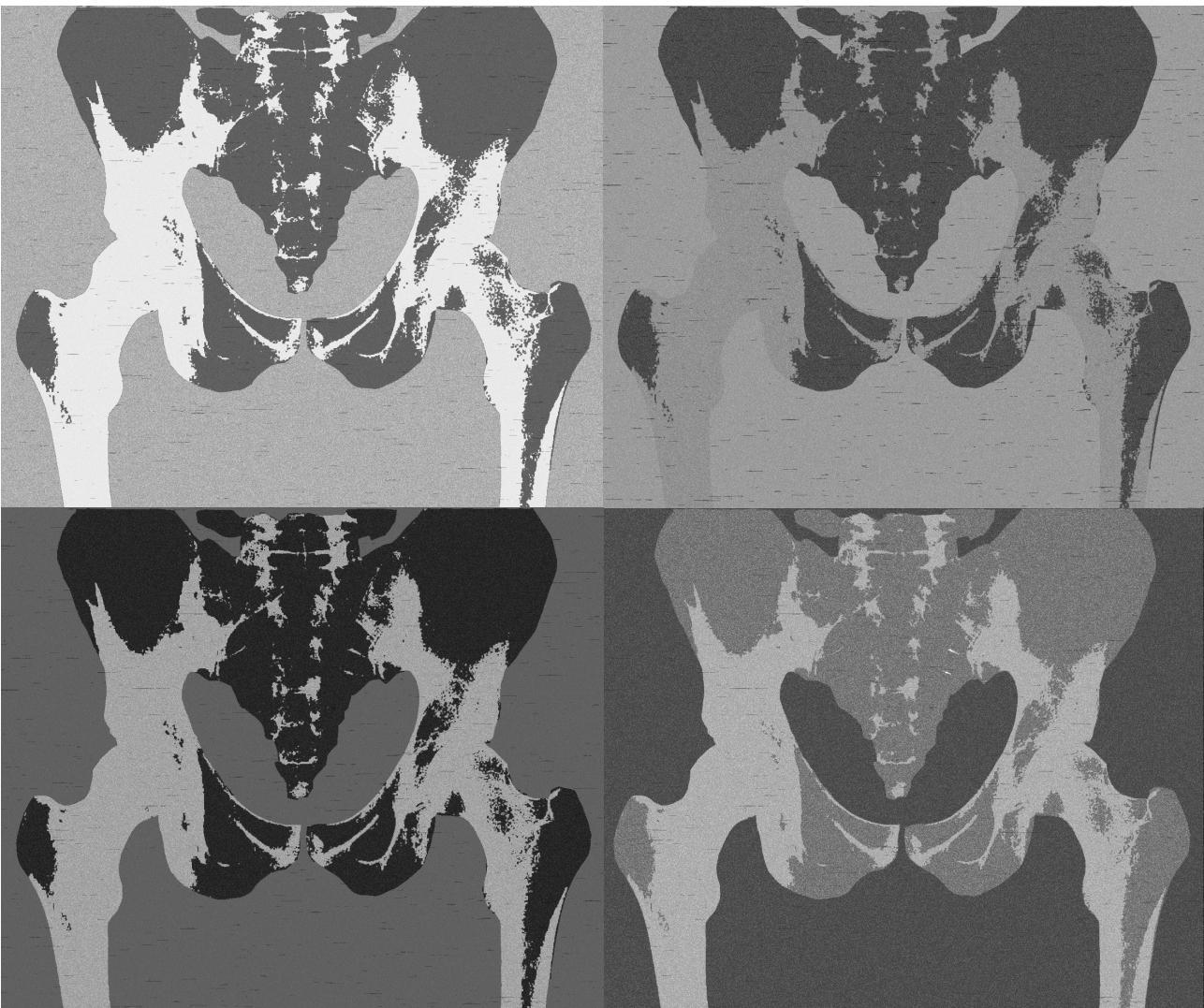
1. van der Linden S, Valkenburg H, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361-8.
2. Feldtkeller E, Khan M, van der Heijde D, *et al.* Age at disease onset and diagnosis delay in HLA-B27 negative vs. positive patients with ankylosing spondylitis. Rheumatol Int 2003;23:61-6.
3. Rudwaleit M, van der Heijde D, Landewé 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.
4. Khan M, van der Linden S, Kushner I, *et al.* Spondylitic disease without radiologic evidence of sacroiliitis in relatives of HLA-B27 positive ankylosing spondylitis patients. Arthritis Rheum 1985;28:40-3.
5. van den Berg R, de Hooge M, Rudwaleit M, *et al.* ASAS modification of the Berlin algorithm for diagnosing axial spondyloarthritis: results from the SPondyloArthritis Caught Early (SPACE)-cohort and from the Assessment of SpondyloArthritis international Society (ASAS)-cohort. Ann Rheum Dis 2013;72:1646-53.
6. Spoorenberg A, de Vlam K, van der Linden S, *et al.* Radiological scoring methods in ankylosing spondylitis. Reliability and change over 1 and 2 years. J Rheumatol 2004;31:125-32.
7. van Tubergen A, Heuft-Dorenbosch L, Schulpen G, *et al.* Radiographic assessment of sacroiliitis by radiologists and rheumatologists: does training improve quality? Ann Rheum Dis 2003;62:519-25.
8. van den Berg R, Stanislawska-Biernat E, van der Heijde DM. Comparison of recommendations for the use of anti-tumour necrosis factor therapy in ankylosing spondylitis in 23 countries worldwide. Rheumatology (Oxford) 2011;50:2270-7.
9. Dougados M, D'Agostino M, 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.
10. Calin A, Porta J, Fries J, *et al.* Clinical history as a screening test for ankylosing spondylitis. JAMA 1977;237:2613-4.
11. Rudwaleit M, Metter A, Listing J, *et al.* Inflammatory back pain in ankylosing spondylitis: a reassessment of the clinical history for application as classification and diagnostic criteria. Arthritis Rheum 2006;54:569-78.
12. Poddubnyy D, Gaydukova I, Hermann K, *et al.* Magnetic resonance imaging compared to conventional radiographs for detection of chronic structural changes in sacroiliac joints in axial spondyloarthritis. J Rheumatol 2013;40:1557-65.
13. Blachier M, Coutanceau B, Dougados M, *et al.* Does the site of magnetic resonance imaging abnormalities match the site of recent-onset inflammatory back pain? The DESIR-cohort. Ann Rheum Dis 2013;72:979-85.
14. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull 1968;70(4):213-20.
15. Dice L. Measures of the amount of ecologic association between species. Ecology 1945;26:297-302.
16. Landis J, Koch G. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
17. http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM366519.pdf. (24 December 2013, date last accessed).
18. van den Berg R, Lenczner G, Thévenin F, *et al.* Concordance between 'a positive MRI of the sacroiliac joints' based on the local reading versus a centralised reading: experience from the DESIR-cohort. Arthritis Rheum 2013;65[10]:S475.
19. 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.
20. Geijer M, Gothlin G, Gothlin J. The clinical utility of computed tomography compared to conventional radiography in diagnosing sacroiliitis. A retrospective study on 910 patients and literature review. J Rheumatol 2007;34:1561-5.

6

Classification of axial SpA based on positive imaging (radiographs and/or MRI of the sacroiliac joints) by local rheumatologists or radiologists versus central trained readers in the DESIR-cohort

Ann Rheum Dis Published Online First: [24 June 2014]
doi:10.1136/annrheumdis-2014-205432

R. van den Berg
G. Lenczner
F. Thévenin
P. Claudepierre
A. Feydy
M. Reijnierse
A. Saraux
A. Rahmouni
M. Dougados
D. van der Heijde



Abstract

Objective

Investigating changes in patient classification (ASAS axSpA criteria) based on evaluation of images of the sacroiliac joints (MRI-SI and X-SI) by local and central readers.

Methods

The DESIR-cohort included patients with inflammatory back pain (IBP; ≥3 months, but <3 years), suggestive of axSpA. Local radiologists/rheumatologists (local reading) and two central readers (central reading) evaluated baseline images. Agreement regarding positive MRI (pos-MRI) between central readers and between local reading and central reading was calculated (kappas). Number of patients classified differently (ASAS criteria) by using local reading instead of central reading was calculated.

Results

Interreader agreement between the two central readers and between local reading and central reading was substantial (κ=0.73 and κ=0.70, respectively). In 89/663 MRI-SIs (13.4%) local reading and central reading disagreed; 38/223 patients (17.0%) with pos-MRI (local reading) were negative by central reading; 51/440 patients (11.6%) with neg-MRI (local reading) were positive by central reading.
In 163/582 patients eligible for applying ASAS criteria (28.0%), local reading and central reading disagreed on positive imaging (MRI-SI and/or X-SI; κ=0.68). In 46/582 patients (7.9%) a different evaluation resulted in a different classification; 18/582 patients (3.1%) classified no SpA (central reading) were axSpA by local reading; 28/582 patients (4.8%) classified axSpA (central reading) were no SpA by local reading. Among axSpA patients (central reading), 16/419 patients (3.8%) fulfilling imaging arm by central reading fulfilled clinical arm by local reading; 29/419 patients (6.9%) fulfilling clinical arm by central reading fulfilled also imaging arm by local reading.

Conclusion

In patients with recent onset IBP, trained readers and local rheumatologists/radiologists agree well on recognizing a pos-MRI. While disagreeing in 28% of the patients on positive imaging (MRI-SI and/or X-SI), classification of only 7.9% of the patients changed based on a different evaluation of images, showing the ASAS axSpA criteria's robustness.

Introduction

The 2009 classification criteria for axial spondyloarthritis (axSpA) by the Assessment of SpondyloArthritis international Society (ASAS) are gaining more awareness and are increasingly being used to guide daily practice and include patients in clinical trials [1-3]. According to the ASAS axSpA criteria it is possible to classify patients with chronic back pain as axSpA via the clinical arm based on the presence of at least two SpA-features in addition to HLA-B27 positivity, or to classify patients via the imaging arm. In the presence of sacroiliitis on plain radiographs (modified New York (mNY) criteria) and/or MRI (ASAS definition of a positive MRI (pos-MRI)), a patient can be classified as axSpA if at least one additional SpA-feature should be present [1, 4, 5]. However, recognition of sacroiliitis on MRI and especially on plain radiographs is challenging [6-8]. It is known that interpretation of findings vary according to the expertise of the physician interpreting the image [7]. In daily practice, local radiologists and/or rheumatologists judge MRIs and radiographs of the SI-joints, frequently with knowledge of the clinical signs and symptoms, while in research cohorts and clinical trials ≥1 trained reader - blinded for clinical information - judge the images. As the classification as axSpA is heavily based on sacroiliitis, the classification of a patient could change as another reader judges the same MRI and/or radiographs differently. The ABILITY-1 trial included patients with non-radiographic axSpA (nr-axSpA), based on readings of the pelvic radiographs by local radiologists or rheumatologists. A post-hoc central reading (for another purpose) was performed and based on this reading, 37% of the patients classified as nr-axSpA by local sites were reclassified as fulfilling the mNY criteria [3, 9]. In another trial, the RAPID-axSpA trial, a similar analysis was performed resulting in reclassification of 36% of the patients (26% reclassified as fulfilling the mNY criteria, and 10% reclassified as nr-axSpA, based on the central reading in contrast to the local reading) [2, 10].

As sacroiliitis by two imaging methods as well as HLA-B27 positivity play an important role in the ASAS axSpA criteria, a patient will not necessarily be classified differently based on another reading of the radiograph and/or MRI of the SI-joints. Therefore, we investigated the change in classification of patients according to the ASAS axSpA criteria based on the evaluation of local and central readers of the same set of images. We performed this investigation in the DESIR (DEvenir des Spondylarthropathies Indifférenciées Récentes)-cohort, which has information on MRIs and radiographs of the SI-joints scored by the local rheumatologist or radiologist and also by two trained central readers.

Methods

Patients

Baseline data from the DESIR-cohort was used for this analysis. The DESIR-cohort is described extensively before [11]. In short, consecutive patients aged 18-50 with Inflammatory Back Pain (IBP) in the thoracic and/or lumbar spine and/or the buttock area (≥3 months, but <3 years) fulfilling either the Calin (4/5 criteria) or the Berlin (2/4 criteria) for IBP and a suspicion of SpA by the rheumatologist with a score of ≥ 5 on a scale of 0 to 10 (where 0 was not suggestive of axSpA and 10 was very suggestive of axSpA) from 25 centers in France were included in this prospective longitudinal cohort [12, 13]. In total, 708 patients were included between December 2007 and April 2010. The study is approved by the appropriate medical ethical committee and fulfilled Good Clinical Practice Guidelines. Before patients were included in the study, they gave written informed consent.

A detailed description of the study protocol is available at the website (http://www.lacohortedesir.fr/desir-in-english/). The research proposal for this particular analysis was approved by the scientific committee of the DESIR-cohort.

Data collection

With the use of a standardized CRF a database was built. According to the DESIR protocol the following data, among others, were collected: physical examination, ongoing treatment, co-morbidities, laboratory tests and questionnaires [11]. The database for the baseline data used for this analysis was locked on October 30th 2012.

Images and scoring methods

In each participating center MRIs of the SI-joints (MRI-SIs) were performed at baseline, with magnetic fields between 1.0 and 1.5T, using T1-FSE and STIR sequences with 12-15 semi-coronal slices of 4 mm thickness, parallel to the long axis of the sacrum, without the use of a contrast agent. All initial MRI-SIs were checked on quality by a central reader in Montpellier, and regular calibration by the manufacturer was required. Plain radiographs of the pelvis (X-SI) were performed in anteroposterior view at baseline.

All available baseline MRI-SIs (n=663) were scored by a local radiologist/rheumatologist who might have had access to all clinical and laboratory data at each participating center (local reading) [14]. Each SI-joint on MRI was assessed on the presence/absence of inflammation by answering the following question on the CRF '*Are there characteristic acute/active inflammatory lesions compatible with axial spondyloarthritis of the sacroiliac joints or entheses, outside the sacroiliac joints? Normal (score 0), doubtful (score 1) or abnormal (score 2).*' On the CRF, inflammatory lesions were defined as 'Bone edema/contrast product uptake in or adjacent to the sacroiliac joints or entheses (compatible with active lesions observed in cases of ankylosis spondylitis/axial spondyloarthritis; STIR and/or T1 sequences with gadolinium injection are required).' In this reading, a pos-MRI was defined as a score of 2 in at least one of the SI-joints.

Two central readers (RvdB and FT), experienced in scoring MRI-SIs, participated in a calibration training on reading MRI-SIs according to the ASAS definition. MRI-SIs were considered positive according to the ASAS definition if BME lesions highly suggestive of SpA were present if ≥1 BME lesion on ≥2 consecutive slices, or if several BME lesions are visible on a single slice. The presence of only synovitis, enthesitis or capsulitis without BME is not sufficient for a positive MRI-SI [5]. During the calibration session, executed by two senior radiologists (MR and AF) and two senior rheumatologists (PC and MD), supervised by an expert in AS and imaging scoring (DvdH), definitions of lesions, examples and pitfalls were discussed, followed by a supervised reading of training cases by the two readers. After this calibration session, 30 blinded MRI-SIs were read independently by the two readers (κ=0.30; positive agreement 73.7%; negative agreement 54.5%). A consensus meeting followed with the same group. Six weeks later, a second set consisting of 20 blinded MRI-SIs, were read independently by the two readers, again followed by a consensus meeting with the same group. After this second training session, agreement between the two readers was considered sufficient so the readers could start reading the DESIR-cohort (κ=0.74; positive agreement 80.0%; negative agreement 93.3%).

All available baseline MRI-SIs were read independently by the two readers, blinded for all clinical and laboratory data, the other imaging modality, as well as the local readings. Agreement on presence/absence of a pos-MRI was calculated and in case of disagreement, one of the senior radiologists involved in the calibration session (MR) served as adjudicator and scored the MRI-SI blinded to the information of the primary readers. An image was marked as pos-MRI (central reading) if 2/3 readers agreed.

The evaluation of X-SIs (n=688) by both local readers and central readers has been described before [8]. In short, the calibration of the two central readers (RvdB and GL) was performed in a similar way as for MRI-SI. Based on the mNY criteria, sacroiliitis was defined as grade ≥2 bilaterally or grade 3-4 unilaterally by central reading (pos-X-SI) [4]. The local readers evaluated X-SIs according to a method derived from the mNY. Since the local readers, who are working

in regular clinical practice, were not trained experts it was considered more appropriate use a scoring system that better resembles common clinical practice than the mNY criteria do. Local readers were asked to rate each SI-joint either as 'normal' or as 'doubtful sacroiliitis' or as 'obvious sacroiliitis' or as 'SI-joint fusion'. In this analysis at least a unilateral rating of 'obvious sacroiliitis' was considered a pos-X-SI for local reading. This has been explained in more detail before [8].

Statistical analysis

Agreement was calculated using cross-tabulation expressed in Cohen's Kappa (κ), agreement on positive cases (positive agreement) and on negative cases (negative agreement) for the following comparisons (online supplementary text 1) [15-18]: interreader agreement between the two central readers, agreement between local reading and central reading and between local reading and the two individual central readers on the presence/absence of a pos-MRI. Central reading was considered the external standard.
Next, the number of patients with a different MRI-SI and/or X-SI read using local reading instead of central reading was calculated, followed by the number of patients classified differently according to the ASAS axSpA criteria. This was done both for overall fulfillment and fulfillment of the imaging versus clinical arm.
SPSS software version 20.0 was used for the statistical analysis.

Results

The mean age of patients with available MRI-SI (n=663) was 31.7 (SD 8.7) years, mean symptom duration was 17.8 (SD 10.5) months, 309 (46.6%) patients were men and 387 (58.4%) were HLA-B27 positive.
Finally, in 15 patients X-SI was missing resulting in 648 patients with complete imaging. In 66/648 patients with complete imaging, IBP onset was >45 years and therefore the ASAS axSpA criteria could not be applied, leaving 582 patients (figure 1). Patient characteristics of these 582 patients were very similar to the patients with complete MRI-SI; mean age was 31.5 (SD 7.2) years, mean symptom duration was 18.3 (SD 10.6) months, 277 (47.7%) patients were men and 350 (60.1%) were HLA-B27 positive.

Agreement on a positive MRI

Interreader agreement between the two central readers regarding a pos-MRI is substantial (κ=0.73; table 1); 84/663 MRI-SIs (12.7%) were adjudicated because of disagreement.

Table 1: Interreader agreement between central reader 1 and central reader 2

	Total n=663	**Central reader 2**	
		MRI-SI+ (ASAS)	**MRI-SI- (ASAS)**
Central reader 1	**MRI-SI+ (ASAS)**	200	56
	MRI-SI - (ASAS)	28	379
	κ (95% CI): 0.73 (0.67-0.78)	Positive agreement: 82.6%	Negative agreement: 90.0%

Positive agreement is the agreement on positive cases. Negative agreement is the agreement on negative cases.

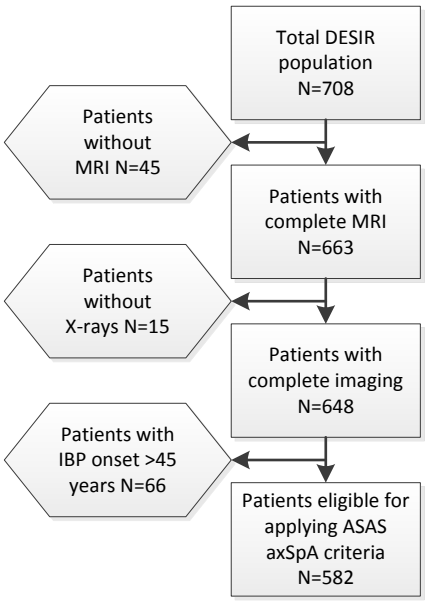


Figure 1: Flowchart of patients included in the DESIR-cohort and included in this analysis.

According to central reading, 236/663 patients (35.6%) had a pos-MRI; according to local reading, 33.6% had a pos-MRI. Agreement between local reading and central reading was also substantial (κ=0.70). In 13.4% of the MRI-SIs, local reading and central reading disagreed; 38/223 patients (17.0%) with a pos-MRI according to local reading, were read negative by central reading; 51/440 patients (11.6%) without a pos-MRI according to local reading, were read positive by central reading (online supplementary text 2). Comparisons of local reading versus the individual central readers show very similar results (table 2). There was no difference in agreement between local reading and central reading if MRI-SIs were read by local rheumatologists (n=174) or by local radiologists (n=457) (n=32 read by both a radiologist and a rheumatologist; data not shown).

Classification of patients according to the ASAS axSpA criteria

In this paragraph we focus only on the 582 patients in which the ASAS axSpA criteria could be applied. In 28.0% of the patients there was a disagreement on pos-imaging, MRI-SI and/or X-SI (κ=0.68). In 15.6% of the patients the disagreement was caused by a different X-SI read only (agreement on MRI-SI); in 10.1% the read of MRI-SI was different only (agreement on X-SI); and in 2.2% both X-SI and MRI-SI were read differently.

In total, 409 patients (70.2%) fulfilled the ASAS axSpA criteria based on local reading and 419 patients (72.0%) based on central reading. In 7.9% of the patients a different evaluation of imaging (MRI-SI and/or X-SI) resulted in a different classification. Eighteen patients were classified no SpA based on central reading but were classified axSpA based on local reading; in 28 patients it was the other way around (figure 2). In 14/18 and 13/28 patients, respectively, the different classification was the result of a different X-SI evaluation, consequently these patients changed from AS to no SpA and vice versa (table 3). The results of the comparison of local reading versus the individual central readers are similar (online supplementary table S2).

Table 2: Agreement between local reading and central reading, and between local reading and the individual central readers regarding presence/absence of sacroiliitis on MRI.

Local reading	**Total n=663**	**Central reading (2/3)**	
		MRI-SI+ (ASAS)	**MRI-SI- (ASAS)**
	MRI-SI+	185	38
	MRI-SI -	51	389
	κ (95% CI): 0.70 (0.65-0.76)	Positive agreement: 80.6%	Negative agreement: 89.7%
Local reading	**Total n=663**	**Central reader 1**	
		MRI-SI+ (ASAS)	**MRI-SI- (ASAS)**
	MRI-SI+	180	43
	MRI-SI -	76	364
	κ (95% CI): 0.61 (0.55-0.67)	Positive agreement: 75.2%	Negative agreement: 86.0%
Local reading	**Total n=663**	**Central reader 2**	
		MRI-SI+ (ASAS)	**MRI-SI- (ASAS)**
	MRI-SI+	177	46
	MRI-SI -	51	389
	κ (95% CI): 0.67 (0.62-0.73)	Positive agreement: 78.5%	Negative agreement: 88.9%

Positive agreement is the agreement on positive cases. Negative agreement is the agreement on negative cases.

Additional discrepancies were seen when interested in whether patients fulfill the imaging arm or the clinical arm within the ASAS axSpA criteria. By definition, patients fulfilling the clinical arm will always fulfill the clinical arm as HLA-B27 status will not change, but could fulfill the imaging arm as well, or not anymore, if a different evaluation of the same imaging set is used. Among the patients classified as axSpA based on central reading (n=419), 16 axSpA patients fulfilled the imaging arm based on central reading but fulfilled the clinical arm only based on local reading (in 8 patients due to a different X-SI read) (figure 2). When solely interested in whether patients fulfilled the imaging arm of the ASAS axSpA criteria or not, 44 patients fulfilled the imaging arm by central reading but not by local reading. Vice versa, 29 axSpA patients fulfilled the clinical arm only based on central reading but fulfilled the imaging arm based on local reading (in 13 patients due to a different X-SI read). Again, when interested in whether patients fulfill the imaging arm or not, 47 patients fulfilled the imaging arm by local reading but not by central reading (table 3). Comparisons of local reading versus the individual readers show similar results (online supplementary table S1).

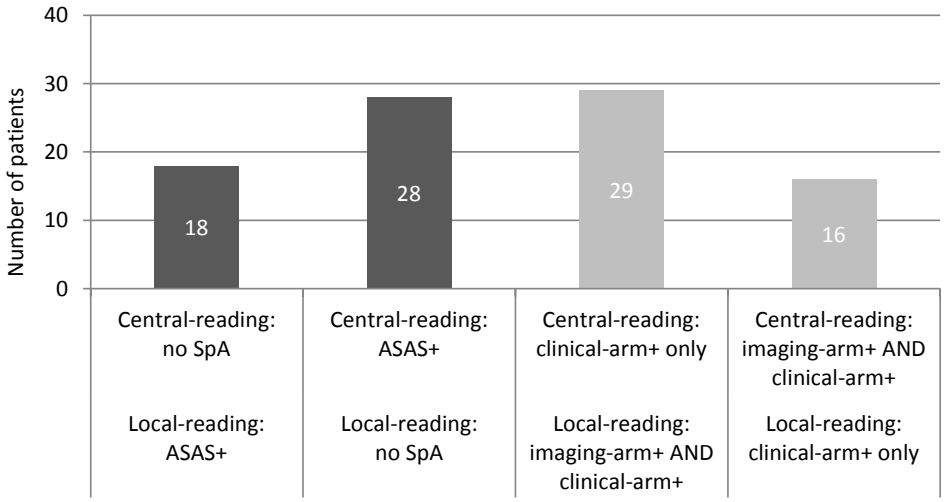


Figure 2: Bar graphs representing patients with SpA (ASAS+) according to one reading (local or central) but without SpA (dark grey) or a different arm of the ASAS criteria (light grey) according to the other reading (local or central) (total n=582, in 163 patients the imaging read was different between local and central reading).

Table 3: Classification of patients according to the ASAS axSpA criteria using Local reading instead of Central reading or the individual central readers.

Patients in which local and central reading (2/3 readers) differed, n=163			Central reading (2/3 readers)							
			Both arms +			Imaging arm+			Clinical arm+	No SpA (ASAS-)
			mNY+ MRI+	mNY+ MRI-	mNY- MRI+	mNY+ MRI+	mNY+ MRI-	mNY- MRI+		
Local reading	Both arms +	mNY+ MRI+	0	4	29				3	
		mNY+ MRI-	8	0	0				10	
		mNY- MRI+	8	1	0				16	
	Imaging arm+	mNY+ MRI+				0	1	13		4
		mNY+ MRI-				3	0	1		10
		mNY- MRI+				4	0	0		4
	Clinical arm+		1	7	8					
	No SpA (ASAS-)					3	10	15		

Both arms; patients fulfil both the imaging arm and the clinical arm. Imaging arm; patients fulfil the imaging arm only. Clinical arm; patients fulfil the clinical arm only. MRI+; sacroiliitis on MRI. mNY+; sacroiliitis on radiograph. Boxes in grey are empty as a patient fulfilling the clinical arm by definition will always fulfil the clinical arm as HLA-B27 status will not change, regardless of a different reading of images.

Discussion

In the DESIR-cohort, agreement between two trained central readers as well as between central reading and local reading on pos-MRI was substantial, thereby comparable to levels of agreement reported in a study designed to test inter- and intrareader agreement between experienced radiologists on a pos-MRI (κ=0.79-0.85) [19]. Though, it should be noted that at the start of the DESIR-cohort, the ASAS definition of a positive MRI-SI was not published yet. The levels of agreement of pos-MRI in the DESIR-cohort were higher than levels of agreement on pos-X-SI in the same cohort (κ=0.46-0.55). In addition, where misclassification by local reading regarding X-SIs almost exclusively consisted of overclassification of positive cases, the disagreement regarding MRI-SI is more balanced (as many positive as negative misclassifications) [8].

Our data provide interesting information of what would happen in case of testing eligibility of patients for clinical trials. Potentially 163/582 patients in which MRI-SI and/or X-SI reading was different between local reading and central reading could have a different classification according to the ASAS axSpA criteria. If patients in the DESIR-cohort would have been included in a clinical trial requiring fulfillment of mNY criteria based on local reading, 76/183 (41.5%) of the patients would not have fulfilled the mNY criteria by central reading. Similarly, 38/505 (7.5%) of the patients would be included based on central reading but not based on local reading [8]. Assuming a requirement of sacroiliitis on MRI according to local reading, 38/223 (17.0%) of the patients included would not be eligible based on central reading; the other way around, 51/440 patients (11.6%) not eligible for inclusion based on local reading would be included based on central reading. However, if inclusion would have been based on fulfillment of the imaging arm of the ASAS axSpA criteria the total percentage of reclassified patients would be 15.6% (91/582); 44 patients (7.6%) eligible based on central reading would not be included based on local reading and 47 patients (8.1%) the other way around. Based on the fulfillment of the entire axSpA criteria this percentage is 7.9% (46/582 patients); 28 patients (4.8%) would be included based on central reading but not on local reading and 18 patients (3.1%) the other way around.

The effect of local versus central reading regarding fulfillment of mNY criteria became recently evident by data provided to the Food and Drug Administration (FDA). In both the ABILITY-I and RAPID-axSpA trial, over 25% of the patients were reclassified as fulfilling mNY criteria based on central reading while they were entered as nr-axSpA based on local reading [2, 3, 9, 10]. The DESIR-cohort confirms this disagreement between local and central readers in the largest cohort addressing this issue. Moreover, there are no data on this aspect for MRI-SI this far, so the data presented in this study are the first data on MRI-SI in a large group of patients. As X-SI reading is so unreliable, the question arises whether it would be an option to only conduct MRI-SI and leave out X-SI completely, especially if structural lesions on MRI-SI are considered as well. More data from other cohorts, including patients with long-standing disease, are necessary to address this question in more detail.

Without knowing the truth of the result of imaging, central reading based on a consensus score of 2/3 readers, is the best approximation of the truth, followed by the reading of one central reader trained in the scoring, followed by local reading, (readers not specifically trained for this purpose). The choice for local reading or central reading for inclusion in clinical trials depends also on the purpose: if the aim is to test a drug in the way it will be applied in clinical practice, local reading would be preferred; if the aim is testing efficacy in the purest population, central reading would be preferred. The latter is mostly required by registration agencies. Furthermore, the European Medical Agency has approved TNF-inhibitors for patients with nr-axSpA only if additional signs of objective inflammation such as elevated CRP and/or a pos-MRI are present, while in patients fulfilling the mNY no additional sign of objective inflammation is required. Looking at all axSpA patients (including patients fulfilling the clinical arm) in the DESIR-cohort and assuming eligibility of all patients

for treatment with TNF-inhibitors (i.e. assuming that patients in the clinical arm had signs of objective inflammation and that all patients had active disease), 18 patients could have had inappropriate treatment with TNF-inhibitors and 28 patients were not treated with TNF-inhibitors based on false classification by local reading in comparison to the external standard of central reading. It should be noted that this situation implies an intrinsic dissimilarity in requirements to start with TNF-inhibitors based on the potentially fallible judgement on the presence or absence of radiographic sacroiliitis.

This study has several strengths we would like to address. The DESIR-cohort consists of a high number of patients, and in every patient both local reading and central reading of the same baseline set of images is available, thereby offering the unique opportunity to investigate the effect of local reading versus central reading. As patients were recruited in 25 centers where several rheumatologists and radiologists are working, local reading is a wide representation of clinical practice. Furthermore, central reading was performed by two independent trained readers and included an adjudication score, ensuring the robustness of central reading.

The main limitation of this study is that the DESIR-cohort only comprises patients with short disease duration. Patients with short symptom duration usually do not show extensive lesions, thereby making recognition of lesions in patients in the DESIR-cohort probably more difficult than in patients with established disease. Thus the results regarding agreement on positive imaging presented in this study might be slightly worse than could be expected in more established diseased patients. Another limitation is the fact that all sites were in France. It is unknown if this is generalizable to other countries. However, the two RCTs with similar percentages of disagreement in X-SI scores included many international sites across the world. Lastly, the role of structural damage on MRI-SI has not been taken into account. It would be interesting to know how the agreement between local and central reading is for this aspect, and if these structural changes could be taken into account in addition to or instead of the X-SI.

In conclusion, substantial levels of agreement between the two central readers and between local reading and central reading indicate that both local rheumatologists/radiologists and trained readers performed well in recognizing a pos-MRI in patients with recent onset IBP. However, when taking into account the reading of X-SI as well, levels of agreement between local reading and central reading are decreasing, yet it is reassuring that only 7.9% of the patients in the DESIR-cohort were classified differently using the full ASAS axSpA criteria, based on a different reading of the same set of images by local reading and central reading. These results point out the robustness of the ASAS axSpA classification criteria to differences in reading of the images, showing that these criteria can be applied reliably in clinical practice.

Supplementary data

Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2014-205432).

References

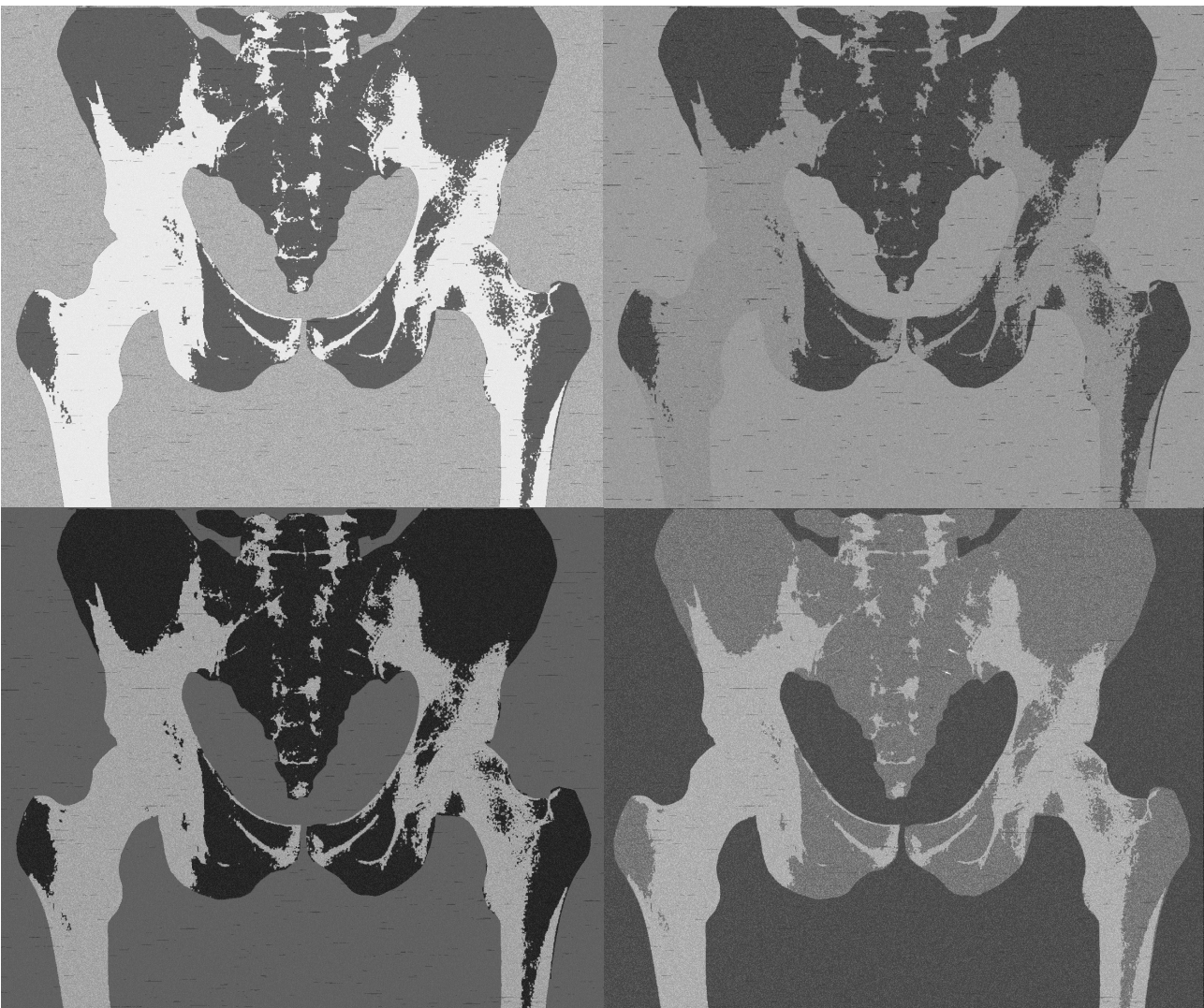
1. Rudwaleit M, van der Heijde D, Landewé R, *et al.* The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 2009;68:777-83.
2. Landewé R, Braun J, Deodhar A, *et al.* Efficacy of certolizumab pegol on signs and symptoms of axial spondyloarthritis including ankylosing spondylitis: 24-week results of a double-blind randomised placebo-controlled Phase 3 study. Ann Rheum Dis 2014;73:39-47.
3. Sieper J, van der Heijde D, Dougados M, *et al.* Efficacy and safety of adalimumab in patients with non-radiographic axial spondyloarthritis: results of a randomised placebo-controlled trial (ABILITY-1). Ann Rheum Dis 2013;72:815-22.
4. van der Linden S, Valkenburg H, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361-8.
5. Rudwaleit M, Jurik A, Hermann K, *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.
6. van Tubergen A, Heuft-Dorenbosch L, Schulpen G, *et al.* Radiographic assessment of sacroiliitis by radiologists and rheumatologists: does training improve quality? Ann Rheum Dis 2003;62:519-25.
7. Marzo-Ortega H, McGonagle D, Bennett A. Magnetic resonance imaging in spondyloarthritis. Curr Opin Rheumatol 2010;22:381-7.
8. van den Berg R, Lenczner G, Feydy A, *et al.* Reading of sacroiliac joints on plain pelvic radiographs: agreement between clinical practice and trained central reading. Results of the DESIR-cohort. Arthritis Rheumatol 2014;66:2403-11.
9. http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM366517.pdf. (24 December 2013, date last accessed).
10. http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM366519.pdf. (24 December 2013, date last accessed).
11. Dougados M, D'Agostino M, 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.
12. Calin A, Porta J, Fries J, *et al.* Clinical history as a screening test for ankylosing spondylitis. JAMA 1977;237:2613-4.
13. Rudwaleit M, Metter A, Listing J, *et al.* Inflammatory back pain in ankylosing spondylitis: a reassessment of the clinical history for application as classification and diagnostic criteria. Arthritis Rheum 2006;54:569-78.
14. Molto A, Paternotte S, van der Heijde D, *et al.* Evaluation of the validity of the different arms of the ASAS set of criteria for axial spondyloarthritis and description of the different imaging abnormalities suggestive of spondyloarthritis: data from the DESIR-cohort. Ann Rheum Dis [online 3-1-2014] doi: 10.1136/annrheumdis-2013-204262.
15. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull 1968;70:213-20.
16. Dice L. Measures of the amount of ecologic association between species. Ecology 1945;26:297-302.
17. Landis J, Koch G. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
18. Cicchetti D, Feinstein A. High agreement but low kappa: II. Resolving the paradoxes. J Clin Epidemiol 1990;43:551-8.
19. Arnbak B, Jensen T, Manniche C, *et al.* Spondyloarthritis-related and degenerative MRI changes in the axial skeleton - an inter- and intra-observer agreement study. BMC Musculoskelet Disord 2013;14:274.

7

Metric properties of the SPARCC-score of the sacroiliac joints – Data from baseline, 3 and 12 months follow-up in the SPACE-cohort

Submitted

R. van den Berg
M. de Hooge
P. Bakker
F. van Gaalen
V. Navarro-Compán
K. Fagerli
R. Landewé
M. van Oosterhout
R. Ramonda
M. Reijnierse
D. van der Heijde



Abstract

Objectives

To evaluate metric properties of the SPARCC-score of the sacroiliac joints.

Methods

Patients ≥16 years with back pain (≥3 months, ≤2 years, onset <45 years) were included in the SPondyloArthritis Caught Early (SPACE)-cohort. Patients with (possible) axial spondyloarthritis (axSpA) had follow-up visits after 3 and 12 months. Patients were treated according to usual clinical practice. MRI-SIs were scored in two independent campaigns (1: baseline to 3 months and 2: baseline to 3 months to 12 months) by two different blinded reader pairs, applying the ASAS definition (positive versus negative MRI-SI) (discordant cases were adjudicated by a third reader) and the SPARCC-score (mean of two agreeing readers) was obtained. Agreement (kappa; positive/negative agreement) between SPARCC-score cut-off values and a consensus judgment of a positive MRI (ASAS definition) as external standard, change in SPARCC-score and smallest detectable changes (SDCs) over 3 and 12 months were calculated.

Results

SPARCC-score ≥2 showed best agreement with a positive MRI (both campaigns). In campaign 1, SPARCC-score changed (increased/decreased) in 70/151 patients; 26/70 change >SDC (3.4) of which 20 on stable treatment. In campaign 2, 20/68 patients changed in SPARCC-score; 11/20 change >SDC (2.1) of which 8 patients on stable treatment (3 months). Over 1 year, 23/74 patients changed in SPARCC-score; 14/23 change >SDC (2.4) of which 7 on stable treatment.

Conclusions

SPARCC-score ≥2 can be used as a surrogate for a consensus judgment of a positive MRI (ASAS definition) in clinical trials. The SDCs ranged from 2.1-3.4 dependent on reader pair and these are close to the proposed minimum important change of 2.5.

Introduction

A positive MRI of the sacroiliac joints according to the ASAS definition ('positive-MRI') [1] is part of the ASAS axial spondyloarthritis (axSpA) criteria [2] and is increasingly used to test eligibility of axSpA patients for clinical trials [3-5]. Within clinical trials, MRI-SI is often repeated over short periods of time (e.g. 12 weeks) to test efficacy of (especially biological) treatment in terms of changes in inflammation. For this efficacy read, the SPondyloArthritis Research Consortium of Canada (SPARCC)-score is frequently used as it measures inflammation on a continuous scale with good sensitivity to change [6, 7]. It is unknown what SPARCC-score cut-off value the equivalent is of a 'positive-MRI', which is needed to link the read for eligibility and the efficacy reading. This information would be useful for example to define groups with MRIs scored according to SPARCC-scores as having either or not a 'positive-MRI', to study differences in treatment response over time [3].

Treatment with biologicals may dramatically influence inflammatory signs on MRI [8-11] but inflammation may also spontaneously change over time in patients without treatment and in patients on stable non-biological treatment [12-14]. However, it is not clear how many SPARCC-score units these spontaneous changes represent [12-14]. Moreover, these spontaneous changes are likely to be different with variable lengths of follow-up. A minimally important change (MIC) of 2.5 SPARCC-units is proposed based on the patient global assessment as external anchor [15]. It is known that interreader reliability of SPARCC-scores at a fixed time point is acceptable to high (ICC 0.69-0.96) [16, 17], but reliability on change in SPARCC-scores over time has sparsely been reported and appeared to be moderate (ICC 0.52) in one small study with 20 patients [7]. Therefore, it would be of additional value to have knowledge about interreader reliability in terms of smallest detectable change (SDC), in order to be able to judge whether the SDC is sufficiently small to detect the proposed MIC.

The aim of this study is threefold: first, to define which SPARCC-score best approximates a 'positive-MRI' judgment; second, to establish an SDC for a 3-month period and for a 1-year period; third, to describe which variation in SPARCC-score over a 3-month and 1-year period can be expected in patients without (change in) treatment.

Methods

Study population

Data from the SPondyloArthritis Caught Early (SPACE)-cohort is used for this analysis. An extensive description of the SPACE-cohort is given elsewhere [18]. In short, the SPACE-cohort is an ongoing cohort started in January 2009, including patients aged 16 years and older with back pain (≥3 months, ≤2 years, onset <45 years) visiting the rheumatology outpatient clinics of five participating centers. Patients were not included if they had other painful conditions (not related to SpA) that could interfere with the evaluation of the disease. After signing informed consent, all patients underwent a diagnostic work-up at baseline, including MRI and plain radiographs of the SI-joints, HLA-B27 testing and examining for other SpA-features. Patients fulfilling the ASAS axSpA criteria or patients with possible axSpA were included for follow-up visits after 3 and 12 months. Possible axSpA was defined as the presence of at least one specific SpA-feature with a high positive likelihood ratio (LR+ above 6) or at least two less specific SpA-features (LR+ below 6), but not fulfilling the ASAS axSpA criteria [19].

MRI-SI

MR imaging was performed on a 1.5T scanner, acquiring T1-weighted Turbo Spin Echo (T1TSE) (TR 550/TE 10) and Short Tau Inversion Recovery (STIR) (TR 2500/TE 60) sequences, obtaining slices of 4mm thickness in coronal oblique view of the SI-joints.

All readers in this study were extensively trained in reading MRIs according to the ASAS definition and the SPARCC-score during a calibration session, supervised by a senior radiologist (MR) and a senior rheumatologist (DvdH), discussing definitions of lesions, examples and pitfalls. Next, all readers independently read 30 blinded MRIs to calculate agreement (κ=0.75 to κ=0.87 for the different pairs of readers), followed by a consensus meeting in which the supervising rheumatologist and radiologist of the calibration session participated too. The agreement was considered sufficiently high to start scoring the SPACE-cohort.

Two reading campaigns were performed, at different moments in time, by different pairs of readers (RvdB and MdH in campaign 1; PB and MdH in campaign 2) with partly overlapping patients and images. Patients in the first reading campaign were included between January 2009 and November 2012 in five different centers and patients in the second reading campaign were included between January 2009 and October 2013 in one center. In campaign 1, baseline and 3-month MRI-SIs were evaluated; in campaign 2, baseline, 3-month and 1-year MRI-SIs were evaluated. In both campaigns, MRI-SIs were independently read by the two trained readers on the fulfilllment of the ASAS definition [1] and according to the SPARCC-score [6], blinded for the time sequence of the MRI-SIs as well as for clinical and laboratory data.

An MRI-SI can be marked positive according to the ASAS definition if ≥1 bone marrow edema (BME) lesion highly suggestive of SpA is present on ≥2 consecutive slices, or if several BME lesions highly suggestive of SpA are visible on a single slice. The presence of only synovitis, enthesitis or capsulitis without BME is not sufficient for a positive MRI-SI [1]. In case the two readers disagreed on the presence of a 'positive-MRI', a third trained reader served as adjudicator (VNC in campaign 1; RvdB in campaign 2).

According to the SPARCC-score, the presence of increased signal corresponding to BME lesions highly suggestive of SpA is marked on the six middle slices of an MRI-SI, representing the largest proportion of the synovial compartment of the SI-joints. Each SI-joint is divided into four quadrants (upper iliac, lower iliac, upper sacrum and lower sacrum). The maximum score for two SI-joints on each slice is eight. In addition, a score for 'intensity' may be assigned to each SI-joint if an 'intense signal' is seen in any quadrant on each slice resulting in a maximum score of 12. The signal from presacral blood vessels defined a lesion that is scored as intense. Furthermore, a score for 'depth' may be assigned to each SI-joint if an homogeneous and unequivocal increase in signal is extending over a depth of at least 1 cm from the articular surface on each slice resulting in a maximum score of 12. A lesion is graded as deep if there is a homogeneous and unequivocal increase in signal extending over at least 1 cm from the articular surface. The total maximum SPARCC-score is 72 [6]. The mean SPARCC-scores of the two readers were used; in case there was a third reader involved, the mean of the SPARCC-scores of the two readers in agreement of a 'positive-MRI' for that particular case were used.

Treatment

Patients in the SPACE-cohort are not treated according to a fixed protocol, but according to usual clinical practice by their rheumatologist. Treatment with NSAIDs was recorded according to the ASAS recommendations, resulting in a 0-100 score whereby 0 means no NSAID intake at all, and 100 means a daily intake at a full dose over the whole period of interest [20]. Treatment with DMARDs and anti-TNF therapy was recorded as present or absent. To investigate variation in SPARCC-scores over time, patients were categorized according to their treatment over the period of interest: no treatment, stable NSAID and/or DMARD intake, and change in NSAID and/or DMARD intake. Patients receiving anti-TNF therapy during the period of interest were excluded from the analysis on variation in SPARCC-scores.

Statistical analysis

Baseline characteristics of patients in both groups were investigated using descriptive statistics. Agreement (Cohen's kappa) between MRI-positivity based on several SPARCC-score cut-off values (≥1, ≥2, ≥3 and ≥4) and the consensus judgment of a 'positive-MRI' , as external standard, was calculated using cross-tabulation. Agreement on positive cases (positive agreement) and on negative cases (negative agreement) was also calculated [21].
Changes in SPARCC-score over the period of interest (baseline - 3 months (both campaigns); baseline - 1 year (campaign 2)), were visualized in cumulative probability plots in which patients were grouped based on treatment. Next, SDCs were calculated based on a 95% level of agreement (95%LoA) between the two readers on the change scores for both baseline to 3-month and baseline to 1-year intervals, using the following formula: SDC = (1.96 * $SD_{\Delta change\text{-}scores}$) / (√2 *√k), whereby k represents the number of readers (equals 2 in this study) [22]. The SDCs are also displayed in Bland Altman plots, that plot the mean SPARCC-score changes of the two readers (X-axis) and the inter-reader differences in SPARCC-score changes (Y-axis). In addition, the mean of the inter-reader differences in SPARCC-score changes (which is a reflection of the systematic error between the two readers) and the 95% levels of agreement (LoA) are presented in these plots. SPSS software version 20.0 was used for statistical analysis.

Results

Patients with available baseline MRI-SI were included in the analysis of the agreement between the SPARCC-score cut-off value and 'positive-MRI' (n=294 (campaign 1) and n=249 (campaign 2)). There is a partial overlap (49.1%) between patients included in campaign 2 and those included in campaign 1. In both campaigns the population is young, with short symptom duration, around 1/3 of the patients is male and around 1/3 fulfilled the ASAS axSpA criteria (table 1).
A 3-month follow-up MRI-SI was available in 154 patients in campaign 1. However, 3/154 patients received anti-TNF therapy during this period and were therefore excluded from the follow-up part of the analysis of the SPARCC-score changes over time and SDCs. In campaign 2, a 3-month follow-up MRI-SI was available in 70 patients and in 76 patients a 1-year follow-up MRI-SI was available. Two patients received anti-TNF therapy, leaving MRI-SIs of 68 (campaign 1) and 74 patients (campaign 2) for follow-up analyses.

SPARCC-score cut-off

In both campaigns, there was a high level of agreement between MRI-positivity based on all tested SPARCC-score cut-off values and the consensus judgment of a 'positive-MRI' as external standard (table 2). A cut-off value of ≥2 showed the highest kappa values (0.94 in campaign 1 and 0.98 in campaign 2) and provided the best balance in terms of misclassifications in comparison to the external standard; 5 false-positive and 1 false-negative classifications in campaign 1; zero false-positive and 1 false-negative classification in campaign 2.

Smallest detectable change of SPARCC-score

Of the patients with available follow-up MRI, the mean SPARCC-score at baseline was 4.0 (SD 8.3) and 2.3 (SD 5.7) (campaign 1 and 2, respectively). At 3 months, the mean SPARCC-score was 3.4 (SD 6.7) and 1.6 (SD 3.8) (campaign 1 and 2, respectively), and at 1 year the mean SPARCC-score was 1.4 (SD 4.0) (campaign 2).
Bland and Altman plots show the mean of the two readers in SPARCC-score changes over the 3-month (campaign 1 and 2) and 1-year period (campaign 2) against the difference between the two readers in SPARCC-score changes over those periods (figure 1).

Table 1: Baseline characteristics of patients in reading campaign 1 and patients in reading campaign 2. A proportion (49.1%) of the patients was included in both campaigns.

	Reading campaign 1, n=294	**Reading campaign 2, n=249**
Age (years) at inclusion, mean ± SD	31.2 ± 10.4	31.1 ± 11.5
Male, n (%)	102 (34.7)	81 (32.5)
Duration of back pain (months), mean ± SD	13.1 ± 7.1	13.3 ± 7.4
HLA-B27 positive, n (%)	113 (38.4)	79 (31.7)
Pos. Fam. History SpA, n (%)	113 (38.4)	89 (35.7)
IBP, n (%)	195 (66.3)	142 (57.0)
Psoriasis, n (%)	28 (9.5)	26 (10.4)
Dactylitis, n (%)	16 (5.4)	8 (3.2)
Enthesitis, n (%)	49 (16.7)	24 (9.6)
Uveitis, n (%)	24 (8.2)	18 (7.2)
IBD, n (%)	20 (6.8)	19 (7.6)
Good response to NSAIDs, n (%)	112 (38.1)	69 (27.7)
Elevated CRP/ESR, n (%)	58 (19.7)	42(16.9)
Asymmetric lower limb arthritis, n (%)	48 (16.3)	26 (10.4)
Radiographic sacroiliitis*, n (%)	23 (7.8)	24 (9.6)
Sacroiliitis MRI**, n (%)	67 (22.8)	31 (12.4)
SPARCC-score, mean ± SD	2.9 ± 7.7	1.3 ± 4.4
CRP, mean ± SD	6.9 ± 13.0	7.3 ± 11.6
ASDAS, mean ± SD	2.6 ± 1.1	2.7 ± 0.8
BASDAI, mean ± SD	4.6 ± 2.5	4.6 ± 2.1
BASFI, mean ± SD	3.0 ± 2.3	3.2 ± 2.4
ASAS axSpA criteria positive, n (%)	119 (40.5)	83 (33.3)

*Radiographic sacroiliitis according to the modified New York criteria [25]. **Sacroiliitis on MRI according to the ASAS definition (consensus judgment) [1]. HLA-B27, Human Leukocyte Antigen; IBP, Inflammatory Back Pain; IBD, Inflammatory Bowel Disease; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate.

The plots show that a large number of observations is clustered around the mean difference of 0, and that differences between readers occur with similar amplitude across the entire range of the SPARCC-score (a homoscedastic pattern). To visualize the high number of overlapping observations, series of ranges were defined in which all observations were grouped into their corresponding range, exponentially displayed on the X-axis. The SDC in campaign 1 over the 3-month period is 3.4 SPARCC-units, depicted in figure 1a as the dark grey area reflecting the SDC of both increased and decreased SPARCC-scores over time. The SDC in campaign 2 over the 3-month period is 2.1 SPARCC-units (figure 1b) and over the 1-year period 2.4 SPARCC-units (figure 1c).

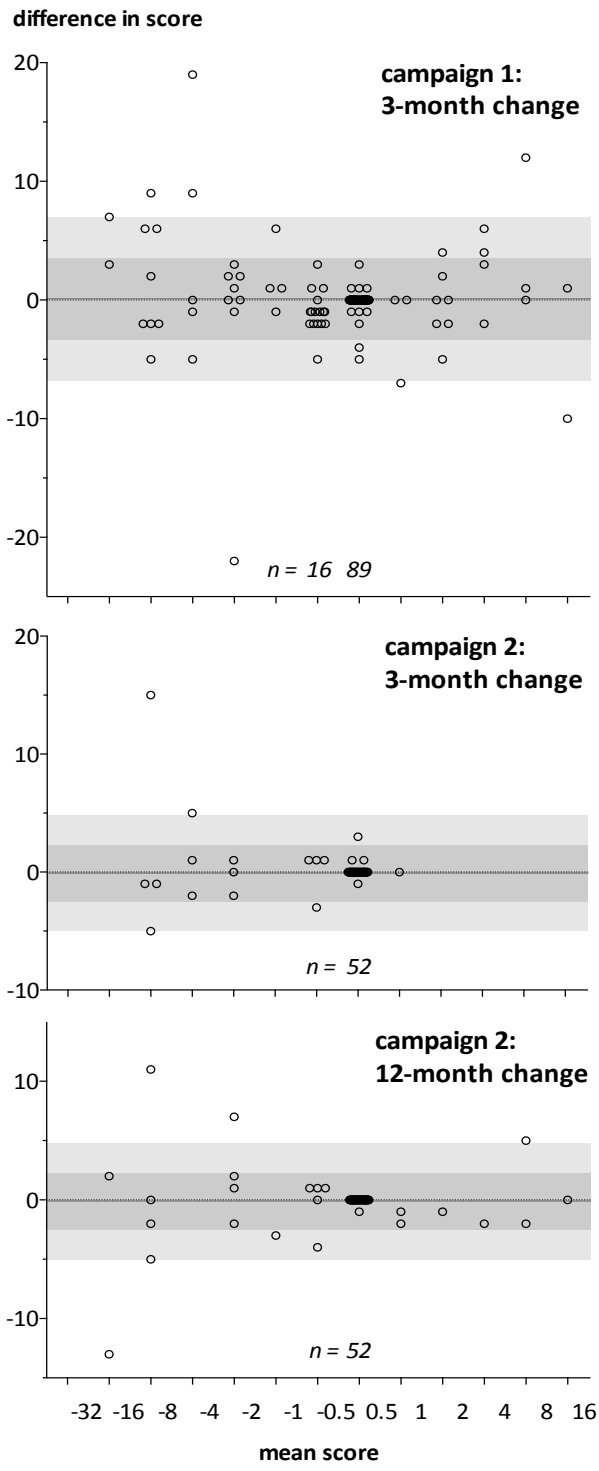


Figure 1: Bland Altman plots showing the mean SPARCC-score change of the two readers (X-axis) versus the delta SPARCC-score changes of the two readers (Y-axis). The large number of overlapping observations clustered around the mean difference of zero are displayed in series of ranges increasing exponentially on the positive side of zero and decreasing exponentially on the negative side (X-axis). The 'n' above the X-axis show the number of observations per group. The solid grey line represent the overall mean of the delta SPARCC-score changes (equivalent to systematic error between the two readers). The light grey area represents the 95% levels of agreement (LoA), and the dark grey area represents the smallest detectable change (SDC) in both directions (increase in SPARCC-score and decrease in SPARCC-score over time). The reader is referred to the text for further clarification.

Figure 1a: mean of the delta SPARCC-scores 0.1 (95% LoA -6.8 to 7.0); SDC 3.4. Observations are clustered in the range -0.5 to 0.5 (n=89) and the range -1 to -0.5 (n=16).

Figure 1b: mean of the delta SPARCC-scores 0.2 (95% LoA -4.0 to 4.4); SDC 2.1. Observations are clustered in the range -0.5 to 0.5 (n=52).

Figure 1c: mean of the delta SPARCC-scores -0.1 (95% LoA -5.0 to 4.8); SDC 2.4. Observations are clustered in the range -0.5 to 0.5 (n=52).

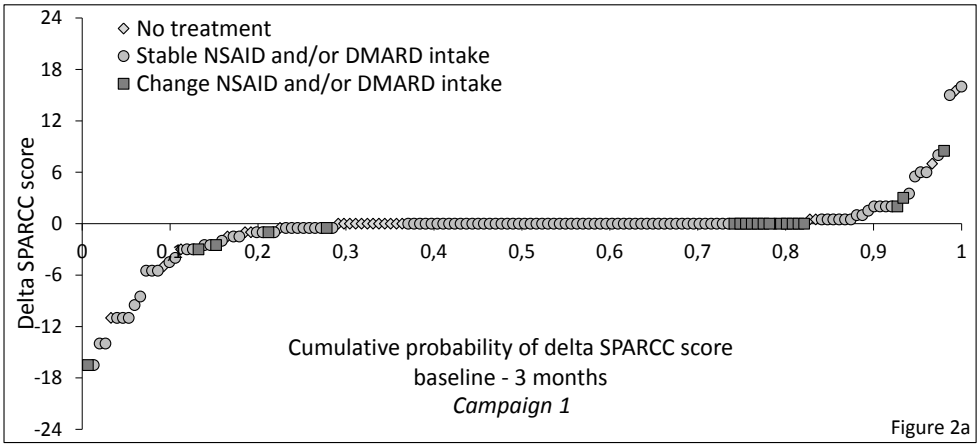
Table 2: Various SPARCC cut-off values tested against the ASAS definition of a positive MRI, in reading campaign 1 and reading campaign 2.

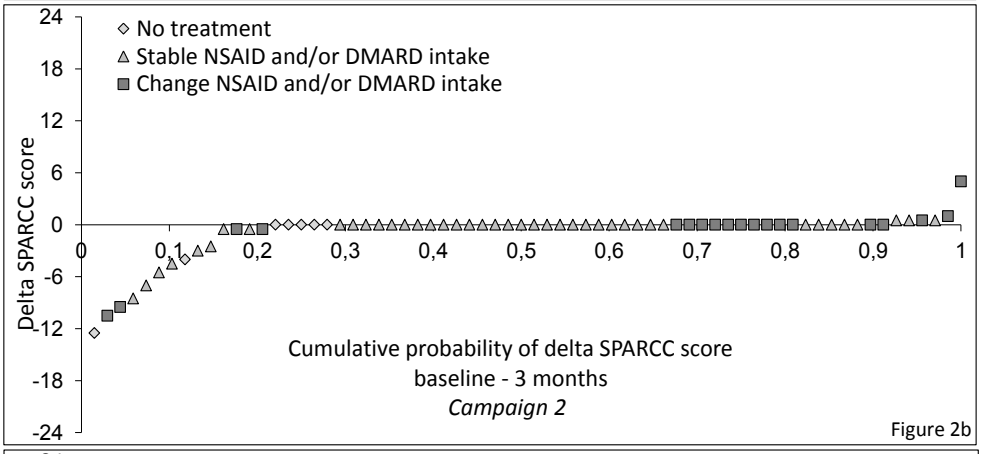
Reading campaign 1 (n=294)		
	positive MRI (ASAS)	**negative MRI (ASAS)**
SPARCC ≥1	67	21
SPARCC <1	0	206
Kappa: 0.82	PA: 95.2%	NA: 86.5%
SPARCC ≥2	66	5
SPARCC <2	1	222
Kappa: 0.94	PA: 98.7%	NA: 95.7%
SPARCC ≥3	57	1
SPARCC <3	10	226
Kappa: 0.89	PA: 97.6%	NA: 91.2%
SPARCC ≥4	47	1
SPARCC <4	20	226
Kappa: 0.77	PA: 95.6%	NA: 81.7%
Reading campaign 2 (n=249)		
	positive MRI (ASAS)	**negative MRI (ASAS)**
SPARCC ≥1	31	5
SPARCC <1	0	213
Kappa: 0.91	PA: 98.8%	NA: 92.5%
SPARCC ≥2	31	1
SPARCC <2	0	217
Kappa: 0.98	PA: 99.8%	NA: 98.4%
SPARCC ≥3	25	0
SPARCC <3	6	218
Kappa: 0.88	PA: 98.6%	NA: 89.3%
SPARCC ≥4	21	0
SPARCC <4	10	218
Kappa: 0.79	PA: 97.8%	NA: 80.8%

PA, positive agreement is the agreement on positive cases. NA, negative agreement is the agreement on negative cases.

Change in SPARCC-scores over 3 months and 1 year

Eighty-one out of 151 patients in campaign 1 (53.6%) showed no change in SPARCC-score over the 3-month period of which 75/81 (92.6%) had a SPARCC-score of 0 at both time points. In the 70 out of 151 patients (46.4%) showing a change in SPARCC-score, 27 increased and 43 decreased (mean change -1.1 (SD 6.3); median change -0.5 (range -16.5 to 16.0)) (figure 2a & table 3). In 26 out of 70 patients (37.1%) with SPARCC-score changes, the change was more than the SDC (3.4); 16 patients decreased (2 patients without treatment, 11 with stable NSAIDs intake, 2 with stable NSAIDs and DMARD intake, 1 patient started





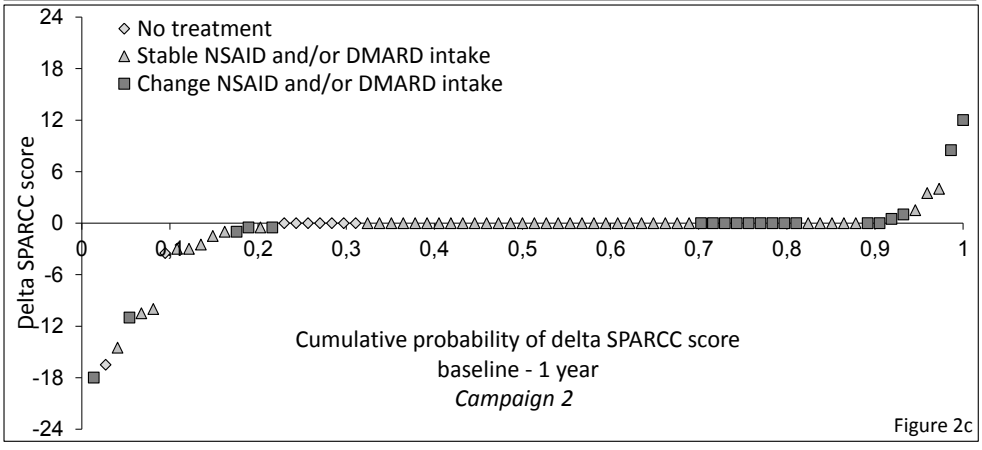


Figure 2: Cumulative probability plots of all delta SPARCC-scores over a 3-month period (2a and 2b) and a 1-year period (2c) with different symbols indicating the treatment over the investigated period.

Table 3: All changes in SPARCC-score in patients grouped according to treatment.

	No treatment	Stable NSAIDs/ DMARDs	Start NSAIDs/ DMARDs	Stop NSAIDs/ DMARDs
Campaign 1 – baseline to 3 months				
No SPARCC-score change	N=13	N=56	N=7	N=5
Increase in SPARCC-score (mean change (SD); range)	N=4 5.9 (SD 7.1) 0.5 to 15.5	N=20 3.7 (4.6) 0.5 to 16.0	N=3 4.5 (3.5) 2 to 8.5	-
Decrease in SPARCC-score (mean change (SD); range)	N=9 -2.8 (3.4) -11.0 to -0.5	N=29 -4.9 (4.9) -16.5 to -0.5	N=5 -4.7 (6.7) -16.5 to -0.5	-
Campaign 2 – baseline to 3 months				
No SPARCC-score change	N=4	N=31	N=8	N=5
Increase in SPARCC-score (mean change (SD); range)	-	N=5 0.6 (0.2) 0.5 to 1.0	N=1 5 (-) -	-
Decrease in SPARCC-score (mean change (SD); range)	N=2 -5.0 (6.4) -12.5 to -0.5	N=10 -4.6 (3.3) -10.5 to -0.5	N=2 -6.5 (8.5) -12.5 to -0.5	-
Campaign 2 – baseline to 1 year				
No SPARCC-score change	N=10	N=28	N=7	N=6
Increase in SPARCC-score (mean change (SD); range)	-	N=3 3.0 (1.3) 1.5 to 4.0	N=1 12 (-) -	N=3 3.3 (4.5) 0.5 to 8.5
Decrease in SPARCC-score (mean change (SD); range)	N=3 -6.8 (8.5) -16.5 to -0.5	N=8 -5.8 (5.1) -14.5 to -1.0	N=1 -0.5 (-)	N=4 -7.6 (8.4) -18.0 to -0.5

NSAIDs intake) and 10 patients increased (2 without treatment, 7 with stable NSAIDs intake, 1 started NSAIDs intake). In the remaining 44 patients (62.9%) the SPARCC-score changes were within the area still compatible with measurement error. intake but continued NSAID intake). In the remaining 9 patients (39.1%) SPARCC-score changes were not beyond measurement error.

In campaign 2, two follow-up intervals for the same patients are available. Over the 3-month period, SPARCC-score did not change in 48 out of 68 patients (70.6%); 46/48 patients (95.8%) had a SPARCC-score of 0 at both time points. In the remaining 20 patients (29.4%) the SPARCC-score changed; 14 patients showed a decrease and 6 patients an increase (mean change -3.1 (SD 4.6); median change -1.5 (range -12.5 to 5) (figure 2b & table 3). Eleven out of 20 patients (55.0%) showed a SPARCC-score change >SDC (2.1); 10 patients decreased (1 without treatment, 6 with stable NSAIDs intake, 2 with stable NSAIDs and DMARD intake, 1 started NSAIDs intake) and 1 patient increased (started NSAIDs intake). The remaining 9 patients (45.0%) had SPARCC-score changes still compatible with measurement error.

The results over the 1-year period in campaign 2 are similar to the results over the 3-month period in campaign 2, although more variation between patients is seen; 51/74 patients (68.9%) did not show a change in SPARCC-score, of which 50 patients (98.0%) had a SPARCC-score of 0 at both time points. The remaining 23 patients (31.1%) showed a change in SPARCC-score; 16 patients decreased and 7 increased (mean change -2.9 (SD 7.5); median

change -1.0 (range -18.0 to 12.0)) (figure 2c & table 3). Fourteen out of the 23 patients (60.9%) showed a SPARCC-score change of more than the SDC (2.4); 10 patients decreased (2 without treatment, 4 with stable NSAID intake, 2 with stable DMARD intake, 1 stopped NSAID intake, 1 started but stopped again NSAID intake) and 4 patients increased (1 with stable NSAID intake, 1 stopped NSAID intake, 1 started NSAID intake, 1 stopped DMARD
The majority of the patients showing changes in SPARCC-score of more than the SDC in both campaigns (20/26 (76.9%; campaign 1), 8/11 (72.7%; 3-month period campaign 2) and 7/14 (50.0%; 1-year period campaign 2)) were on stable NSAID and/or DMARD intake.

Discussion

This study performed in the SPACE-cohort has shown in two campaigns that a cut-off value of 2 SPARCC-units is best compatible with a consensus judgment of a positive versus negative MRI according to the ASAS definition. These results were not unexpected as the ASAS definition of a positive MRI-SI includes - apart from a qualitative part (BME lesions highly suggestive of spondyloarthritis) - a quantitative part that requires at least one BME lesion visible on at least 2 consecutive slices or several lesions on a single slice [1]. However, in theory, a SPARCC-score can be high because of the presence of several small lesions (highly suggestive of SpA), scattered over several slices (e.g. one lesion on slice 1, another lesion on slice 4 and another lesion on slice 6) but still not fulfilling the ASAS definition. A SPARCC-score can also be high if one lesion is assigned as 'intense' or 'deep', while it is only visible on 1 slice. Moreover, the SPARCC-score prescribes that lesions are scored in the six middle slices, while the ASAS definition takes all slices into account [1, 6]. Occasionally, part of a lesion may be visible on only one of the six middle slices, while the remaining part of the lesion is visible outside those six middle slices, or a slice outside those middle six shows several lesions. However, these considerations are mainly theoretical and do not appear very frequently. Therefore, a SPARCC cut-off level of 2 units may serve as a surrogate for the ASAS definition of a positive MRI and could be used in clinical trials with central efficacy reading in order to derive a dichotomy (positive versus negative) for prognostic reasons.
The SDCs in campaign 2 (2.1 SPARCC-units over 3 months and 2.4 over 1 year) are close to the proposed MIC of 2.5 SPARCC-units, which was calculated using pooled changes over 12 and 52 weeks [15], but the SDC of campaign 1 (3.4) is slightly higher. This suggests that the previously proposed MIC is close to measurement error in our study based on two different reader pairs and different periods of follow-up.
A large proportion of the SPARCC-score changes seen in the patients in both reading campaigns could be considered as noise as these changes are smaller than the SDCs (62.9% and 45% (3-months, campaign 1 and 2) and 39.1% (1-year in campaign 2). To investigate the influence of non-biological treatment on inflammation on MRI-SI, only patients with SPARCC-score changes greater than the SDC were taken into account. Somewhat surprisingly, the majority of patients with a change in SPARCC-score were on stable NSAID and/or DMARD treatment. Some patients taking stable doses of NSAIDs increased in SPARCC-score while others who were also on stable NSAIDs intake decreased in SPARCC-score. These results are in line with the results found in trials where patients using NSAIDs – either in an open label trial or in a placebo group – showed also both increased and decreased inflammation scores on MRI-SI over 6 and 16 weeks, respectively [14, 23]. Moreover, also patients with stable background treatment in the placebo group of the ABILITY-1 trial slightly decreased in SPARCC-score at group level, like we found in this study [3].
Although too few patients in the SPACE-cohort used DMARDs to draw conclusions on the effect of DMARDs, comparable effects can be expected. The comparator group in the ESTHER trial using sulfasalazine showed a mean decrease of 1.7 and 1.9 SPARCC-units over 24 and 48 weeks, respectively [13]. In the comparator group of another trial where patients used methotrexate, a mean of 1.4 (95%CI -0.8 to 3.5) inflammatory lesions resolved over

30 weeks [24]. Although an overall decrease in inflammation score was seen in these trials, some patients increased in inflammation score on MRI-SI when looking at the individual level [13, 24]. These results indicate that in patients on stable treatment changes in BME on MRI-SI that are beyond measurement error may occur, which may point to true fluctuation in inflammatory activity over time.

The direct comparisons of our results with the results of drug efficacy trials is difficult as the SPACE-cohort is an observational cohort including unselected patients with back pain of short duration resulting in a heterogeneous patient population, with low numbers of a 'positive-MRI' and low baseline mean SPARCC-scores, while drug efficacy trials select patients with high levels of disease activity. In patients selected because of a high level of disease activity a decrease in scores is more likely (regression to the mean) in comparison to an unselected group of patients. Thus, the patients in the SPACE-cohort will likely not be representative of patients in trials. Nevertheless, we have also observed an overall decrease in the SPACE-cohort, just as in the trials. This might be due to the fact that patients preferably seek help in case of maximum complaints, which is by default the time point of inclusion in the SPACE-cohort. It is possible that the results would have been different if this study had been performed in a long-standing or severely diseased group of patients. Furthermore, the SPACE-cohort is not designed to investigate the effects of treatment on inflammation on MRI. For example, and in contrast to drug efficacy trials, there is not a good relation between the start date of therapy and the date of the MRI.

Another possible limitation is that the readers have given their judgement based on the ASAS definition immediately after the evaluation according to the SPARCC-score. Since the quantitative part of the ASAS definition resembles a SPARCC-score of 2, the choice of the value of 2 as the best SPARCC-score to serve as cut-off level for negative and positive MRI may not be entirely independent. It would have been better if different scores were acquired independently, or even by different readers, as is frequently the case in clinical trials.

In conclusion, a SPARCC-score of 2 as cut-off value best reflects the caesura between a positive and negative MRI according to the ASAS definition. This cut-off can be used (in clinical trials) in order to create a dichotomous MRI variable of potential prognostic interest. The SDCs we have obtained in our two experiments are close enough to the proposed MIC of 2.5 SPARCC-units, which adds credibility to a cut-off level of 2.5 units in that it represents a true difference rather than only measurement error. Surprisingly, while patients are on stable treatment, true (>SDC) changes in SPARCC-score over time (both increases and decreases) were frequently observed. This observation strongly suggests that MRI-activity fluctuates over time.

References

1. Rudwaleit M, Jurik A, Hermann K, *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.
2. Rudwaleit M, van der Heijde D, Landewé 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.
3. Sieper J, van der Heijde D, Dougados M, *et al.* Efficacy and safety of adalimumab in patients with non-radiographic axial spondyloarthritis: results of a randomised placebo-controlled trial (ABILITY-1). Ann Rheum Dis 2013;72:815-22.
4. Sieper J, Lenaerts J, Wollenhaupt J, *et al.* Efficacy and safety of infliximab plus naproxen versus naproxen alone in patients with early, active axial spondyloarthritis: results from the double-blind, placebo-controlled INFAST study, Part 1. Ann Rheum Dis 2014;73:101-7.
5. Landewé R, Braun J, Deodhar A, *et al.* Efficacy of certolizumab pegol on signs and symptoms of axial spondyloarthritis including ankylosing spondylitis: 24-week results of a double-blind randomised placebo-controlled Phase 3 study. Ann Rheum Dis 2014;73:39-47.
6. Maksymowych W, Inman R, Salonen D, *et al.* Spondyloarthritis research Consortium of Canada magnetic resonance imaging index for assessment of sacroiliac joint inflammation in ankylosing spondylitis. Arthritis Rheum 2005;53:703-9.
7. Landewé R, Hermann K, van der Heijde D, *et al.* Scoring sacroiliac joints by magnetic resonance imaging. A multiple-reader reliability experiment. J Rheumatol 2005;32:2050-5.
8. Braun J, Baraliakos X, Hermann K, *et al.* Golimumab reduces spinal inflammation in ankylosing spondylitis: MRI results of the randomised, placebo- controlled GO-RAISE study. Ann Rheum Dis 2012;71:878-84.
9. Baraliakos X, Brandt J, Listing J, *et al.* Outcome of patients with active ankylosing spondylitis after two years of therapy with etanercept: clinical and magnetic resonance imaging data. Arthritis Rheum 2005;53:856-63.
10. Lambert R, Salonen D, Rahman P, *et al.* Adalimumab significantly reduces both spinal and sacroiliac joint inflammation in patients with ankylosing spondylitis: a multicenter, randomized, double-blind, placebo-controlled study. Arthritis Rheum 2007;56:4005-14.
11. Braun J, Landewé R, Hermann K, *et al.* Major reduction in spinal inflammation in patients with ankylosing spondylitis after treatment with infliximab: results of a multicenter, randomized, double-blind, placebo-controlled magnetic resonance imaging study. Arthritis Rheum 2006;54:1646-52.
12. van Onna M, Jurik A, van der Heijde D, *et al.* HLA-B27 and gender independently determine the likelihood of a positive MRI of the sacroiliac joints in patients with early inflammatory back pain: a 2-year MRI follow-up study. Ann Rheum Dis 2011;70:1981-5.
13. Song I, Hermann K, Haibel H, *et al.* Effects of etanercept versus sulfasalazine in early axial spondyloarthritis on active inflammatory lesions as detected by whole-body MRI (ESTHER): a 48-week randomised controlled trial. Ann Rheum Dis 2011;70:590-6.
14. Jarrett S, Sivera F, Cawkwell L, *et al.* MRI and clinical findings in patients with ankylosing spondylitis eligible for anti-tumour necrosis factor therapy after a short course of etoricoxib. Ann Rheum Dis 2009;68:1466-9.
15. Maksymowych W, Lambert R, Brown L, *et al.* Defining the Minimally Important Change for the SpondyloArthritis Research Consortium of Canada Spine and Sacroiliac Joint Magnetic Resonance Imaging Indices for Ankylosing Spondylitis. J Rheumatol 2012;39:1666-74.
16. van der Heijde D, Sieper J, Maksymowych W, *et al.* Spinal inflammation in the absence of sacroiliac joint inflammation on magnetic resonance imaging in patients with active nonradiographic axial spondyloarthritis. Arthritis Rheumatol 2014;66:667-73.
17. Weber U, Maksymowych W, Jurik A, *et al.* Validation of whole-body against conventional magnetic resonance imaging for scoring acute inflammatory lesions in the sacroiliac joints of patients with spondylarthritis. Arthritis Rheum 2009;61:893-9.

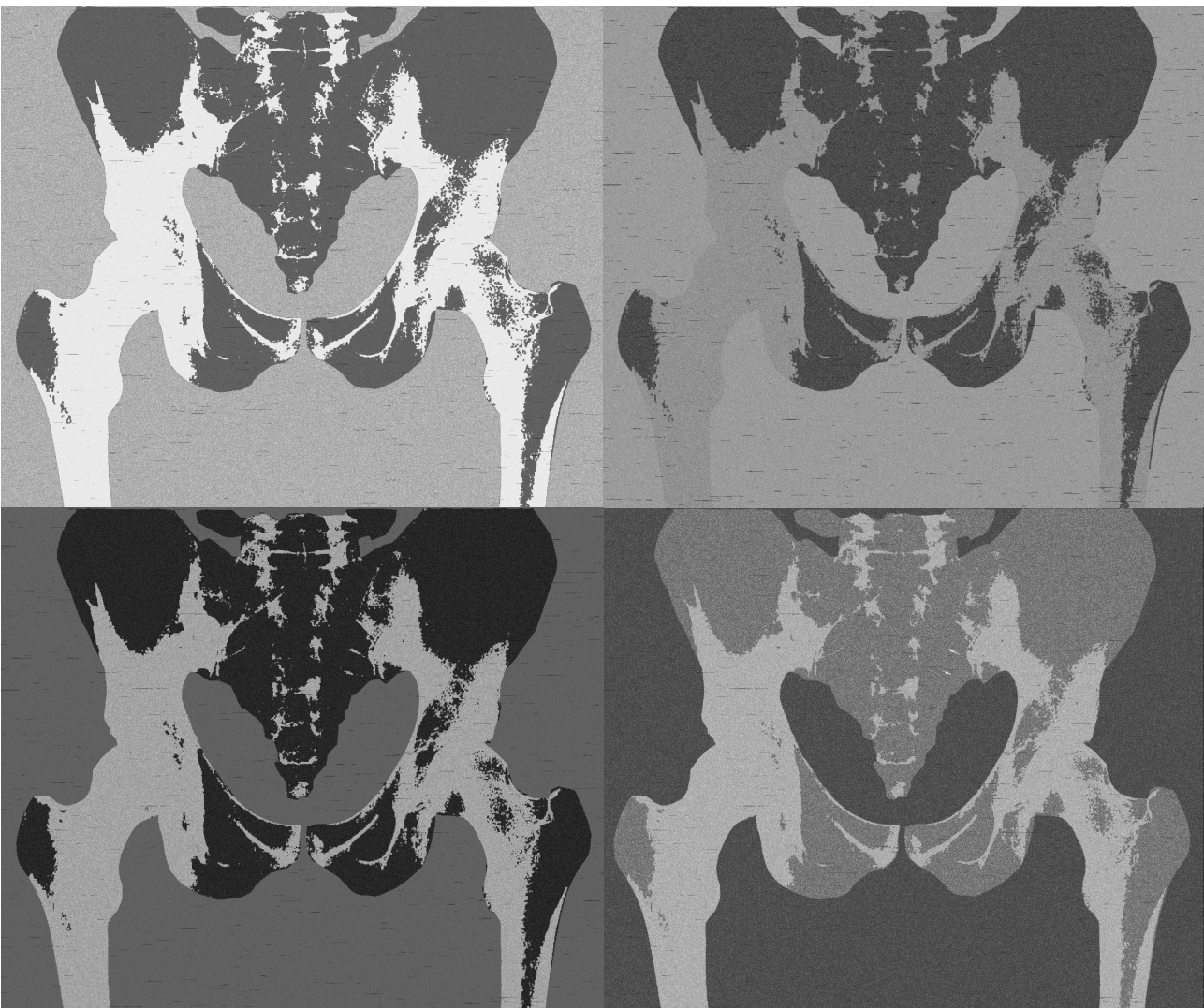
18. van den Berg R, de Hooge M, van Gaalen F, *et al.* Percentage of patients with spondyloarthritis in patients referred because of chronic back pain and performance of classification criteria: experience from the Spondyloarthritis Caught Early (SPACE) cohort. Rheumatology (Oxford) 2013;52:1492-9.
19. Rudwaleit M, van der Heijde D, Khan M, *et al.* How to diagnose axial spondyloarthritis early. Ann Rheum Dis 2004;63:535-43.
20. Dougados M, Simon P, Braun J, *et al.* ASAS recommendations for collecting, analysing and reporting NSAID intake in clinical trials/epidemiological studies in axial spondyloarthritis. Ann Rheum Dis 2011;70:249-51.
21. Cicchetti D, Feinstein A. High agreement but low kappa: II. Resolving the paradoxes. J Clin Epidemiol 1990;43:551-8.
22. Bruynesteyn K, Boers M, Kostense P, *et al.* Deciding on progression of joint damage in paired films of individual patients: smallest detectable difference or change. Ann Rheum Dis 2005;64:179-82.
23. Barkham N, Keen H, Coates L, *et al.* Clinical and imaging efficacy of infliximab in HLA-B27-Positive patients with magnetic resonance imaging-determined early sacroiliitis. Arthritis Rheum 2009;60:946-54.
24. Marzo-Ortega H, McGonagle D, Jarrett S, *et al.* Infliximab in combination with methotrexate in active ankylosing spondylitis: a clinical and imaging study. Ann Rheum Dis 2005;64:1568-75.
25. van der Linden S, Valkenburg H, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361-8.

8

Comparison of recommendations for the use of anti-tumour necrosis factor therapy in ankylosing spondylitis in 23 countries worldwide

Rheumatology (Oxford) 2011;50:2270-7

R. van den Berg
E. Stanislawska-Biernat
D. van der Heijde



Abstract

Objective

To give an overview of the recommendations for the use of anti-TNF-α therapy in AS in 23 countries worldwide

Methods

The recommendations were collected, translated and a summary was checked by Assessment of SpondyloArthritis International Society (ASAS) members from the respective countries. The recommendations were compared with the ASAS recommendations (2006) on three aspects: patient selection for initiation of treatment (diagnosis, disease activity, previous treatment and contraindications), assessment of disease and assessment of response.

Results

The majority of the recommendations are similar to the ASAS recommendation with regard to patient selection, assessment of disease and treatment response. Additional objective assessments of disease activity are required in eight countries, leading to a more strict indication to start anti-TNF-α therapy.

Conclusion

Most national recommendations follow the international ASAS recommendations, suggesting that the latter are widely implemented. This might contribute to comparable access with anti-TNF-α treatment across countries. This article shows that general consensus exists about the use of anti-TNF-α therapy in AS across the world, although some countries require additional objective signs of inflammation and/or more pre-treatment, which limits access.

Introduction

AS is a chronic, progressive inflammatory, rheumatic disease that generally starts in the second or third decade of life [1-3]. The most characteristic features of AS are inflammatory back pain (IBP) due to sacroiliitis and spondylitis, and the formation of syndesmophytes leading to ankylosis of the spine [1, 4]. In addition, AS is frequently associated with enthesitis, acute anterior uveitis, inflammatory bowel disease (IBD), psoriasis, peripheral (oligo) arthritis predominantly of the lower extremities, and cardiovascular and pulmonary abnormalities [1, 5, 6].

For decades, AS was mainly treated with NSAIDs, physiotherapy and to a lesser extent with DMARDs [3, 4]. And this is still the basis for treatment according to the Assessment of SpondyloArthritis International Society (ASAS)/European League Against Rheumatism (EULAR) recommendations for the management of AS [1]. Even though NSAIDs often give quick symptomatic relief [7], the effects on the longterm outcome are limited and there are reservations with safety in relation to long-term use [2, 7, 8]. Moreover, DMARDs are largely ineffective in axial AS and have limited efficacy on peripheral arthritis in AS [3, 7, 8].

The treatment armamentarium is broadened since the discovery of anti-TNF-α agents as an effective therapy. The anti-TNF-α agents infliximab [8, 9], etanercept [10, 11], adalimumab [12] and golimumab [13] have shown to be effective in the treatment of AS in short-term as well as intermediate to long-term evaluations [2, 14]. Anti-TNF-α agents are very effective in the treatment of AS; nevertheless, they are associated with high costs and risks of side effects and might not be suitable for all patients.

Therefore, it is important that recommendations are available to support the appropriate use of anti-TNF-α agents within individual countries.

In 2003, the ASAS proposed recommendations for the use of anti-TNF-α treatment in AS for rheumatologists and other experts in the management of AS, as well as payers [3, 14]. There was an update of the recommendations in 2006 [15]. Many countries developed national guidelines, whether or not based on the ASAS recommendations. The aim of the present report is to give an overview of the recommendations for the use of anti-TNF-α therapy in AS in 23 countries worldwide, with a focus on the similarities and differences compared with the ASAS recommendations.

In concordance with the advice of EULAR, we use the general term of recommendations throughout the manuscript, although some countries publish their recommendations as guidelines.

Methods

The recommendations of the following countries (presented alphabetically grouped by continent) were presented and translated: Australia, Hong Kong, Korea, Canada, Colombia, Mexico, Belgium, Czech Republic, Finland, France, Germany, Greece, Hungary, Italy, the Netherlands, Norway, Poland, Portugal, Slovakia, Spain, Sweden, Switzerland and the UK.

A summary of the translated recommendations was sent to ASAS members from the specific countries included in this overview. They were asked to check the correctness of the summary. The recommendations were compared with the 2006 version of the ASAS recommendations [15] as a standard to be able to easily compare discrepancies.

ASAS recommendations

The ASAS recommendations are divided into the following three parts: patient selection for initiation of treatment including diagnosis, disease activity, previous treatment and contraindications; assessment of disease; and assessment of response (table 1).

Table 1: International ASAS consensus statement for the use of anti-TNFα agents in patients with AS.

PATIENT SELECTION	
Diagnosis	Patients normally fulfilling modified New York criteria for definitive ankylosing spondylitis Modified New York criteria 1984: Radiological criterion: Sacroiliitis, grade > II bilaterally or grade III to IV unilaterally Clinical criteria (two of the following three): low back pain and stiffness for more than three months which improves with exercise but is not relieved by rest; limitation of motion of the lumbar spine in both the sagittal and frontal planes; limitation of chest expansion relative to normal values correlated for age and sex
Active disease	Active disease for >4 weeks BASDAI >4 (0-10) and an expert* opinion** *The expert is a physician, usually a rheumatologist, with expertise in inflammatory back pain and the use of biological agents. Expert should be locally defined. **The expert should consider clinical features (history and examination), serum acute phase reactant levels and/or imaging results, such as radiographs demonstrating rapid progression or MRI indicating ongoing inflammation.
Treatment failure	All patients should have had adequate therapeutic trials of at least two NSAIDs. An adequate therapeutic trial is defined as: Treatment for at least 3 months at maximum recommended or tolerated anti-inflammatory dose unless contraindicated Treatment for <3 months where treatment was withdrawn because of intolerance, toxicity, or contraindications Patients with pure axial manifestations do not have to take DMARDs before anti-TNFα treatment can be started Patients with symptomatic peripheral arthritis should have an insufficient response to at least one local corticosteroid injection if appropriate Patients with persistent peripheral arthritis must have had a therapeutic trial of sulfasalazine* Patients with symptomatic enthesitis must have failed appropriate local treatment *Sulfasalazine: treatment for at least four months at standard target dose or maximally tolerated dose unless contraindicated or not tolerated. Treatment for less than four months, where treatment was withdrawn because of intolerance or toxicity or contraindicated.
Contra-indications	Women who are pregnant or breast feeding; effective contraception must be practiced Active infection Patients at high risk of infection including: Chronic leg ulcer Previous tuberculosis (note: please follow local recommendations for prevention or treatment) Septic arthritis of a native joint within the past 12 months Sepsis of a prosthetic joint within the past 12 months, or indefinitely if the joint remains in situ Persistent or recurrent chest infections Indwelling urinary catheter History of lupus or multiple sclerosis Malignancy or pre-malignancy states excluding: Basal cell carcinoma Malignancies diagnosed and treated more than 10 years previously (where the probability of total cure is very high)

Table 1: Continued

ASSESSMENT OF DISEASE	
ASAS core set for daily practice	Physical function (BASFI or Dougados functional index) Pain (VAS, past week, spine at night, from ankylosing spondylitis and VAS, past week, spine, from ankylosing spondylitis) Spinal mobility (chest expansion and modified Schober and occiput to wall distance and lateral lumbar flexion) Patient's global assessment (VAS, past week) Stiffness (duration of morning stiffness, spine, past week) Peripheral joints and entheses (number of swollen joints (44 joints count), enthesitis score such as developed in Maastricht, Berlin, or San Francisco) Acute phase reactants (ESR or CRP) Fatigue (VAS)
BASDAI	VAS overall level of fatigue/tiredness, past week VAS overall level of ankylosing spondylitis neck, back, or hip pain, past week VAS overall level of pain/swelling in joints other than neck, back or hips, past week VAS overall discomfort from any areas tender to touch or pressure, past week VAS overall level of morning stiffness from time of awakening, past week Duration and intensity (VAS) of morning stiffness from time of awakening (up to 120 minutes)
ASSESSMENT OF RESPONSE	
Responder criteria	BASDAI: 50% relative change or absolute change of 20 mm (on a scale between 0 and 100) and expert opinion in favour of continuation
Time of evaluation	Between 6 and 12 weeks

ASAS, Assessment of Spondyloarthritis International Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; VAS, visual analogue scale; ESR, erythrocyte sedimentation rate; CRP, C-reactive protein; NSAID, non-steroidal anti-inflammatory drugs; DMARD, disease modifying anti-rheumatic drugs; TNF, tumour necrosis factor; MRI, magnetic imaging resonance.

Results

Table 2 gives an overview of the recommendations of the 23 countries (references of the recommendations in appendix 1, available as supplementary data at Rheumatology Online). They are presented alphabetically grouped by continent. The recommendations of Canada, Mexico, France, Italy, Portugal, Spain and Sweden (n=7) (table 2) were developed by the professional rheumatologic community as treatment recommendations. In Australia, Hong Kong, Korea, Colombia, Belgium, Finland, Greece, Norway, Poland and Switzerland (n=10) (table 2), the recommendations were developed for reimbursement purposes. The recommendations of the Czech Republic, Germany, Hungary, the Netherlands, the UK and Slovakia (n=6) (table 2) were developed for both purposes.

Diagnosis

According to the ASAS recommendations, patients should normally fulfill the modified New York criteria for AS (table 1) [15]. Most recommendations (n=16) follow the ASAS recommendations and qualify patients for treatment if they fulfill the modified New York criteria [16]. In five recommendations, MRI and/or CT, instead of X-rays, are approved to reveal sacroiliitis [16]. In Hong Kong and Colombia, a diagnosis of SpA according to the Amor or ESSG criteria is sufficient for the diagnostic part for initiation of anti-TNF-α therapy (table 2).

Disease activity

The ASAS recommendations define active AS as having active disease for >4 weeks based on a BASDAI score ≥4 (scale 0-10) and an expert opinion of active AS (table 1) [15]. According to all recommendations, except the Finnish recommendation, disease activity should be measured with the BASDAI. In 19 recommendations, the disease activity is qualified as high when the BASDAI is ≥4. In two other recommendations (Hong Kong and Norway), the BASDAI is also used to measure disease activity, but no qualification of active disease is given.
An expert opinion to determine disease activity is required in 13 countries (table 2).
In eight recommendations, additional assessments of disease activity are required, such as laboratory parameters for inflammation (CRP and/or ESR), (spinal) pain [visual analogue scale (VAS)] (n=4), patient and physician global health (n=2 and n=1, respectively), and/or inflammation on MRI (n=1), or limitation in spinal mobility (n=1) (table 2). In particular, the request for additional elevated acute-phase reactants or inflammation on MRI increases the threshold to start a TNF-blocker substantially. In one instance (Hong Kong), a large increase is required (ESR >50 mm/h or CRP >50 mg/l). Moreover, the requirement for limitation in spinal mobility is remarkable, as this can be caused by the severity of the disease without active inflammation.

Failure of standard treatment

ASAS offers a description of conventional treatment failure specified for the predominant localization of the disease (axial, peripheral arthritis and enthesitis) (table 1).
Most recommendations follow the ASAS recommendations and give specified descriptions of treatment failure.
In general, the recommendations describe failure of conventional treatment for predominantly axial localization as failure of two or more NSAIDs administered for a period of 1-3 months (n=18). In Hong Kong, Canada and France, patients should fail at least three NSAIDs. Conventional treatment failure for a predominantly peripheral localization is in 18 recommendations described as a failure of one or two DMARDs (in most recommendations specified as MTX and/or SSZ) administered for a period of 2-3 months, and as a failure of IA injections of CSs (n=16). Conventional treatment failure of CS injections for enthesitis is described in 12 recommendations (table 2).

Table 2: The recommendations of 23 countries.

Country	Diagnosis	Disease activity	Treatment failure	Assessment of disease	Assessment of response
Asia Pacific Region					
Australia[1, R1]	Sacroiliitis (X-ray) grade II bi- or grade III unilateral	BASDAI ≥4 & abnorm/al lab tests (ESR >25, CRP >10)	In the preceding 3 months: 2 different NSAIDs & a specified exercise program (both stretching & daily aerobic exercise)	At least 2 out of 3: LBP & stiffness ≥ 3 months relieved by exercise but not by rest; ↓ lumbar flexion (sagittal & frontal planes as a score of at least 1 on relevant measures in BASMI); ↓ chest expansion	↓ BASDAI ≥ 2 points & normalized lab tests (or 20% ↑ on baseline lab tests) 12 weeks
Hong Kong[1, R2]	Modified NY criteria or ESSG	'Persistent active disease' (BASDAI & patient & physician GH) & ESR ≥50 mm/hr & CRP ≥50mg/L	3 NSAIDs (different chemical classes), ≥4 weeks each & ≥2 DMARDs (SSZ/ MTX/Arava) ≥ 3 months (peripheral joint)	According to ASAS	According to ASAS, or 50% or 2 points (VAS) ↓ in patient & physician GH or ↓ < 30% TJC or SJC 16 weeks
Korea[1, R3]	According to ASAS	BASDAI ≥ 4	2 DMARDS or NSAIDS, 3 months	According to ASAS	According to ASAS 12 weeks
Americas					
Canada[2, R4]	Expert opinion & 'unequivocal evidence of sacroiliitis or spinal inflammation' on X-ray/CT/ MRI	2 out of the 3 following: BASDAI ≥4 & ↑ CRP and/or ESR & inflammatory lesions SI joint and/or spine on MRI	≥3 NSAIDs 2 weeks & coriticosteroid injections may be considered & SSZ ≥3 months in peripheral arthritis & MTX ≥3 months in peripheral arthritis	Disease manifestations & level of symptoms, clinical findings/prognostic indicators & disease activity/ inflammation& pain & function, disability, handicap & structural damage, hip involvement, spinal deformities & general clinical status & patients' wishes / expectations	According to ASAS 16 weeks

Table 2: Continued

Colombia[1, R5]	According to ASAS, or AMOR or ESSG	According to ASAS, duration not specified	According to ASAS & >2 infiltration intra steroids in peripheral arthritis & >2 injections corticosteroids in enthesitis	According to ASAS & not measuring fatigue & stiffness & extra chest radiography	↓ BASDAI ≥50% 3 months
Mexico[2, R6]	According to ASAS	According to ASAS	According to ASAS	According to ASAS & assessment of safety	According to ASAS
Europe					
Belgium[1, R7]	Modified NY criteria & expert opinion	BASDAI >4 & elevated CRP	'Insufficient response' on at least ≥2 NSAIDs, optimum dosage ≥3 months or contraindication for NSAIDs	Not mentioned	↓ BASDAI ≥50% or 2 points ENT & ADA <14 weeks IFX <12 weeks
Czech Rep.[1, 2, R8]	According to ASAS, or MRI instead of X-ray	BASDAI ≥4 & CRP ≥10 at 2 consecutive FU visits separated by ≥4 weeks	According to ASAS	According to ASAS	According to ASAS 12 weeks
Finland[1, R9]	Not mentioned	'Active disease' not specified	2 NSAIDs & MTX & SSZ 6-12 months & intolerability/ lack of efficacy DMARDs	Not mentioned	Expert opinion
France[2, R10]	According to ASAS, or MRI/ CT instead of X-ray Or cervical syndesmophytes without any sacroiliac structure lesion	According to ASAS & peripheral: TJC & SJC (≥3 of 76-78 joints)	According to ASAS & ≥3 NSAIDs instead of 2 NSAIDs & enthesitis not specified	According to ASAS	Axial: ↓ BASDAI >2 points Peripheral: >30% decrease TJC & SJC FU varies with drug & route of administration
Germany[1, 2, R11]	'Secured diagnosis of AS'	According to ASAS & disease symptomatic ≥6 months	According to ASAS & enthesitis not specified	'Clinical rheumatologic findings using validated scores'	'If there is no response (not specified) after 3 months, no continuation treatment'
Greece[1, R12]	Clinical & laboratory & radiological findings	According to ASAS	According to ASAS & MTX ≥2 months in peripheral arthritis & ≥2 topical infusions of corticosteroids in enthesitis	According to ASAS	According to ASAS 12-16 weeks

Table 2: Continued

Hungary[1, 2, R13]	According to ASAS	According to ASAS	According to ASAS & ≥2 intra-articular steroid injections, SSZ or other DMARD ≥4 months in peripheral arthritis	BASDAI	According to ASAS 14 weeks
Italy[2, R14]	According to ASAS	According to ASAS	According to ASAS	According to ASAS	According to ASAS
Netherlands[1, 2, R15]	According to ASAS	According to ASAS	According to ASAS	According to ASAS	According to ASAS
Norway[1, R16]	'Conventional diagnosis' Prescription from department with >2 specialists (rheumatologists). Approval based on disease history, previous treatment & current status.	ESR/CRP, clinical status and imaging. Peripheral arthritis: joint counts & disease activity indices. Axial: BASDAI & BASFI. Both axial & peripheral: pain, fatique & globals on VAS. (not specified)	Axial: 2 NSAIDs & peripheral arthritis: DMARD (pref. SSZ) If relevant: intra-articular steroidal injections	According to ASAS - spinal mobility - stiffness + DAS28 + BASFI	Treatment stop after 3-6 months, if estimated treatment response has not been achieved.
Poland[2, R17]	According to ASAS	2 of 3 following parameters: BASDAI ≥4, pain VAS ≥4, CRP >10mg/dl in 12 weeks interval on stable treatment. One or more parameters limited in 1 month interval: chest expansion, occiput-to-wall distance, Schober test	According to ASAS & >2 intra-articular glucocorticosteroids injections in peripheral arthritis	BASFI & BASMI & BASDAI, & spinal pain (VAS) & CRP/ ESR	↓ BASDAI <4 No time frame
Portugal[2, R18]	According to ASAS, or MRI/ CT instead of X-ray	According to ASAS & in case of BASDAI <4: expert opinion	According to ASAS & 4 weeks instead of 3 months	Physical function (BASFI) & pain & patient's global assessment & stiffness & BASDAI	↓ BASDAI ≥50% or ≥ASAS 20% improvement 12 weeks
Slovakia[1, 2, R19]	According to ASAS, or MRI instead of X-ray	BASDAI ≥4 & CRP ≥10 at 2 consecutive FU visits separated by ≥4 weeks	According to ASAS	According to ASAS	According to ASAS 12 weeks

Table 2: Continued

Spain[2, R20]	Expert opinion	BASDAI >4 & 1 out of 3: spinal pain or patient global assessment (VAS >4) or ↑ ESR/CRP, for >3 months	According to ASAS & any DMARD (pref. SSZ) instead of only SSZ	Pain & patient & physician global assessment & physical function & structural damage	↓ BASDAI >50% & 50% or 2 points (VAS) ↓ in patient's general assessment, and/or ESR/CRP 3-4 months
Sweden[2, R21]	According to ASAS	According to ASAS	According to ASAS	According to ASAS	According to ASAS
Switzerland[1, R22]	Expert opinion	Expert opinion	Expert opinion & mandatory consent of the consultant physician of the health insurance company	Not mentioned	Expert opinion, IFX after 6 weeks ADA after 12 weeks ETN unlimited
UK[1, 2, R23]	According to ASAS	BASDAI >4 & spinal pain (VAS >4) on two occasions >4 weeks apart and no change in treatment	>2 NSAIDs for 4 weeks	Not mentioned	According to ASAS & ↓ spinal pain (VAS ≥2cm) 6-12 weeks

R1-R23, references of the recommendations (online appendix 1); 1, reimbursement recommendation; 2, professional recommendation; ASAS, Assessment of Spondyloarthritis International Society; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; VAS, visual analogue scale; SJC, swollen joint count; TJC, tender joint count; ESR, erythrocyte sedimentation rate; CRP, C-reactive protein; MTX, methotrexate; SSZ, sulfasalazine; IFX, infliximab; ADA, adalimumab; ETN, etanercept; NSAID, non-steroidal anti-inflammatory drugs; DMARD, disease modifying anti-rheumatic drugs; MRI, magnetic resonance imaging; CT, computed tomography; FU, follow-up; GH, general health; LBP, low back pain.

Contraindications

To minimize treatment risks, ASAS has specified a list of contraindications (table 1) basically similar to contraindications of the treatment of anti-TNF-α therapy for other indications [15]. Almost all recommendations (n=17) list active infections, especially tuberculosis (TB), as contraindications.

Several recommendations mention some types of malignancy or pre-malignancy (n=10), a history of lupus (n=8), multiple sclerosis or other demyelinating diseases (n=11) and pregnancy/breastfeeding (n=9) as contraindications, in accordance with the ASAS recommendations.

A frequently mentioned contraindication (n=11) not referred to in the ASAS recommendations [15] is heart failure stages 3-4 as defined by the New York Heart Association (NYHA) [17].

Remarkably, the recommendations of the Czech Republic and Slovakia report that an advanced or terminal radiographic stage of the disease is a contraindication for applying anti-TNF-α therapy. Four recommendations do not mention contraindications at all.

Monitoring and withdrawal

ASAS recommends using the ASAS core set for daily practice [18] and the BASDAI to assess the activity of the disease (table 1) [15]. Most countries (n=19) recommend the ASAS core set for daily practice as well, or at least a part of the ASAS core set. However, four countries do not specify how to assess the disease (table 2).

An assessment of the treatment response should be conducted 6-12 weeks after the start of the treatment, according to ASAS (table 1) [15]. In 16 recommendations, the same time frame is advised. However, in seven recommendations the response is assessed after >12 weeks (range 14-16 weeks).

At this assessment point, a decision should be made about either continuation or discontinuation of anti-TNF-α therapy. ASAS advises considering discontinuation in patients not showing a 50% relative or absolute change of 2cm (scale 0-10 cm) in the BASDAI score [15]. Eighteen recommendations use these criteria to determine a good treatment response. In some recommendations other criteria to assess response to treatment are obligatory, such as normalized or improved lab tests (n=3) and improvement in pain (n=2) or BASDAI <4 (n=1). Furthermore, ASAS advises a positive opinion by the expert to continue treatment. This criterion is used in 14 recommendations as well.

Discussion and Conclusion

This report provides an overview of the recommendations developed in 23 countries across the world. ASAS developed recommendations for the management of anti-TNF-α therapy in patients with AS [3, 15]. As internationally developed recommendations, the ASAS recommendations might contribute to comparable access with anti-TNF-α treatment across countries [19].

Indeed, this aim is (largely) reached, since the recommendations in AS are quite similar worldwide, in contrast to the recommendations in RA, which vary greatly between countries in Europe [19]. This can be explained by the lack of European guidance for initiation of anti-TNF-α therapy in RA [19], unlike the situation in AS [15]. Another explanation might be the considerably varying goals of RA treatment with anti-TNF-α agents [19]. Other possible explanations for the differences in recommendations across countries that apply to both RA and AS are variations regarding different methods for funding health-care provision and the level of recognition of recommendations [19].

Despite the similarities between the recommendations in AS across countries, differences exist. These differences are mostly based on the fact that some countries use objective

assessment, such as acute-phase reactants, to measure disease activity for initiation and to monitor treatment response. This puts a major limitation on access to TNF-a blockers for patients in these countries, as only about half of the patients with active disease have elevated acute-phase reactants [20]. Although patients with elevated acute-phase reactants have a higher likelihood to show response, this difference is too small to withhold patients with a normal acute-phase reactant treatment with TNF-a blockers. Other differences exist in the required pre-treatment for NSAIDs (more and/or longer) and DMARDs (also required in axial disease and not only SSZ in peripheral disease). Moreover, several countries evaluate the efficacy of treatment after ≥12 weeks.

In conclusion, it can be said that despite some differences, there is general consensus about the recommendations to use anti-TNF-a therapy in AS across the world, except for the stricter requirement of objective signs of inflammation in some countries. The observation that most national recommendations follow the international ASAS recommendations seems to indicate that the latter are widely accepted and implemented. The information acquired by this comparison will also be taken into account in the next update of the ASAS recommendations.

Supplementary data

Supplementary data are available at Rheumatology Online.

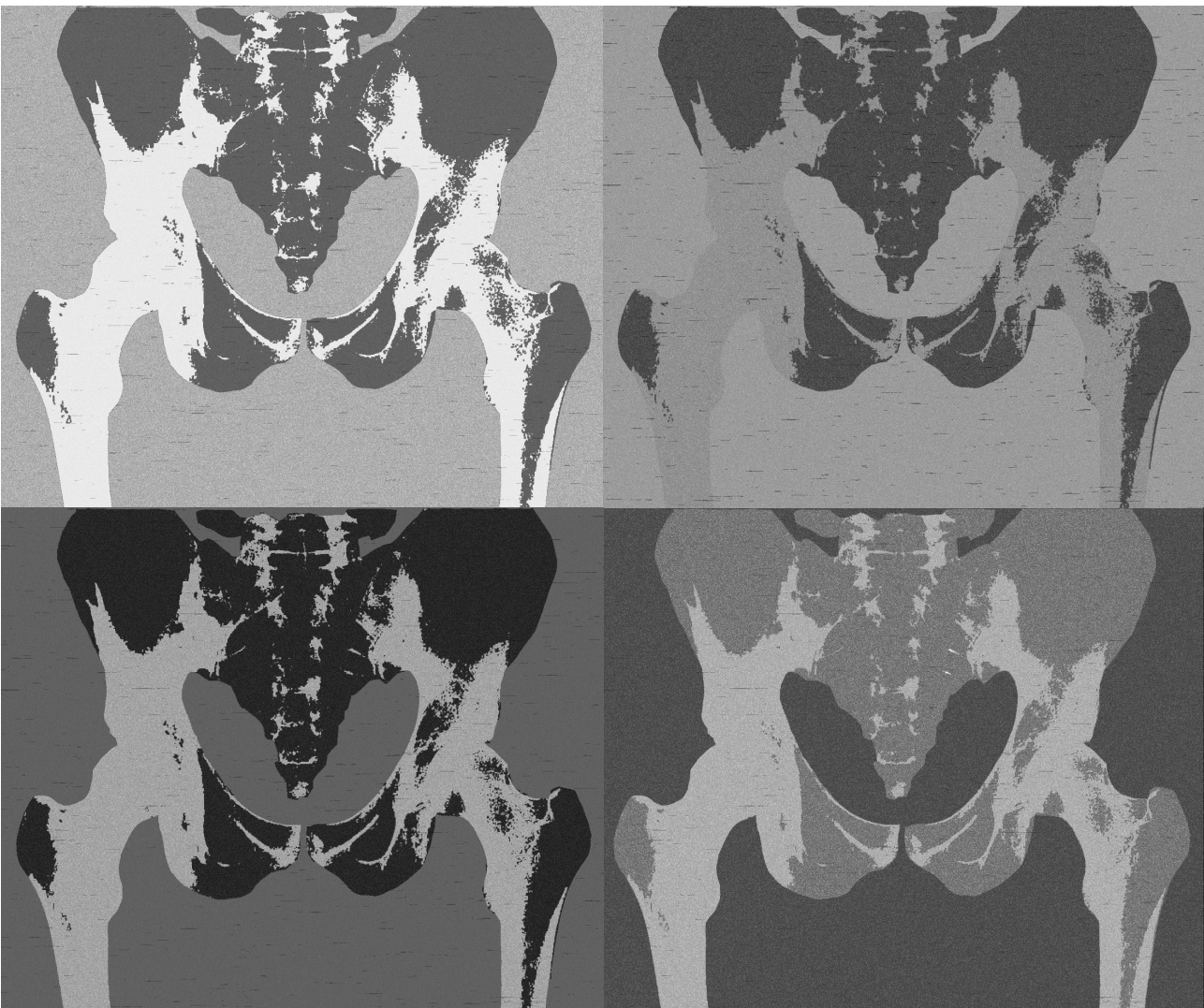
References

1. Zochling J, van der Heijde D, Burgos-Vargas R, *et al.* ASAS/EULAR recommendations for the management of ankylosing spondylitis. Ann Rheum Dis 2006;65:442-52.
2. van der Heijde D, Landewé R, Baraliakos X, *et al.* Radiographic findings following two years of infliximab therapy in patients with ankylosing spondylitis. Arthritis Rheum 2008;58:3063-70.
3. Braun J, Pham T, Sieper J, *et al.* International ASAS consensus statement for the use of anti-tumour necrosis factor agents in patients with ankylosing spondylitis. Ann Rheum Dis 2003;62:817-24.
4. Baraliakos X, Listing J, Rudwaleit M, *et al.* Radiographic progression in patients with ankylosing spondylitis after 2 years of treatment with the tumour necrosis factor alpha antibody infliximab. Ann Rheum Dis 2005;64:1462-6.
5. Khan M. Update on spondyloarthropathies. Ann Intern Med 2002;136:896-907.
6. Braun J, Bollow M, Remlinger G, *et al.* Prevalence of spondylarthropathies in HLA-B27 positive and negative blood donors. Arthritis Rheum 1998;41:58-67.
7. Henderson C, Davis J. Drug insight: anti-tumor-necrosis-factor therapy for ankylosing spondylitis. Nat Clin Pract Rheumatol 2006;2:211-8.
8. Braun J, Brandt J, Listing J, *et al.* Treatment of active ankylosing spondylitis with infliximab: a randomised controlled multicentre trial. Lancet 2002;359:1187-93.
9. van der Heijde D, Dijkmans B, Geusens P, *et al.* Efficacy and safety of infliximab in patients with ankylosing spondylitis: results of a randomized, placebo-controlled trial (ASSERT). Arthritis Rheum 2005;52:582-91.
10. Davis J, van der Heijde D, Braun J, *et al.* Recombinant human tumor necrosis factor receptor (etanercept) for treating ankylosing spondylitis: a randomized, controlled trial. Arthritis Rheum 2003;48:3230-6.
11. Gorman J, Sack K, Davis J. Treatment of ankylosing spondylitis by inhibition of tumor necrosis factor alpha. N Engl J Med 2002;346:1349-56.
12. van der Heijde D, Kivitz A, Schiff M, *et al.* Efficacy and safety of adalimumab in patients with ankylosing spondylitis: results of a multicenter, randomized, double-blind, placebo-controlled trial. Arthritis Rheum 2006;54:2136-46.
13. Inman R, Davis J, van der Heijde D, *et al.* Efficacy and safety of golimumab in patients with ankylosing spondylitis results of a randomized, double-blind, placebo-controlled, phase III trial. Arthritis Rheum 2008;58:3402-12.
14. Baraliakos X, Listing J, Fritz C, *et al.* Persistent clinical efficacy and safety of anti-TNF alpha therapy with infliximab in patients with ankylosing spondylitis over 8 years. Ann Rheum Dis 2009;68(Suppl3):634.
15. Braun J, Davis J, Dougados M, *et al.* First update of the international ASAS consensus statement for the use of anti-TNF agents in patients with ankylosing spondylitis. Ann Rheum Dis 2006;65:316-20.
16. van der Linden S, Valkenburg H, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361-8.
17. The Criteria Committee of the New York Heart Association. Nomenclature and criteria for diagnosis of diseases of the heart and great vessels, 9th edn. Boston, MA: Little, Brown & Co, 1994.
18. Zochling J, Braun J. Assessments in ankylosing spondylitis. Best Pract Res Clin Rheumatol 2007;21:699-712.
19. Emery P, van Vollenhoven R, Ostergaard M, *et al.* Guidelines for initiation of anti-tumour necrosis factor therapy in rheumatoid arthritis: similarities and differences across Europe. Ann Rheum Dis 2009;68:456-9.
20. Spoorenberg A, van der Heijde D, de Klerk E, *et al.* Relative value of erythrocyte sedimentation rate and C-reactive protein in assessment of disease activity in ankylosing spondylitis. J Rheumatology 1999;26:980-4.

First update of the current evidence for the management of ankylosing spondylitis with non-pharmacological treatment and non-biological drugs: a systematic literature review for the ASAS/EULAR management recommendations in ankylosing spondylitis

Rheumatology (Oxford) 2012;51:1388-96

R. van den Berg
X. Baraliakos
J. Braun
D. van der Heijde



Abstract

Objective

To perform a systematic literature review as a basis for the update of the Assessment in SpondyloArthritis International Society and European League Against Reumatism (ASAS/EULAR) recommendations for the management of AS with non-pharmacological interventions and non-biologic drugs.

Methods

The search was performed in PubMed, EMBASE, PEDro and Cochrane between 1 January 2005 and 1 December 2009, and in abstracts of EULAR and ACR meetings (2007-09). Effect sizes for outcomes on pain, disease activity, spinal mobility and physical function and level of evidence were presented.

Results

Of 2383 papers, 35 with complete data were included. Physical therapy exercises in various modalities have positive effects on BASFI, BASDAI, pain and mobility function. Various NSAIDs including coxibs improve BASDAI, disease activity and BASFI. No effect of SSZ and MTX on any variable was found. Surgical interventions of the spine and the hip can give excellent results by restoring function.

Conclusion

This concise summary of current evidence for non-pharmacological interventions and non-biologic drugs formed the basis for the update of the ASAS/EULAR recommendations for the management of AS.

Introduction

AS is a chronic, inflammatory rheumatic disease, generally starting early in life [1-4]. Inflammatory back pain due to sacroiliitis and spondylitis, and formation of syndesmophytes leading to ankylosis of the spine, characterize AS [4, 5]. Although AS is difficult to treat, the treatment armamentarium of AS has been broadened since the discovery of anti-TNF-α agents as effective treatments [6-8].

Clinicians need to be aware of the relative benefits and risks of the available treatments, and need to have evidence-based information about the most efficacious strategies in particular patient settings [4].

In 2005 Zochling *et al.* [4] performed a systematic literature search for evidence based recommendations by the Assessment in SpondyloArthritis International Society and European League Against Reumatism (ASAS/EULAR) for the management of AS. In 2010 an update of this systematic literature search was performed to serve as a base for the development of an update of the ASAS/EULAR recommendations [9]. The details and results of the performed systematic review on non-pharmacological interventions and non-biologic drugs are presented in this article. The results on biologics are presented in the article by Baraliakos elsewhere in this journal.

Methods

Participants and outcome measures

Participants were defined as patients with a diagnosis of AS or axial spondyloarthritis. The required treatments were non-pharmacological interventions and non-biologic drugs. There were no restrictions with regard to type of non-pharmacological intervention, or to dose, duration or route of administration of non-biologic drugs.

The primary outcomes of interest include pain, disease activity (including BASDAI), spinal mobility (including BASMI) and physical function (including BASFI).

Inclusion/exclusion criteria

Randomized controlled trials (RCTs) and controlled trials are the ideal study designs for inclusion in this review.

However, the aim of this review is to provide evidence of all types of non-pharmacological interventions and non-biologic drugs. Since not all types of treatment can be studied within RCTs alone, the main focus of interest was also on systematic reviews, uncontrolled trials/cohort studies, case-control studies and cross-sectional studies. Studies about non-axial spondyloarthritides and other inflammatory joint conditions, animal studies, non-clinical outcome studies and non-treatment studies, narrative review articles, commentaries, guidelines, case reports, letters and editorials and studies in other languages than English, Dutch and German were excluded.

Studies about biologic drugs were also excluded because those studies will be included in a search performed by Baraliakos.

Systematic literature search

A search strategy was built in collaboration with an experienced librarian, based on the previous search of Zochling *et al*. [4]. The systematic literature search for published papers was performed in the electronic databases PubMed, EMBASE, PEDro and Cochrane between 1 January 2005, which is the end date of the literature search by Zochling *et al*. [4], and 1 December 2009. The complete search strategies for the databases are provided in supplementary appendix S1, available as supplementary data at Rheumatology Online.

Abstracts of rheumatology scientific meetings (EULAR, ACR) from the years 2007, 2008 and 2009 were searched by hand to ensure that all potential studies were identified for this review. Furthermore, references of relevant reviews and included papers were hand searched for information on any other relevant studies.

Selection of studies

One reviewer (RvdB) assessed each title and abstract on suitability for inclusion in the review, according to the inclusion and exclusion criteria described above. Papers not addressing the topic of interest were excluded and reasons recorded. The full-text paper was assessed when further information was required to determine if the inclusion criteria were met.

Data extraction and categorizing evidence

The included papers were assessed using the full-text paper by one reviewer (RvdB) to extract relevant data, including patient characteristics and details of treatment.
If necessary, authors were contacted to provide any required additional information. The results were reported to the ASAS/EULAR expert committee at the beginning of the recommendation development process. All included papers were categorized according to their level of evidence (see legend in table 1) [3]. The assigned levels are shown in table 1.

Data analysis

Since different types of studies are included about various types of treatments, the results are very heterogeneous and therefore the results cannot be pooled. Yet, the results are analysed and presented per type of treatment.

Estimation of effectiveness

Per treatment group, the Cohen's effect size (Cohen's ES; mean change in score divided by the baseline SD) was calculated, and the standardized response mean (SRM; mean change divided by the SD of the change) was calculated where possible [36]. To compare the effect between treatment groups, treatment ES was calculated (mean change in the index group minus the mean change in the comparator group divided by a pooled baseline SD). For each ES, the corresponding 95% CI was constructed. An ES of 0.2 or 0.3 is considered a small change, around 0.5 as moderate and >0.8 as a large change, and a negative ES indicates worse.

Results

Treatment modalities and types of research evidence

The general search revealed 3179 papers; 1638 in PubMed, 1486 in EMBASE, 14 in PEdro, 34 in Cochrane and 7 abstracts. After eliminating duplicates, 2383 papers remained. Of those, 2347 papers were excluded (supplementary appendix S2, available as supplementary data at Rheumatology Online) and 35 papers were included, of which 3 are Cochrane reviews and 1 abstract (supplementary appendix S3, available as supplementary data at Rheumatology Online). An overview of the included papers is shown in the supplementary appendix S4, available as supplementary data at Rheumatology Online.

Non-pharmacological treatment

No studies on treatments about diet, education, self-help groups or lifestyle modification were present within this search.

Table 1: Cohen's effect size (Cohen's ES) with 95% confidence interval (95% CI) for various outcome measures and different management modalities.

Intervention	Assessment point	No. of patients	Level of evidence	Cohen's ES BASFI (95% CI)	Cohen's ES BASDAI (95% CI)	Cohen's ES Pain (95% CI)	Cohen's ES Disease act. (95% CI)	Cohen's ES BASMI (95% CI)
Exercise therapy								
Group exercise [15]	6 weeks	22	2b	0.27 (-0.33, 0.86)	-	0.50 (-0.10, 1.10)	-	0.26 (-0.34, 0.86)
Home exercise		16		0 (-0.69, 0.69)	-	0.20 (-0.50, 0.89)	-	0.05 (-0.64, 0.69)
Exercise group [13]	6 weeks	22	3	0.10 (-0.49, 0.69)	-	-	-	-
Exercise Group [14]	8 weeks	16	1b	0.41 (-0.29, 1.11)	0.38 (-0.31, 1.08)	-	-	
Control group		16		0.22 (-0.47, 0.92)	0.24 (-0.46, 0.93)	-	-	0.10 (-0.60, 0.79)
Hospital exercise [17]	12 weeks	23	1b	***0.94 (0.34, 1.55)***	***1.00 (0.39, 1.61)***	***0.76 (0.16, 1.36)***	-	-
Home exercise		23		0.22 (-0.36, 0.80)	***0.88 (0.28, 1.49)***	0.11 (-0.47, 0.69)	-	-
Hospital exercise	6 months	23		***0.78 (0.18, 1.38)***	***0.84 (0.24, 1.45)***	0.48 (-0.10, 1.07)	-	-
Home exercise		23		0.48 (-0.11, 1.06)	***1.12 (0.50, 1.74)***	0.48 (-0.10, 1.07)	-	-
Home exercise [18]	12 weeks	25	2b	***0.63 (0.07, 1.20)***	***0.94 (0.35, 1.52)***	-	-	-
Control group		18		0.18 (-0.48, 0.83)	0.19 (-0.48, 0.84)	-	-	-

Table 1: continued

Exercise GPR method [19]		20		0.57 (-0.06, 1.21)	***1.07 (0.41, 1.73)***	***1.28 (0.60, 1.96)***	-	-
Conventional exercise	12 weeks	21	2b	***0.66 (0.04, 1.28)***	***1.20 (0.54, 1.86)***	***1.03 (0.38, 1.67)***	-	-
Control group		15		0.13 (-0.59, 0.85)	0.32 (-0.40, 1.04)	0.19 (-0.53, 0.91)	-	-
Balneotherapy								
Exercise + Stangerbath [20]	3 weeks	29	1b	1.05 (-0.08, 2.17)	***1.71 (0.33, 3.09)***	-	-	0.30 (-0.44, 1.05)
Exercise		28		0.20 (-0.48, 0.88)	0.36 (-0.43, 1.15)	-	-	0.06 (-0.52, 0.63)
Balneotherapy [22]		20		-	-	0.81 (-0.26, 1.88)	-	-
Balneotherapy + NSAIDs	3 weeks	21		-	-	0.83 (-0.26, 1.92)	-	-
NSAID		20		-	-	0.58 (-0.38, 1.54)	-	-
Balneotherapy		20	1b	-	-	0.97 (-0.17, 2.11)	-	-
Balneotherapy + NSAIDs	6 months	21		-	-	1.10 (-0.10, 2.10)	-	-
NSAID		20		-	-	0.96 (-0.18, 2.09)	-	-
Rehabilitation [24]	3 weeks	52	3	***2.23 (1.74, 2.72)***	-	***5.10 (4.31, 5.89)***	-	-
	6 weeks			***1.70 (1.25, 2.14)***	-	***3.62 (3.00, 4.24)***	-	-
	12 weeks			***1.16 (0.75, 1.58)***	-	***2.06 (1.58, 2.54)***	-	-

Table 1: continued

NSAIDs								
Celecoxib (200 mg) [27]	12 weeks	153	1b	***0.35 (0.12, 0.57)***	***0.60 (0.37, 0.83)***	***1.96 (1.69, 2.23)***	***1.11 (0.87, 1.35)***	***0.52 (0.29, 0.75)***
Celecoxib (400 mg)		150		***0.41 (0.18, 0.64)***	***0.80 (0.56, 1.03)***	***1.79 (1.52, 2.06)***	***1.29 (1.05, 1.54)***	0.13 (-0.10, 0.36)
Diclofenac (150 mg)		155		***0.35 (0.12, 0.57)***	***0.83 (0.60, 1.06)***	***1.84 (1.57, 2.10)***	***1.28 (1.03, 1.52)***	0.18 (-0.04, 0.40)
Etoricoxib (90 mg) [25]	6 weeks	22	3	***1.08 (0.45, 1.71)***	***1.08 (0.45, 1.71)***	***1.44 (0.77, 2.10)***	***0.93 (0.31, 1.56)***	-
DMARDs								
Sulfasalazine (2g) [32]	24 weeks	120	1b	0.19 (-0.06, 0.44)	***1.11 (0.84, 1.38)***	-	-	-
Sulfasalazine (3g) [33]	16 weeks	187	1b	***0.29 (0.08, 0.49)***	***1.24 (1.02, 1.46)***	***0.95 (0.74, 1.17)***	-	0.08 (-0.13, 0.28)
Etanercept (50 mg)		379		***0.48 (0.33, 0.62)***	***1.96 (1.79, 2.14)***	***1.68 (1.51, 1.84)***	-	***0.34 (0.20, 0.48)***
Sulfasalazine (2g) [34]	52 weeks	16	3	-	0.43 (-0.28, 1.13)	-	-	-
MTX (15mg up to 20mg) [31]	16 weeks	20	3	0.01 (-0.61, 0.63)	0.00 (-0.62, 0.62)	0.00 (-0.62, 0.62)	-	0.23 (-0.12, 1.14)
Other								
Radium-224 [36]	6 months	278	3	***0.68 (0.39, 0.97)***	***0.86 (0.57, 1.15)***	***1.14 (0.85, 1.44)***	***1.22 (0.92, 1.52)***	-

Table 1: continued

Surgery	No. of patients	Results
Hip resurfacing [37]	23 (38 hips)	Both groups ROM ↑: resurfacing group significantly better than THR group
Total hip replacement (THR)	25 (41 hips)	
Posterior correction and fixation without anterior fusion (posterior opening-wedge osteotomy) [39]	30	Neurologic deficit ↑ good correction of kyphosis
Open wedge osteotomy (OWO) [38]	51	Both techniques have good clinical outcome, patient quality of life ↑, high patient satisfaction: no difference between the two techniques
Closed wedge osteotomy (CWO)	66	
Cervicothoracic extension osteotomy [40]	26	Neck pain and swallowing problems ↑ Both groups restored horizontal gaze
Cervical extension osteotomy: Conventional technique [41]	114	High patient satisfaction, function ↑, psycho-social body image ↑: no difference in outcome between the two groups
Cervical extension osteotomy: Current technique	17	
Smith-Peterson osteotomy [42]	12	Pain and neurologic deficits ↑, high patient satisfaction
Lumbar closing wedge osteotomy [43]	11	Horizontal gaze restored in all patients
Cervical osteotomy: sitting position [45]	11	Good correction of kyphosis, no loss of correction
Cervical osteotomy: prone position	5	
Cervical decancellation closing wedge osteotomy [46]	8	Horizontal gaze ↑, good subjective outcome
Closing wedge osteotomy [44]	21	Quality of life ↑, good functional outcome

Significant ES in ***bold italics***. Category of evidence: Ia = meta-analysis of randomized controlled trials (RCT); Ib = RCT; IIa = controlled study without randomization; IIb = quasi-experimental study; III = non-experimental descriptive studies (comparative, correlation and case-control studies); IV = expert committee reports or opinion or clinical experience of respected authorities, or both.

Exercise therapy

The effect of physiotherapy has been reviewed in a Cochrane review in 2008 [37]. The results of this review show that individual home-based or supervised exercise programmes are better than no intervention at all on pain, physical function, spinal mobility and patient global assessment, and that supervised group physiotherapy is better than home exercise [37]. Besides the Cochrane review, nine papers were identified [10-15, 38-40] of which three were already included in the Cochrane review [38-40]. In the six additional papers, the effects of various exercises in AS patients are compared (supplementary appendix S4, available as supplementary data at Rheumatology Online). The results of these six studies confirm the results of the Cochrane review. Various types of exercise [supervised group, home and Global Posture Reeducation (GPR) method exercise] have moderate to good effects on BASFI, BASDAI, pain and mobility, as shown by the calculated Cohen's ES and SRM (table 1 and figure 1). The calculated treatment ES showed that supervised group physiotherapy is better than home exercise on BASFI, pain and mobility, and slightly better on BASDAI. Home exercise is better than no exercise at all on BASFI and BASDAI (table 2).

Although most papers had level 1b evidence, the studies investigated various exercises with variable durations and had small patient samples. Therefore, many ES are not statistically significant, showing only a trend (table 1).

Balneotherapy, spa therapy and rehabilitation

The same Cochrane review also revealed that combined inpatient spa exercise therapy followed by group physiotherapy is better than group physiotherapy alone [37]. In addition to the Cochrane review, four RCTs [16, 17, 42, 43] about various types of balneotherapy and spa therapy in AS patients were identified (level 1b evidence), of which two were already presented in the Cochrane review [42, 43].

As in the exercise therapy studies, the studies about balneotherapy included only small patient numbers in various therapies, resulting in not statistically significant ES.

However, the trend shows that balneotherapy in all its modalities is (moderate) effective on BASFI, BASDAI and pain, as shown by the calculated Cohen's ES and treatment ES (tables 1 and 2). The effect of balneotherapy on pain is equal to the effect of NSAIDs (either mono or combined) [17]. Stangerbath therapy combined with exercises is effective on BASFI and BASDAI, only directly after therapy [16] (figure 1). One level 3 evidence study about the effect of inpatient rehabilitation was identified that showed a strong effect on BASFI, pain and OWD (table 1) [18].

NSAIDs

Three studies about the effects of different NSAIDs in AS patients were identified [19, 20, 41]. The effect of celecoxib (200 and 400mg daily) in comparison with diclofenac (150mg daily) (level 1b evidence) [19], the effect of etoricoxib (90mg daily) (level 3 evidence) [20] and the effect of NSAIDs in continuous usage in comparison with NSAID usage on demand (level 1b evidence) [41] were investigated.

The latter study is a follow-up study of a double-blind RCT about the effect of celecoxib 200mg versus ketoprofen 200mg versus placebo after 6 weeks [44]. This study was already included in the review of Zochling *et al*. [4], showing a significant improvement in pain and function after 6 weeks of use of both NSAIDs in comparison with placebo [44]. The follow-up study showed that measures of disease activity, including pain and BASDAI, were stable over a time period of 24 months in both the continuous and on-demand groups and not statistically significant between the groups [41]. Although the clinical effects of both treatment strategies are similar, inhibition of structural damage progression in the spine is better with continuous use than with on-demand use [41].

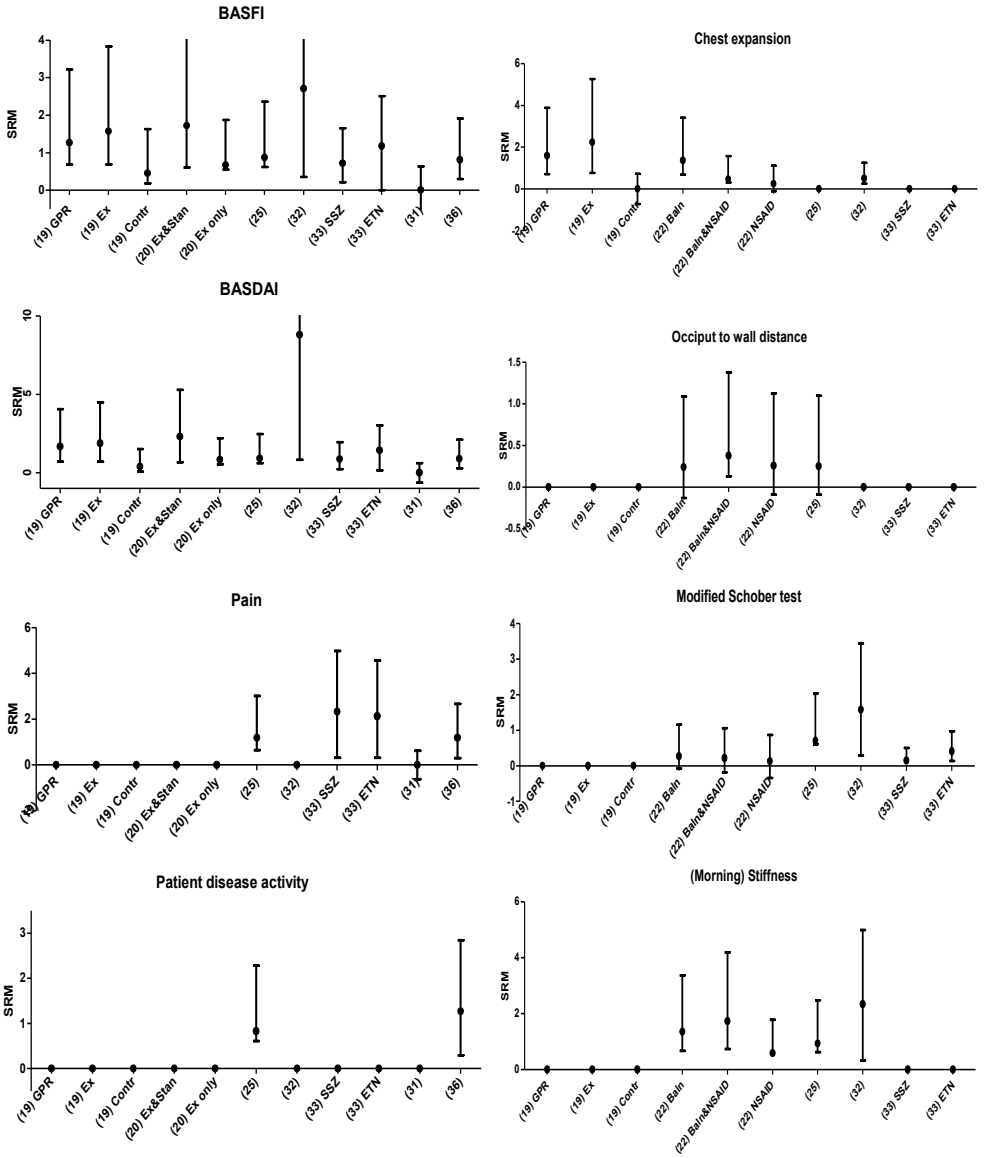


Figure 1: SRM of different outcome parameters.

The calculated Cohen's ES and SRM of the other two studies showed that all NSAIDs have statistically significantly moderate to good effect on BASFI, BASDAI, disease activity and pain (table 1 and figure 1). Various NSAIDs have a similar effect as assessed by treatment ES (table 2). Furthermore, no new signs of toxicity were discovered.

DMARDs

In 2005 and 2006, two Cochrane reviews summarized the effects of MTX and SSZ, respectively [45, 46]. The MTX review showed that there is no evidence to support any benefit of MTX in the treatment of AS. One additional open-label study about the effect of MTX was found besides the Cochrane review. The calculated Cohen's ES did not show any improvement on BASFI, BASDAI, pain or mobility (table 1 and figure 1) [24].

The SSZ Cochrane review showed some benefit of SSZ in reducing ESR and easing morning stiffness, yet no benefit in physical function, pain, spinal mobility and disease activity [46]. These results are confirmed by three additional identified SSZ studies (level 1b and 3 evidences) not included in the SSZ Cochrane review [21-23] (supplementary appendix S4, available as supplementary data at Rheumatology Online). Although the calculated Cohen's ES revealed moderate to good effect on BASDAI and pain (table 1, figure 1), the calculated treatment ES showed that the effect of SSZ on these outcome parameters is not better than the effect of placebo, as shown in one study, and that the effect of SSZ was statistically significantly worse than the effect of etanercept (ETN) (table 2). No new signs of toxicity for SSZ and MTX were found.

Other therapies

Two studies about other types of therapy were identified. One study investigated the effect of probiotics compared with placebo (level 1b evidence) [47] (supplementary appendix S4, available as supplementary data at Rheumatology Online). The calculated ES showed that probiotics do no better on global well-being and functional index than placebo (tables 1 and 2). The other study investigated the effect of radium chloride on BASFI, BASDAI, pain and disease activity in an uncontrolled design (level 3 evidence) (table 1). The calculated ES demonstrated a moderate effect on BASFI and good effect on BASDAI, pain and disease activity (table 1 and figure 1) [25].

Surgical interventions

Total hip replacement

Total hip replacement (THR) is a frequently used procedure in AS patients with hip involvement. This search revealed one study about hip surgery and resurfacing of the hip (Zimmer, Wintherthur, Switzerland). The authors proposed that hip resurfacing might be an option instead of THR for young AS patients with hip involvement. They compared the effects of resurfacing with THR on pain relief, function and mobility in 38 resurfaced hips (23 AS patients) and 41 THRs (25 AS patients) over a mean follow-up time of 34.5 months. Both groups showed significant pain relief and good restoration of function and mobility [26] (table 1).

Spine

Although spinal surgery to resolve fixed kyphotic deformity is accompanied by severe risks, it can give excellent functional results by restoring balance and horizontal vision, as shown by all nine included papers in this search [27-35]. These papers review the different available techniques. All included papers are case series, and therefore low-quality studies (level of evidence 3) (supplementary Appendices S4 and S5, available as supplementary data at Rheumatology Online).

One study compared open wedge osteotomy (OWO) of the cervical spine with closed wedge osteotomy (CWO).

No difference in correction of kyphosis between the two techniques was found [28].

Table 2: Treatment ES with 95% confidence interval (95% CI) for the different interventions. The interventions are grouped together. The level of evidence, duration and point of measurement are described in this table.

Intervention	Comparator	Assessment point	Treatment ES BASFI (95% CI)	Treatment ES BASDAI (95% CI)	Treatment ES Pain (95% CI)	Treatment ES Disease act. (95% CI)	Treatment ES BASMI (95% CI)
Exercise therapy							
Group exercise [15]	Home exercise	6 weeks	0.30 (-0.35, 0.95)	-0.06 (-0.71, 0.58)	0.31 (-0.34, 0.96)	-	0.20 (-0.44, 0.85)
Exercise Group [14]	Control group	8 weeks	0.17 (-0.52, 0.87)	0.00 (-0.69, 0.69)	-	-	0.53 (-0.17, 1.24)
Hospital exercise [17]	Home exercise	12 weeks	***0.59 (0.01, 1.16)***	0.22 (-0.36, 0.80)	***0.68 (0.10, 1.26)***	-	-
Hospital exercise	Home exercise	6 months	0.15 (-0.43, 0.72)	-0.17 (-0.74, 0.41)	0.04 (-0.54, 0.61)	-	-
Home exercise [18]	Control group	12 weeks	0.49 (-0.12, 1.12)	***0.84 (0.21, 1.47)***	-	-	-
Exercise (GPR method) [19]	Conventional exercise	12 weeks	0.01 (-0.60, 0.62)	-0.04 (-0.65, 0.57)	0.17 (-0.44, 0.79)	-	-
Conventional exercise	Control group		0.44 (-0.23, 1.11)	***0.82 (0.13, 1.51)***	***0.93 (0.24, 1.63)***	-	-
Exercise (GPR method)	Control group		0.41 (-0.26, 1.09)	***0.75 (0.06, 1.44)***	***1.15 (0.43, 1.87)***	-	-
Balneotherapy							
Exercise + stangerbath [20]	Exercise	3 weeks	***0.65 (0.12, 1.18)***	***1.49 (0.91, 2.08)***	-	-	0.21 (-0.31, 0.73)

Table 2: continued

Balneotherapy [22]	Balneotherapy + NSAIDs		-	-	0.08 (-0.53, 0.70)	-	-
Balneotherapy + NSAIDs	NSAIDs	3 weeks	-	-	0.36 (-0.26, 0.98)	-	-
Balneotherapy	NSAIDs		-	-	0.43 (-0.18, 1.04)	-	-
Balneotherapy	Balneotherapy + NSAIDs		-	-	0.00 (-0.61, 0.61)	-	-
Balneotherapy + NSAIDs	NSAIDs	6 months	-	-	0.31 (-0.31, 0.92)	-	-
Balneotherapy	NSAIDs		-	-	0.28 (-0.32, 0.89)	-	-
NSAIDs							
Continuous NSAID use (200-400 mg) [26]	On-demand NSAID use (200-400 mg)	104 weeks	-0.15 (-0.42, 0.12)	-	-0.053 (-0.32, 0.21)	0.1 (-0.17, 0.37)	-
Celecoxib (200 mg) [27]	Diclofenac (150 mg)	12 weeks	0 (-0.22, 0.22)	-0.30 (-0.52, 0.07)	-0.08 (-0.31, 0.14)	-0.17 (-0.39, 0.06)	-
Celecoxib (400 mg)	Diclofenac (150 mg)		0.04 (-0.18, 0.27)	-0.10 (-0.32, 0.13)	-0.07 (-0.29, 0.16)	-0.06 (-0.28, 0.17)	-
DMARDs							
Sulfasalazine (2 g) [32]	Placebo	24 weeks	0.08 (-0.17, 0.33)	0.15 (-0.10, 0.40)	0.05 (-0.20, 0.31)	-	-
Sulfasalazine (3 g) [33]	Etanercept (50 mg)	16 weeks	***-0.19 (-0.37, -0.01)***	***-0.76 (-0.94, -0.58)***	***-0.70 (-0.88, -0.52)***	-	***-0.23 (-0.41, -0.06)***

The interventions are grouped together. The level of confidence, duration and point of measurement are described in this table. Significant ES in ***bold italics***.

Another study compared the conventional technique of cervical extension osteotomy with a new technique in which the patients have a modified larger lateral resection area than with the conventional technique. Again, no differences between the two techniques were found concerning functional improvement, satisfaction or complications [30]. Similarly, a prone or a sitting position during the procedure demonstrated no difference in correction [33].
For thoracolumbar deformities, polysegmental wedge osteotomy might be associated with lower risks. However, the correction is often insufficient in the case of calcified intervertebral discs. Theoretically, CWO is superior to OWO in terms of efficiency and minimal loss of correction and lower accompanied risks, although technically difficult [28, 34].
For pseudoarthrosis, posterior correction is an effective treatment (posterior opening wedge osteotomy), as well as fixation without anterior fusion [27]. The data from the included papers do not show whether a specific technique gives better results for any specific indication.

Discussion

This systematic review is an update of the review by Zochling *et al.* [4] and identified available nonpharmacological and non-biologic pharmacological treatments effective for symptomatic control of AS. The results of this search confirm the 2005 findings for physiotherapy [4]; exercises in various modalities, individually at home or in a group and under supervision, land or water based, have positive effects on BASFI, BASDAI, pain and mobility function. However, the small numbers of participants, the heterogeneity of the interventions and outcome measures, and deficiency in reporting data result in wide intervals and lack of strong evidence.
Zochling *et al.* [4] revealed that different kinds of NSAIDs and coxibs improve spinal and peripheral joint pain and function. The current search confirmed these results by showing that various NSAIDs including coxibs improve BASDAI, disease activity and BASFI.
In 2005 no effect of SSZ or MTX on back pain and function was demonstrated [4], which is confirmed by new research. The current search revealed no effect of SSZ and MTX on pain, nor on BASFI and BASDAI.
THR is still the standard procedure in AS patients with hip involvement. Although a small study showed positive effects of hip resurfacing techniques [26], it must be carefully considered whether resurfacing techniques are indeed a good alternative for THR given the recent developments and accompanying problems with the resurfacing techniques from another brand. The articular surface replacement hip prosthesis from the manufacturer DePuy (Warsaw, IN, USA) has been recalled from the market because of failing of the prosthesis. Metal debris from wear of the implant led to a reaction that destroyed the soft tissues surrounding the joint, causing long-term disability and a high revision rate of 12% over 5 years [48].
Surgical interventions of the spine give excellent results by restoring horizontal gaze and function yet are considered with high risks. Furthermore, it is still unclear which procedure of spine surgery is the best for any specific indication.

Conclusion

This review presents a concise summary of the current evidence available for therapeutic interventions for the management of AS, both non-pharmacological and pharmacological, excluding biologics. This overview formed the basis for the update of the ASAS/EULAR recommendations for the management of AS.

Supplementary data

Supplementary data are available at Rheumatology Online.

References

1. Braun J, Pham T, Sieper J, *et al.* International ASAS consensus statement for the use of anti-tumour necrosis factor agents in patients with ankylosing spondylitis. Ann Rheum Dis 2003;62:817-24.
2. van der Heijde D, Landewé R, Baraliakos X, *et al.* Radiographic findings following two years of infliximab therapy in patients with ankylosing spondylitis. Arthritis Rheum 2008;58:3063-70.
3. Zochling J, van der Heijde D, Burgos-Vargas R, *et al.* ASAS/EULAR recommendations for the management of ankylosing spondylitis. Ann Rheum Dis 2006;65:442-52.
4. Zochling J, van der Heijde D, Dougados M, *et al.* Current evidence for the management of ankylosing spondylitis: a systematic literature review for the ASAS/EULAR management recommendations in ankylosing spondylitis. Ann Rheum Dis 2006;65:423-32.
5. Baraliakos X, Listing J, Rudwaleit M, *et al.* Radiographic progression in patients with ankylosing spondylitis after 2 years of treatment with the tumour necrosis factor alpha antibody infliximab. Ann Rheum Dis 2005;64:1462-6.
6. Braun J, Brandt J, Listing J, *et al.* Treatment of active ankylosing spondylitis with infliximab: a randomised controlled multicentre trial. Lancet 2002;359:1187-93.
7. Davis J, van der Heijde D, Braun J, *et al.* Recombinant human tumor necrosis factor receptor, (etanercept) for treating ankylosing spondylitis—a randomized, controlled trial. Arthritis Rheum 2003;48:3230-6.
8. van der Heijde D, Kivitz A, Schiff M, *et al.* Efficacy and safety of adalimumab in patients with ankylosing spondylitis: results of a multicenter, randomized, double-blind, placebo-controlled trial. Arthritis Rheum 2006;54:2136-46.
9. Braun J, van den Berg R, Baraliakos X, *et al.* update of the ASAS/EULAR recommendations for the management of ankylosing spondylitis. Ann Rheum Dis 2011;70:896-904.
10. Karapolat H, Akkoc Y, Sari I, *et al.* Comparison of group-based exercise versus home-based exercise in patients with ankylosing spondylitis: effects on Bath Ankylosing Spondylitis Indices, quality of life and depression. Clin Rheumatol 2008;27:695-700.
11. Ortancil O, Sarikaya S, Sapmaz P, *et al.* The effect(s) of a six-week home-based exercise program on the respiratory muscle and functional status in ankylosing spondylitis. J Clin Rheumatol 2009;15:68-70.
12. Widberg K, Hossein K, Hafstrom I. Self- and manual mobilization improves spine mobility in men with ankylosing spondylitis—a randomized study. Clin Rehabil 2009;23: 599-608.
13. Cagliyan A, Kotevoglu N, Onal T, *et al.* Does group exercise program add anything more to patients with ankylosing spondylitis? J Back Musculoskelet Rehabil 2007;20:79-85.
14. Durmus D, Alayli G, Cil E, *et al.* Effects of a home-based exercise program on quality of life, fatigue, and depression in patients with ankylosing spondylitis. Rheumatol Int 2009;29:673-7.
15. Durmus D, Alayli G, Uzun O, *et al.* Effects of two exercise interventions on pulmonary functions in the patients with ankylosing spondylitis. Joint Bone Spine 2009;76:150-5.
16. Gurcay E, Yuzer S, Eksioglu E, *et al.* Stanger bath therapy for ankylosing spondylitis: illusion or reality? Clin Rheumatol 2008;27:913-7.
17. Yurtkuran M, Ay A, Karakoc Y. Improvement of the clinical outcome in ankylosing spondylitis by balneotherapy. Joint Bone Spine 2005;72:303-8.
18. Lubrano E, D'Angelo S, Parsons W, *et al.* Effectiveness of rehabilitation in active ankylosing spondylitis assessed by the ASAS response criteria. Rheumatology 2007;46: 1672-5.
19. Sieper J, Klopsch T, Richter M, *et al.* Comparison of two different dosages of celecoxib with diclofenac for the treatment of active ankylosing spondylitis: results of a 12-week randomised, double-blind, controlled study. Ann Rheum Dis 2008;67:323-9.
20. Jarrett SJ, Sivera F, Cawkwell L, *et al.* MRI and clinical findings in patients with ankylosing spondylitis eligible for anti-tumour necrosis factor therapy after a short course of etoricoxib. Ann Rheum Dis 2009;68:1466-9.
21. Braun J, Zochling J, Baraliakos X, *et al.* Efficacy of sulfasalazine in patients with inflammatory back pain due to undifferentiated spondyloarthritis and early ankylosing spondylitis: a multicentre randomised controlled trial. Ann Rheum Dis 2006;65:1147-53.

22. Braun J, van der Horst-Bruinsma I, Huang F, *et al.* Clinical efficacy and safety of etanercept versus sulfasalazine in patients with ankylosing spondylitis: a randomized, double-blind trial. Arthritis Rheum 2011;63:1543-51.
23. Genc H, Duyur CB, Nacir B, *et al.* The effects of sulfasalazine treatment on enthesal abnormalities of inflammatory rheumatic diseases. Clin Rheumatol 2007; 26:1104-10.
24. Haibel H, Brandt H, Song I, *et al.* No efficacy of subcutaneous methotrexate in active ankylosing spondylitis: a 16-week open-label trial. Ann Rheum Dis 2007;66:419-21.
25. Alberding A, Stierle H, Brandt J, *et al.* Effectiveness and safety of radium chloride in the treatment of ankylosing spondylitis. Results of an observational study. Z Rheumatol 2006;65:245-51.
26. Li J, Xu W, Xu L, *et al.* Hip resurfacing arthroplasty for ankylosing spondylitis. J Arthroplasty 2009;24:1285-91.
27. Chang K, Tu M, Huang H, *et al.* Posterior correction and fixation without anterior fusion for pseudoarthrosis with kyphotic deformity in ankylosing spondylitis. Spine 2006;31:E408-13.
28. Chang K, Chen Y, Lin C, *et al.* Closing wedge osteotomy versus opening wedge osteotomy in ankylosing spondylitis with thoracolumbar kyphotic deformity. Spine 2005;30:1584-93.
29. Belanger T, Milam R, Roh J, *et al.* Cervicothoracic extension osteotomy for chin-on-chest deformity in ankylosing spondylitis. J Bone Joint Surg Am 2005;87:1732-8.
30. Simmons E, DiStefano R, Zheng Y, *et al.* Thirty-six years experience of cervical extension osteotomy in ankylosing spondylitis: techniques and outcomes. Spine 2006;31:3006-12.
31. Kim K, Lee S, Suk K, *et al.* Spinal pseudarthrosis in advanced ankylosing spondylitis with sagittal plane deformity: clinical characteristics and outcome analysis. Spine 2007;32:1641-7.
32. Min K, Hahn F, Leonardi M. Lumbar spinal osteotomy for kyphosis in ankylosing spondylitis: the significance of the whole body kyphosis angle. J Spinal Disord Tech 2007;20:149-53.
33. Langeloo D, Journee H, Pavlov PW\, *et al.* Cervical osteotomy in ankylosing spondylitis: evaluation of new developments. Eur Spine J 2006;15:493-500.
34. Tokala D, Lam K, Freeman B, *et al.* C7 decancellisation closing wedge osteotomy for the correction of fixed cervico-thoracic kyphosis. Eur Spine J 2007;16:1471-8.
35. Brox J, Helle A, Sorensen R, *et al.* Functional outcome after lumbar closing wedge osteotomy in ankylosing spondylitis. Int Orthop 2009;33:1049-53.
36. Wu J, Jiang G, Wei W. Confidence intervals of effect size in randomized comparative parallel-group studies. Stat Med 2006;25:639-51.
37. Dagfinrud H, Kvien T, Hagen K. Physiotherapy interventions for ankylosing spondylitis. Cochrane Database Syst Rev 2008; CD002822.
38. Fernandez-de-Las-Penas C, Alonso-Blanco C, Morales-Cabezas M, *et al.* Two exercise interventions for the management of patients with ankylosing spondylitis: a randomized controlled trial. Am J Phys Med Rehabil 2005;84:407-19.
39. Fernandez-de-Las-Penas C, Alonso-Blanco C, Alguacil-Diego IM, *et al.* One-year follow-up of two exercise interventions for the management of patients with ankylosing spondylitis: a randomized controlled trial. Am J Phys Med Rehabil 2006; 85:559-67.
40. Ince G, Sarpel T, Durgun B, *et al.* Effects of a multimodal exercise program for people with ankylosing spondylitis. Phys Ther 2006;86:924-35.
41. Wanders A, van der Heijde D, Landewé R, *et al.* Nonsteroidal antiinflammatory drugs reduce radiographic progression in patients with ankylosing spondylitis: a randomized clinical trial. Arthritis Rheum 2005;52:1756-65.
42. Codish S, Dobrovinsky S, Abu S, *et al.* Spa therapy for ankylosing spondylltis at the Dead Sea. Isr Med Assoc J 2005;7:443-6.
43. Altan L, Bingol U, Aslan M, *et al.* The effect of balneotherapy on patients with ankylosing spondylitis. Scand J Rheumatol 2006;35:283-9.
44. Dougados M, Behier J, Jolchine I, *et al.* Efficacy of celecoxib, acyclooxygenase 2-specific inhibitor, in the treatment of ankylosing spondylitis—a six-week controlled study with comparison against placebo and against a conventional nonsteroidal antiinflammatory drug. Arthritis Rheum 2001;44:180-5.
45. Chen J, Liu C, Lin J. Methotrexate for ankylosing spondylitis. Cochrane Database Syst Rev 2006;CD004524.

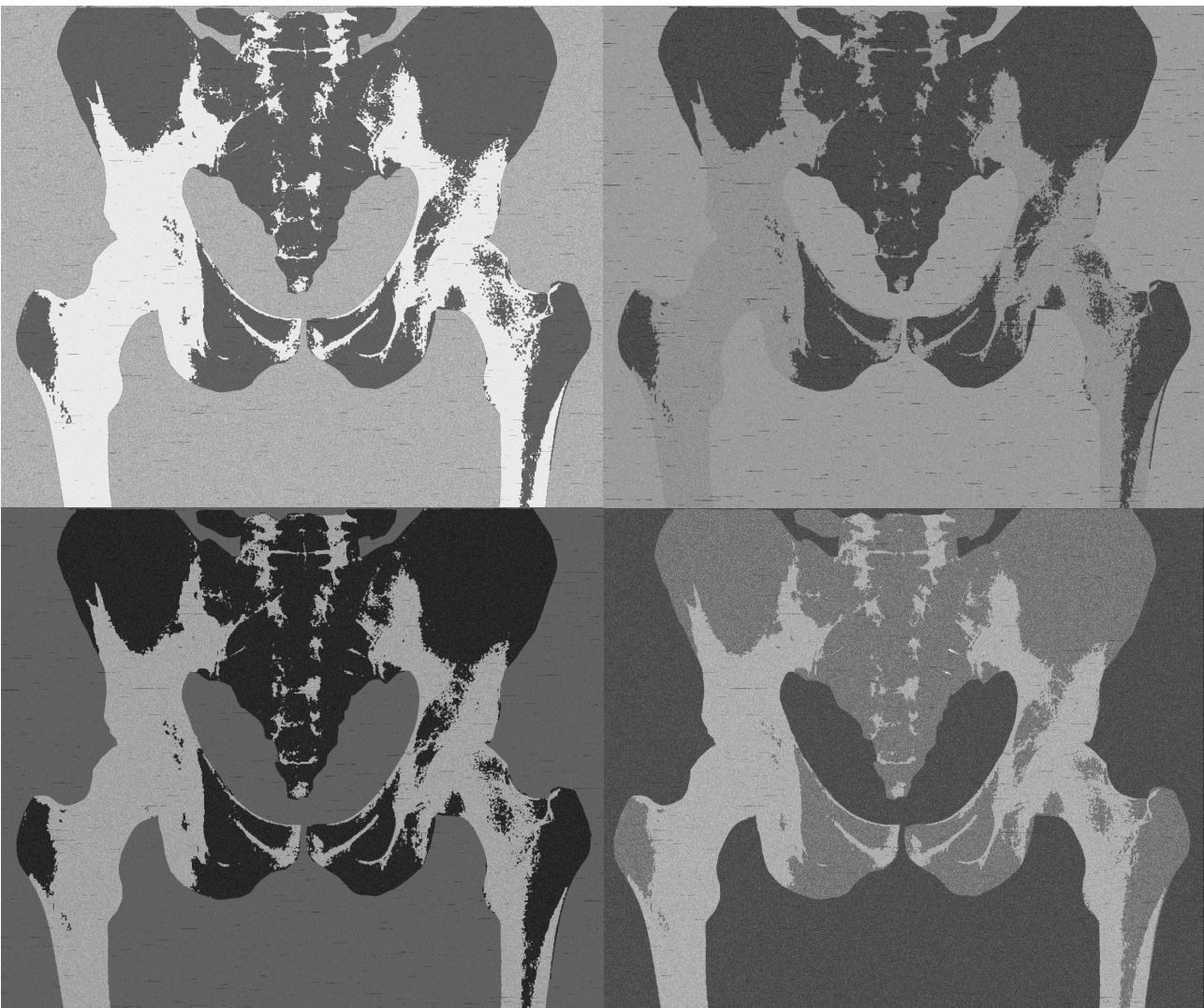
46. Chen J, Liu C. Sulfasalazine for ankylosing spondylitis. Cochrane Database Syst Rev 2005;CD004800.
47. Brophy S, Burrows CL, Brooks C, *et al.* Internet-based randomised controlled trials for the evaluation of complementary and alternative medicines: probiotics in spondyloarthropathy. BMC Musculoskelet Disord 2008;9:4.
48. Cohen D. Out of joint: the story of the ASR. BMJ 2011;342: d2905.

10

Update of the literature review on the treatment with biologics as basis for the first update of the ASAS/EULAR management recommendations of ankylosing spondylitis

Rheumatology (Oxford) 2012;51:1378-87

X. Baraliakos
R. van den Berg
J. Braun
D. van der Heijde



Abstract

Objective

To perform a literature review as basis for the update of the Assessment in SpondyloArthritis international Society/European League Against Rheumatism (ASAS/EULAR) treatment recommendations with biologics in AS.

Methods

A literature search of all publications found in MedLine, Embase and Cochrane database between 2005 and 2009 and in the EULAR/ACR meetings between 2007 and 2009 was performed. The research evidence and strength of recommendation (SOR) for biologics were provided.

Results

Out of 247 reports on AS treatment with biologics, 98 contained efficacy data and 25 had complete data for analysis. The treatment effect sizes (95% CI) for anti-TNF versus placebo varied between 0.34 (0.08, 0.6) and 1.5 (0.45, 2.5) for BASDAI and 0.33 (0.07, 0.59) and 2.5 (1.3, 3.7) for BASFI. The calculation of the numbers needed to treat all the different outcomes varied between 2.3 and 3.0 patients for all ASAS outcomes and between 2.7 and 6.5 patients for ASAS partial remission. Data on biologics other than anti-TNF and for TNF blockers on juvenile SpA were limited. The incidence rates of uveitis during anti-TNF treatment varied between 4.4/100 patient-years (pys) and 15.6/100 pys during placebo ($p<0.05$). The incidence rates of IBD flares were significantly less during infliximab treatment (0.2/100 pys). The rate of infections was higher in patients treated with anti-TNF as compared with placebo, but there was no difference in the incidence of serious infections for treatment with anti-TNF versus placebo.

Conclusions

The overall evidence was very high for anti-TNF treatment (1b, SOR: A) with respect to efficacy and safety, while it was low for biologic treatment other than anti-TNF (3, SOR: C).

Introduction

Treatment with anti-TNF has shown short- and long-term efficacy without major safety issues in clinical trials of patients with active AS. At the moment, four different anti-TNF agents are available and approved for the treatment of AS (infliximab, etanercept, adalimumab and golimumab).

In 2003, the Assessment in SpondyloArthritis international Society (ASAS) proposed recommendations for the use of anti-TNF agents in patients with AS, based on a Delphi questionnaire, published data, clinical expertise and a consensus meeting among experts [1, 2]. In 2006, the ASAS/European League Against Rheumatism (EULAR) management recommendations of AS were published.

These include guidance on non-pharmacological and pharmacological treatment including the use of TNF blockers. The recommendations for the use of anti-TNF agents and the ASAS/EULAR management recommendations are complimentary.

The recommendations for the use of anti-TNF agents were updated in 2006 [3], since it was felt that the research had rapidly evolved in this area after the first publication. In the first update, several aspects of treatment with anti-TNF agents, such as the initiation, use and withdrawal of anti-TNF treatment, based on data on the efficacy and safety of those agents were taken into account. In 2009 it was decided that a second update of both the ASAS/EULAR recommendations for the management of AS and the recommendations for the use of anti-TNF agents should be performed. Two systematic literature reviews were performed to search for the underlying evidence: one on biologics and one on non-pharmacological and pharmacological (excluding biologics) treatment.

The primary outcome of interest for this systematic literature review was the evidence on the long-term efficacy and safety of TNF blockers in AS. This includes information on a possible distinction between the different TNF blockers, information about switching between TNF blockers in case of inefficacy or safety concerns, efficacy and safety of other biologics than TNF blockers and the efficacy of biologics including TNF blockers in patients fulfilling the ASAS classification criteria for axial SpA but not yet the modified New York criteria for AS.

Methods

Included study designs

Randomized controlled trials (RCTs) were considered as the ideal study design for calculation of the intended analyses.

However, since a low number of RCTs was anticipated, all possible studies (quasi-randomized studies, non-randomized studies, case-control studies) as well as abstracts from the EULAR and ACR annual meetings for the years 2007-09 were included.

Systematic literature search

A systematic literature search for published articles was performed for the time period 1 January 2005 (which represents the date after the end of the last systematic literature review on this topic [4]) to 1 December 2009, using the PubMed, Embase and Cochrane databases with the assistance of an experienced librarian. Furthermore, a search of published abstracts in the online abstract libraries of the EULAR and the ACR annual meetings for the years 2007-09 for additional relevant but still unpublished studies was performed by hand. The terms that were used for each search were 'ankylosing spondylitis', 'spondyloarthritis', 'anti-TNF', 'biologics', 'infliximab', 'etanercept', 'adalimumab', 'anakinra', 'abatacept', 'rituximab' in all possible combinations of at least two of the terms and up to all terms together. The complete search strategies for the database searches are provided in supplementary appendix 1, available as supplementary data at Rheumatology Online.

Selection of studies

All reports (published papers and abstracts of meetings) had to deal with patients fulfilling the modified New York criteria for AS [5] or the ASAS classification criteria for SpA [6].

After collection, each title and abstract was examined for suitability in the review by excluding these studies that met the following exclusion criteria: duplicates of papers, incomplete data, reports that had longer follow-ups available in other papers than the ones found (in this case, the longer follow-up papers were included in the final analysis), case reports without follow-up information and publications or reports with 'wrong outcome' (e.g. listing AS or SpA as keywords but not reporting about these diseases in particular) (figure 1). The full papers that were excluded from the analyses are listed in supplementary appendix 2, available as supplementary data at Rheumatology Online.

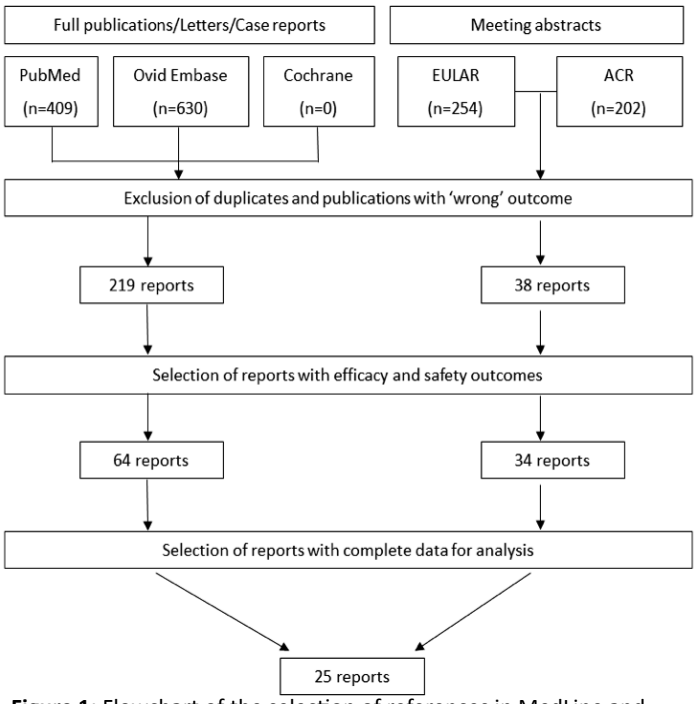


Figure 1: Flowchart of the selection of references in MedLine and Embase database as well as in the abstract books of the EULAR and ACR meetings, which served as the basis for this literature search. During the process, duplicates of papers, incomplete data, reports with longer follow-ups available in other papers, case reports without follow-up information and publications or reports with 'wrong outcome' were excluded.

Data extraction, data analysis and quality appraisal

From the studies that could be included in the analysis, all relevant efficacy and safety data were extracted and entered into standard data extraction forms in a Microsoft Excel- file according to the key components of the PICO (Participants, Interventions, Comparisons and Outcomes) method (supplementary appendix 1, available as supplementary data at Rheumatology Online). Calculations were made for the effect sizes (ESs, mean change in score divided by the baseline SD) for treatment [treatment effect (TE), the mean change in the index group minus the mean change in the comparator group divided by a pooled baseline SD] and for the Guyatt's ES (mean change in the index group divided by the SD of the change in the placebo group) according to all reported measures: disease activity (BASDAI [7]), metrology (BASMI [8]) and function (BASFI [9], CRP, ESR), but also the number needed to treat (NNT) for response to treatment according to the ASAS definitions (ASAS response [10]). The latter is used for assessment of the efficacy of study drugs by using the ASAS group core set of criteria for symptomatic improvement in AS [10] and is measured by a 20 and 40% response according to the ASAS criteria [10] and an improvement in the '5 out of 6 criteria' [11]. ASAS 20 response is defined as an improvement of not <20% and an absolute improvement of at least 1U (on a scale of 0-10) in at least three of the following

four domains: patient's global assessment, pain, function (represented by the BASFI score) and inflammation [represented by the mean of the two morning stiffness-related BASDAI numeric rating scale (NRS) scores]. Furthermore, there must be an absence of deterioration, which is defined as worsening of not >20% and net worsening of not >1U (on a scale of 0-10), in the remaining domain. Similarly, ASAS 40 response is defined as an improvement of not <40% and an absolute improvement of at least 1U (on a scale of 0-10) in at least three of the four domains mentioned above, while there should be no worsening in any of the domains. To meet the '5 out of 6 criteria', a 20% improvement in any five of the following six domains is required: the four domains used for ASAS 20 plus the CRP value and spinal mobility (assessed by the BASMI score).
Since we decided to include not only RCT alone, but also other types of studies (see above) in this review, a greater heterogeneity of the results was expected for all analyses.
A further assessment was made for each included study according to the Oxford Center for Evidence-based Medicine (CEBM) level of evidence, which gives studies a score for 'level of evidence' (1a-5) and a score for 'grade of recommendation' (A-D [12]). Analysis of safety data and adverse events (AEs) was done in a descriptive way in summary tables.
The results were finally presented to the ASAS/EULAR expert committee during the process of the update of the ASAS/EULAR management recommendations of AS and the update of the recommendations for the start of TNF-blocking agents.

Results

Process of the literature review

Overall, 409 reports were identified in MedLine and 630 reports were identified in the Embase database, while no report was found in the Cochrane database.
The search of the abstract meetings revealed 254 reports at EULAR and 202 reports at the ACR meeting. After exclusion of duplicates, 257 reports remained for validation, 64 reports were found to be dealing with efficacy and/or safety outcomes of patients and finally, 25 papers were found to have useful data for analysis (figure 1).

Efficacy

Calculation of ESs for treatment outcomes

The comparison between anti-TNF treatment versus placebo showed superior outcome for the treatment effect in favour of the anti-TNF treatment [13-21]. For the evaluation of the BASDAI, the ES (95% CI) varied between studies from 0.34 (0.08, 0.6) to 1.5 (0.45, 2.5) (table 1).
For evaluation of the BASFI, the ES (95% CI) varied between 0.07 (-0.21, 0.34) and 2.5 (1.3, 3.7), whereas for evaluation of the BASMI, the ES (95% CI) was only available for golimumab [0.08 (-0.20, 0.31)] (table 1).
Furthermore, data for different other outcomes such as occiput-to-wall measurements, chest expansion, physician's global and patient's global were only available in some of the studies [18, 20]. The treatment effect for patient's global assessment was 0.53 (0.17, 0.89), for physician's global assessment 1.3 (0.67, 1.9), for chest expansion zero (-0.35, 0.35), for occiput-to-wall it varied between -0.22 (-0.52, 0.09) and 0.01 (0.34, 0.37) and for modified Schober's test between 0.06 (-0.29, 0.42) and 0.28 (-0.03, 0.58).
The treatment effect for continuous versus on-demand anti-TNF treatment could only be calculated for infliximab, with an ES (95% CI) of 0.76 (0.44, 1.1) for BASDAI and 0.74 (0.42, 1.1) for BASFI, 0.53 (0.22, 0.84) for patient's global assessment, 0.03 (-0.28, 0.34) for the physical component of the short form 36 (SF-36) questionnaire and 0.19 (-0.12, 0.5) for the

mental component of SF-36.
The Guyatt's ES could only be calculated for golimumab in AS [19] and infliximab in non-radiographic SpA [21].

Calculation of numbers needed to treat

The calculation of the NNT for achieving all different treatment outcomes revealed only minor variations between the TNF blockers but superiority as compared with placebo [14, 18-26], with NNTs of 2.3-2.7 for ASAS 20, 2.9-3.7 for ASAS 40, 2.4-2.8 for ASAS 5/6, 2.5 for BASDAI 50 and 4.7-5.9 for ASAS partial remission.
Similar NNTs were found for patients with nonradiographic axial SpA, with 2.3 for ASAS 20, 1.6-2.4 for ASAS 40, 3.2 for ASAS 5/6 and 2.3-2.7 with ASAS partial remission (table 2).
In the comparison of continuous versus on-demand treatment with TNF blockers, the NNTs for ASAS 20 response were 4.2 versus 9.1 patients, for ASAS 40 response 6.7 versus 8.3 patients and for ASAS partial remission 5.9 versus 20.0 patients, respectively. For the differentiation between patients with versus without total spinal ankylosis, the NNTs varied between 2.4 for ASAS partial remission and 9.1 for ASAS 5/6 (table 3).

Efficacy of TNF blockers on extraspinal manifestations of the disease

One study from patients diagnosed as SpA according to the Amor criteria [27] provided data on the efficacy of TNF blockers in peripheral manifestations of the disease.
Patients with refractory disabling heel enthesitis were treated with etanercept or placebo. Patient's global assessment, heel pain and WOMAC improved significantly in the etanercept group as compared with placebo, already after 2 weeks of treatment.

Treatment with biologics other than TNF blockers

Overall, only small studies on biologics other than TNF blockers were available, and all of these studies included patients with advanced disease [28-32]. None of the studies was placebo-controlled. The compounds used were rituximab, anakinra or abatacept. All of the compounds showed only minor improvement in disease-related indices, and because there are no control groups, the level of improvement is difficult to interpret.
For rituximab in anti-TNF naïve patients [28], there were significant within-group improvements in BASDAI (p=0.047), pain as reported by the patient (p=0.021) and improvement in CRP (p=0.017). Further data published in the full paper of this abstract in 2010 showed a good improvement of all assessed parameters (50% in BASDAI50, 40% in ASAS 40) as compared with a poor response in those patients who had failed TNF blocker therapy before rituximab treatment (10% in ASAS40, none in BASDAI 50). For anakinra [31], the rate of patients showing sufficient ASAS response was reported as 25% for ASAS 20 and 20% for ASAS 40, while BASDAI improved from 5.8 to 4.6 and there was no change in CRP, as compared with baseline. The data of this study were included in abstract form in the first version of the recommendations [4], whereas the full paper is now available for the current report. For abatacept, there was only minor response of single patients [30].

TNF blockers in juvenile SpA

Only one small study published in abstract form [33] including patients with juvenile SpA patients with established AS (n=5 patients) and undifferentiated SpA (n=19) treated with infliximab could be used for data analysis. In this study, the amount of active joints, tender entheses, pain, CRP and HAQ showed significant decrease after 1 year in all patients. The mean amount of active joints decreased from 4.7 (1.7) to 0, the mean amount of tender entheses from 11.9 (10.7) to 0, the mean CRP from 24.8 (10) to 1.3 (3.1), the pain (mean of NRS) from 7.2 (2.0) to 1.7 (2.7), while the mean score in the childhood HAQ did not show

changes in the patients who were initially treated with infliximab and remained on this treatment.

Level of evidence and strength of recommendations for treatment with biologics in AS

The overall research evidence for all TNF blockers is rated with 1b+ (table 4), including two studies with patients with non-radiographic axial SpA [18, 21], which showed similar outcomes as compared with studies of patients with established AS.

Furthermore, the research evidence for the use of infliximab on demand and for the use of etanercept in a dose of 1x25 mg/week in patients with low disease activity was also rated with 1b+. There are no data on dose adjustment for adalimumab at the current time point.

The strength of recommendation (SOR) for the use of all available TNF blockers in AS in the recommended dose is rated with A, with the exception of treatment with etanercept 1x25 mg/week, where the SOR is rated with B (table 4).

The research evidence for the treatment of patients with DMARDs concomitant to TNF blockers as well as switching between TNF blockers is 3+, while the SOR was rated with C. Although the analyses for switching between anti-TNF compounds have been based on patients treated with infliximab after failure of treatment with etanercept, it is expected that other combinations among other TNF blockers would reveal similar outcomes.

For treatment with biologics other than TNF blockers, the available data showed a research evidence of 3 for anakinra based on the same study as already included in the previous review; however, this result remains to be confirmed by further studies. Data for abatacept and rituximab are scarce and did not allow for any conclusions, while no data for tocilizumab were available within the period of analysis in this update. The SOR for the use of anakinra in AS was rated with C.

For the use of biologics in patients with juvenile onset of SpA, only data on infliximab were available. The research evidence was 3, which can be translated to SOR rated with C (table 4).

Incidence of concomitant extra-articular manifestations in AS during treatment with TNF blockers

TNF blockers showed beneficial effect on the treatment of extra-articular manifestations (EAMs) of AS as compared with treatment with placebo. Data were available for infliximab, etanercept and adalimumab, while data from studies with golimumab were not available at this time point.

Two main concomitant EAMs were recognized: anterior uveitis (AU) and IBDs.

As suggested in a meta-analysis for the treatment of AU, which included only patients with infliximab and etanercept (adalimumab data were not available at this time point), the incident rates during anti-TNF treatment were 4.4 (range 1.1-8.0) per 100 patient-years (pys) as compared with 15.6 (7.8-27.9)/100 pys during placebo treatment (all $p<0.05$) [34, 35]. In a more recent paper [35], the incidence of AU flares under open-label adalimumab treatment was 7.4/100 pys and statistically significantly lower than the incidence rate of AU during the previously performed placebo-controlled period of the same trial with 15.0 AU flares/100 pys ($p=0.001$).

Another follow-up study with patients treated open label with etanercept showed similar superiority of etanercept (13 AU flares/100 pys), as compared with the numbers known from the placebo-controlled period of the same trial [36]. Similar data were shown in a meta-analysis that was available in abstract form [37] (the full paper was published in 2010).

Table 1: Effect sizes (ES) for treatment effect and Guyatt´s ES with 95% confidence interval (95% CI) for BASDAI, BASFI and BASMI outcomes.

			BASDAI		BASFI		BASMI	
Study	**n patients (treatment / comparator)**	**Treatment duration**	**Treatment effect (95% CI)**	**Guyatt'S ES (95% CI)**	**Treatment effect (95% CI)**	**Guyatt'S ES (95% CI)**	**Treatment effect (95% CI)**	**Guyatt'S ES (95% CI)**
ETN 2x25mg vs. Placebo [15]	9 / 11	24 weeks	1.5 (0.45 – 2.5)	--	2.5 (1.3 – 3.7)	--	--	--
ADA 40mg vs. Placebo [18]	22 / 24	12 weeks	1.2 (0.57 – 1.8)	--	0.07 (-0.21 – 0.34)	--	--	--
GOL 50/100mg vs. Placebo [19]	278 / 78	24 weeks	0.34 (0.08 – 0.60)	1.3 (1.0 – 1.6)	0.33 (0.07 – 0.59)	0.95 (0.68 – 1.2)	0.08 (-0.20 – 0.31)	0.60 (0.34 – 0.86)
INF 5mg/kg cont. vs. on demand [20]	124 / 61	58 weeks	0.76 (0.44 – 1.1)	--	0.74 (0.42 – 1.1)	--	--	--
INF 5mg/kg on demand +/- MTX [20]	62 / 61	58 weeks	-0.15 (-0.50 – 0.21)	--	0.27 (-0.08 – 0.63)	--	--	--
ETN 2x25mg vs. 1x25mg [17]	20/21	6 months	-1.0 (-0.35 – 1.7)	--	--	--	--	--
INF 5mg/kg vs. Placebo [21]	20/20	16 weeks	--	1.41 (0.72 – 2.1	--	1.2 (0.53 – 1.87)	--	--

A value of <0,6 indicates a small, a value of <0,8 indicates a moderate, a value of ≥0,8 indicates a large change. INF = infliximab, ETN = etanercept, ADA = adalimumab, GOL = golimumab.

Table 2: Calculation of numbers needed to treat (NNT) for the comparison of treatment between TNF blockers and placebo.

Study	Intervention vs. Placebo	Duration of follow-up	n patients (treatment / placebo)	Calculated NNT for different treatment outcomes				
				ASAS 20	ASAS 40	ASAS 5/6	BASDAI 50	Part. Rem.
van der Heijde et al [14]	ADA 40mg	12 weeks	208/107	2.7	3.7	2.8	--	5.9
Haibel et al [18]	ADA 40mg (no sacroiliitis on X-rays)	12 weeks	22/24	2.3	2.4	--	--	2.7
Inman et al [19]	GOL 50/100mg	24 weeks	278/78	2.6	3.0	--	--	--
Barkham et al [21]	INF 5mg/kg	16 weeks	20/20	--	1.6	3.2	--	2.3
van der Heijde et al [22]	INF 5mg/kg	24 weeks	201/78	2.4	2.9	2.4	2.5	4.7

Part. Rem.: ASAS partial remission; INF: infliximab; ADA: adalimumab; GOL: golimumab.

A summary of the studies dealing with the occurrence of AU in patients with anti-TNF during the time period analysed in this update is shown in table 5.

For the incidence of IBD, other differences between the TNF blockers were found, with significantly lower incidence rates during treatment with infliximab, as compared with etanercept or adalimumab [38] (table 6).

Safety

AEs

The incidence of AEs between treatment with TNF blocker and placebo, between TNF blockers in different treatment doses or during treatment with TNF blockers with or without concomitant treatment with other compounds is shown in table 7. Overall, the incidence of AEs as reported in the present updated review is in line with those reported in the first version of the recommendations [4].

Infections

In a meta-analysis comparing the risk difference between TNF blockers and placebo [39], the incidence rate of non-serious infections was 84.5 (58.4)/100 pys in patients treated with TNF blockers during the randomized control phases of the trials (RCTs) and reduced to 64.4 (56.7)/ 100 pys during the open-label phases. The latter was similar to the incidence of non-serious infections registered in the placebo arm of the RCTs, with an incidence of 63.6 (63.0) non-serious infections/100 pys.

In contrast, the analysis of serious infections showed an incidence of 2.3 (4.0) under TNF blockers during the RCTs and of 1.4 (2.8) during the open-label phases, as compared with an incidence of 1.4 (2.83) serious infections under placebo. An overview on the available data of the relative risk for infections in patients with AS is shown in table 7.

Table 3: Calculation of numbers needed to treat (NNT) for the comparison of treatment between TNF blockers and other comparators or between different groups with the same treatment.

Study	Intervention	Duration of follow-up	n patients (treatment / comparator)	Calculated NNT for different treatment outcomes				
				ASAS 20	ASAS 40	ASAS 5/6	BASDAI 50%	Part. Rem.
Breban [20]	INF 5mg/kg, continuous use	58 weeks	124/123	4.2	6.7	--	--	5.9
	INF 5mg/kg on demand use, addition of MTX	58 weeks	61/62	9.1	8.3	--	--	20.0
van der Heijde et al [23]	ETN 1x50mg	12 weeks	206/201	2.7	3.9	2.3	2.5	3.9
van der Heijde et al [23]	ETN 2x25mg	12 weeks	201/206	2.9	4.8	2.3	2.6	6.5
Marzo-Ortega [24]	MTX+Plac vs. MTX+Inf	30 weeks	28/14	3.1	--	--	--	--
Pérez-Guijo [25]	INF 5mg/kg + MTX 7.5 mg/ week	30 weeks	9/10	1.5	--	--	--	3.0
Cheung [26]	INF 5mg/kg with vs. without spinal ankylosis	54 weeks	11/16	5.6	5.0	9.1	5.6	2.4

Part. Rem.: ASAS partial remission; INF = infliximab; ETN = etanercept.

Formation of antibodies against TNF blockers

Only a few studies were dealing with the issue of antibody formation during treatment with TNF blockers in AS. In one study with infliximab [40] - patients who discontinued and re-started TNF blockade in the same treatment regimen - immunogenicity had no influence on the response to re-treatment or on safety outcomes in the long-term follow-up. While antibody formation due to immunogenicity was not detected during and after treatment with etanercept [41], antibody formation correlated well with undetectable serum trough levels, with inefficacy and infusion or injection reactions in patients treated with infliximab [42] or adalimumab [43] in two small studies.

Table 4: Research evidence for treatment with different biologic compounds and dosages in patients with AS and SpA for the years 2006 - 2010, compared to the last published version of the recommendations from the years 2001 - 2005 [4].

Intervention	Research Evidence		SOR (A–D)
	2005	2010	
INF 5mg/kg cont. *(only AS)*	1b+	1b+	A
INF 5mg/kg cont. *(non-radiographic SpA)*	NA		
INF 5mg/kg on dem.	NA		A
INF 5mg/kg on dem. + MTX	NA		A
ETN 1x50mg	NA	1b+	A
ETN 2x25mg	1b+		A
ETN 1x25mg	NA		B
ADA 1x40mg	3+ *(only AS)*	1b+ *(both AS and non-radiographic SpA)*	A
GOL 1x50mg/100mg	NA	1b+	A
Switch (INF to ETN)	NA	3+	C
Anakinra, Abatacept, Rituximab	3± *(anakinra only)*	3±	C
INF 5mg/kg in JuvSpA	--	3	C

+ = supportive, - = not supportive, ± = uncertain. SOR = strength of recommendations; NA = no data available.

Discussion

This report is a systematic literature review that was performed in order to obtain the detailed data for the second update of the ASAS/EULAR recommendations for the management of AS, with a special topic of interest being the treatment with biologics. After the first version of the recommendations published by ASAS in 2003 [2] and a first update in 2006 [3], a substantial number of new publications with long-term data on TNF blockers and reports on other biologics made this second update necessary.

On the basis of the published data on efficacy and safety, the research evidence is determined and the SOR is provided.

More data on all TNF blockers approved for the treatment of AS were available for the time January 2005-December 2009, as compared with the time before 2005, where the first version and the first update of the recommendations were available. Overall, all anti-TNF blockers proved to be efficacious in AS and SpA with a high level of research evidence (1b+).

Table 5: Effect of TNF blockers on anterior uveitis (AU) and incidence of anterior uveitis during treatment with TNF blockers as compared to treatment with placebo in patients with AS.

Trial	Type of trial	INF (n/100py)	ETN (n/100py)	ADA (n/100py)	Placebo (n/100py)
Braun J et al [35]	Meta-analysis n=717	4.4 (1.1-8) p=0.005 vs Plac.	7.9 (5.5-11.1) p=0.05 vs Plac.	--	15.6 (7.8-27.9) p=0.01 vs. INF+ETN
Rudwaleit et al [36]	Open-label ADA n=1250	--	--	7.4 p=0.001	15.0
Davis [37]	Double-blind + Open label ETN n=406	--	13	--	22
Sieper J et al [38]	Meta-analysis Double-blind ETN n=508	--	8.6 (4.5-14.2) p=0.031 vs. Plac.	--	19.3 (11.0-29.8) n=249
	Meta-analysis ETN vs. SSZ n=379	--	10.7 (5.5-17.4) p=0.486 vs. SSZ	--	SSZ: 14.7 (6.5-26.5) n=187
	Meta-analysis Double-blind + Open label ETN n=1074	--	12.0 (10-14.1)	--	--

Numbers indicate the occurence of AU per 100pys under treatment with either TNF blockers or placebo. ETN = etanercept; ADA = adalimumab.

Table 6: Incidence of acute inflammatory bowel disease (IBD) in AS patients treated with anti-TNF. P values were p<0.001 for infliximab versus. etanercept, 0.02 for infliximab versus. adalimumab and 1.0 for etanercept versus. adalimumab [39].

Treatment	Incidence of IBD / 100 patient years	Total number of treated patients	Numbers of IBD cases
Placebo	1.3	434	2
Infliximab	0.2	366	1
Etanercept	2.3	419	14
Adalimumab	2.3	295	3

Table 7: Calculation of relative risk (RR) and 95% confidence intervals (95% CI) for infections in patients with AS, as pooled relative risk for the placebo-controlled trials and for trials with comparators other than placebo.

Type of study	Adverse event	Study	RR (95% CI)
Studies comparing TNF treatment with placebo	Serious infections	Meta-analysis of anti-TNF vs. placebo treatment [40]	Risk difference: 0.4% (95% CI -8% to 1.6%)
Studies comparing TNF treatment with other comparators than placebo	All adverse events (for comparison)	Etanercept 1x50mg vs. 2x25mg [23]	0.84 (0.39 - 1.81)
	Infusion reactions	Infliximab continuous vs. on demand [20]	2.23 (1.00 - 4.94)
		Infliximab on demand, without vs. with MTX [20]	3.04 (0.64 - 14.52)

In comparison, the data of the last recommendations were only based on patients with established disease, proposing a research evidence level of 1b+ for continuous infliximab (5 mg/kg/6 weeks) and for etanercept (2-25 mg/week) but a research evidence level of 3+ for adalimumab (40 mg/2 weeks), while there were no data for the treatment with etanercept in the dose of 1-50 mg/week or for golimumab. In this update, also a research level evidence of 1+ can be given to adalimumab and golimumab in the approved doses.

The SOR for the use of all available TNF blockers in AS in the dose recommended by the label of each compound is rated with A. However, the present calculations also support treatment with infliximab on demand and treatment with etanercept in the decreased dose of 1-25mg for patients with established AS who remain on low disease activity (research evidence 1b+). For the latter, the SOR is rated with B. There are no data on dose adjustment for adalimumab at the current time point.

In contrast to the previous version of the recommendations, data on the treatment with DMARDs concomitant to TNF blockers as well as switching between TNF blockers are now available. The calculated research evidence is 3+, while the SOR was rated with C. Although the analyses for switching between anti-TNF compounds have been based on patients treated with infliximab after failure of etanercept treatment, it is expected that other combinations among other TNF blockers would reveal similar outcomes.

Data from new biologic compounds other than TNF blockers were also available this time. However, only studies with anakinra provided information that could be used for calculations, showing a research evidence of 3 (SOR rated with C). Data for abatacept, rituximab and tocilizumab were scarce and did not allow for any conclusions.

For the use of biologics in patients with juvenile onset of SpA, only limited data were

available. There, infliximab showed research evidence on a level of 1b+, which can be translated to a SOR rated with A.
Finally, the available data indicate a beneficial effect of TNF blockers for the treatment of the two main EAMs in the same patients, AU and inflammatory bowel diseases, with only minor differences between the available compounds.
With respect to safety, the overall incidence of AEs was not different to what had been reported previously.
However, treatment with TNF blockers showed a somewhat higher infection rate as compared with placebo, although there was no difference between the treatments in the comparison for serious infections. Nevertheless, it seems that the overall incidence of infections during treatment with TNF blockers decreased with longer duration of the studies, which might be due to selection of patients who stay in the study. In the short-term follow-up studies with patients treated with biologics other than TNF blockers, no major safety issues were reported. Finally, the formation of antibodies against TNF blockers has been reported in some studies and has correlated with low serum levels of the compounds, mainly in studies with mAbs. Nevertheless, immunogenicity had no influence on the response to re-treatment or on safety outcomes in one small study. More data are necessary to determine the clinical relevance of the formation of anti-drug antibodies.
In conclusion, the analysis of all available literature data support the use of the currently available TNF blockers for the treatment of patients with advanced AS who are fulfilling the ASAS recommendations for such treatment.
Furthermore, data from first studies from patients with non-radiographic SpA show a similar response to TNF blockers. Overall, biologics other than TNF blockers cannot be recommended at the current time because of lack of sufficient evidence. DMARDs do not add to efficacy or safety as concomitant treatment with anti-TNF in patients with AS. TNF blockers show good evidence in patients with juvenile onset of SpA, but these data are based on a limited number of studies.

Supplementary data

Supplementary data are available at Rheumatology Online.

References

1. Pham T, van der Heijde D, Calin A, *et al.* Initiation of biological agents in patients with ankylosing spondylitis: results of a Delphi study by the ASAS Group. Ann Rheum Dis 2003;62:812-6.
2. Braun J, Pham T, Sieper J, *et al.* International ASAS consensus statement for the use of anti-tumour necrosis factor agents in patients with ankylosing spondylitis. Ann Rheum Dis 2003;62:817-24.
3. Braun J, Davis J, Dougados M, *et al.* First update of the international ASAS consensus statement for the use of anti-TNF agents in patients with ankylosing spondylitis. Ann Rheum Dis 2006;65:316-20.
4. Zochling J, van der Heijde D, Burgos-Vargas R, *et al.* ASAS/EULAR recommendations for the management of ankylosing spondylitis. Ann Rheum Dis 2006;65:442-52.
5. van der Linden S, Valkenburg H, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361-8.
6. Rudwaleit M, van der Heijde D, Landewé 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.
7. Garrett S, Jenkinson T, Kennedy L, *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.
8. Jenkinson T, Mallorie P, Whitelock H, *et al.* Defining spinal mobility in ankylosing spondylitis (AS). The Bath AS Metrology Index. J Rheumatol 1994;21:1694-8.
9. Calin A, Garrett S, Whitelock H, *et al.* A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. J Rheumatol 1994;21:2281-5.
10. Anderson J, Baron G, van der Heijde D, *et al.* Ankylosing spondylitis assessment group preliminary definition of short-term improvement in ankylosing spondylitis. Arthritis Rheum 2001;44:1876-86.
11. Brandt J, Listing J, Sieper J, *et al.* Development and preselection of criteria for short-term improvement after anti-TNF alpha treatment in ankylosing spondylitis. Ann Rheum Dis 2004;63:1438-44.
12. Oxford Centre for Evidence-based Medicine. Levels of evidence. http://www.cebm.net/index.aspx?o=1025 (28 October 2010, date last accessed).
13. Davis J, van der Heijde D, Braun J, *et al.* Sustained durability and tolerability of etanercept in ankylosing spondylitis for 96 weeks. Ann Rheum Dis 2005;64: 1557-62.
14. van der Heijde D, Kivitz A, Schiff M, *et al.* Efficacy and safety of adalimumab in patients with ankylosing spondylitis: results of a multicenter, randomized, double-blind, placebo-controlled trial. Arthritis Rheum 2006;54:2136-46.
15. Inman R, Clegg D, Davis J, *et al.* Etanercept in adult patients with early onset ankylosing spondylitis. J Rheumatol 2006;33:1634-6.
16. Davis J, Revicki D, van der Heijde D, *et al.* Health-related quality of life outcomes in patients with active ankylosing spondylitis treated with adalimumab: results from a randomized controlled study. Arthritis Rheum 2007;57:1050-7.
17. Berthelot J, Varin S, Cormier G, *et al.* 25mg etanercept once weekly in rheumatoid arthritis and spondylarthropathy. Joint Bone Spine 2007;74:144-7.
18. Haibel H, Rudwaleit M, Listing J, *et al.* Efficacy of adalimumab in the treatment of axial spondylarthritis without radiographically defined sacroiliitis: results of a twelve-week randomized, double-blind, placebo-controlled trial followed by an open-label extension up to week fifty-two. Arthritis Rheum 2008;58:1981-91.
19. Inman R, Davis J, Heijde D, *et al.* Efficacy and safety of golimumab in patients with ankylosing spondylitis: results of a randomized, double-blind, placebo-controlled, phase III trial. Arthritis Rheum 2008;58:3402-12.
20. Breban M, Ravaud P, Claudepierre P, *et al.* Maintenance of infliximab treatment in ankylosing spondylitis: results of a one-year randomized controlled trial comparing systematic versus on-demand treatment. Arthritis Rheum 2008;58:88-97.

21. Barkham N, Keen H, Coates L, *et al.* Clinical and imaging efficacy of infliximab in HLA-B27-positive patients with magnetic resonance imaging-determined early sacroiliitis. Arthritis Rheum 2009;60:946-54.
22. van der Heijde D, Dijkmans B, Geusens P, *et al.* Efficacy and safety of infliximab in patients with ankylosing spondylitis: results of a randomized, placebo-controlled trial (ASSERT). Arthritis Rheum 2005;52:582-91.
23. van der Heijde D, Da Silva J, Dougados M, *et al.* Etanercept 50mg once weekly is as effective as 25mg twice weekly in patients with ankylosing spondylitis. Ann Rheum Dis 2006;65:1572-7.
24. Marzo-Ortega H, McGonagle D, Jarrett S, *et al.* Infliximab in combination with methotrexate in active ankylosing spondylitis. A clinical and imaging study. Ann Rheum Dis 2005;64:1568-75.
25. Perez-Guijo V, Cravo A, Castro Mdel C, *et al.* Increased efficacy of infliximab associated with methotrexate in ankylosing spondylitis. Joint Bone Spine 2007;74:254-8.
26. Cheung P, Tymms K, Wilson B, *et al.* Infliximab in severe active ankylosing spondylitis with spinal ankylosis. Intern Med J 2008;38:396-401.
27. Dougados M, Combe B, Braun J, *et al.* A randomised, multicentre, double-blind, placebo-controlled trial of etanercept in adults with refractory heel enthesitis in spondyloarthritis: the HEEL trial. Ann Rheum Dis 2010;69:1430-5.
28. Song I, Heldmann F, Rudwaleit M, *et al.* Major clinical response of rituximab in active TNF-blocker naive patients with ankylosing spondylitis but not in TNF-blocker-failure patients - an open label clinical trial. Ann Rheum Dis 2009;68(Suppl3):74.
29. Nocturne G, Dougados M, Constantin A, *et al.* Lack of efficacy of rituximab in spondylarthropathies: data of 8 patients prospectively followed in the French AIR ('auto-immunity and rituximab') registry. Ann Rheum Dis 2009;68(Suppl3):626.
30. Berner B, Schedel J, Guenaydin I, *et al.* Abatacept for therapy of spondyloarthritis due to therapy failure or contraindications of TNF-alpha antagonists. Ann Rheum Dis 2009;68(Suppl3):623.
31. Haibel H, Rudwaleit M, Listing J, *et al.* Open label trial of anakinra in active ankylosing spondylitis over 24 weeks. Ann Rheum Dis 2005;64:296-8.
32. Bennett A, Tan A, Coates L, *et al.* Sustained response to anakinra in ankylosing spondylitis. Rheumatology 2008; 47:223-4.
33. Burgos-Vargas R, Casasola-Vargas J, Gutiérrez-Suárez R, *et al.* An open, observational, extension study of a three-month, randomized, placebo-controlled trial to assess the long-term efficacy and safety of infliximab in juvenile-onset spondyloarthritis (Jo-Spa). Arthritis Rheum 2008;58(9 Suppl):S578.
34. Braun J, Baraliakos X, Listing J, *et al.* Decreased incidence of anterior uveitis in patients with ankylosing spondylitis treated with the anti-tumor necrosis factor agents infliximab and etanercept. Arthritis Rheum 2005;52:2447-51.
35. Rudwaleit M, Rodevand E, Holck P, *et al.* Adalimumab effectively reduces the rate of anterior uveitis flares in patients with active ankylosing spondylitis: results of a prospective open-label study. Ann Rheum Dis 2009;68:696-701.
36. Davis J, van der Heijde D, Braun J, *et al.* Efficacy and safety of up to 192 weeks of etanercept therapy in patients with ankylosing spondylitis. Ann Rheum Dis 2008;67:346-52.
37. Sieper J, Koenig A, Baumgartner S, *et al.* Lower rates of uveitis with etanercept or sulphasalazine versus placebo in clinical studies in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68(Suppl3):73.
38. Braun J, Baraliakos X, Listing J, *et al.* Differences in the incidence of flares or new onset of inflammatory bowel diseases in patients with ankylosing spondylitis exposed to therapy with anti-tumor necrosis factor alpha agents. Arthritis Rheum 2007;57:639-47.
39. Fouque-Aubert A, Jette-Paulin L, Tebib J, *et al.* Risk of infections in patients with ankylosing spondylitis with TNF-blockers: systematic literature review. Ann Rheum Dis 2008;67(Supp. II):385.
40. Baraliakos X, Listing J, Brandt J, *et al.* Clinical response to discontinuation of anti-TNF therapy in patients with ankylosing spondylitis after 3 years of continuous treatment with infliximab. Arthritis Res Ther 2005;7:R439-44.

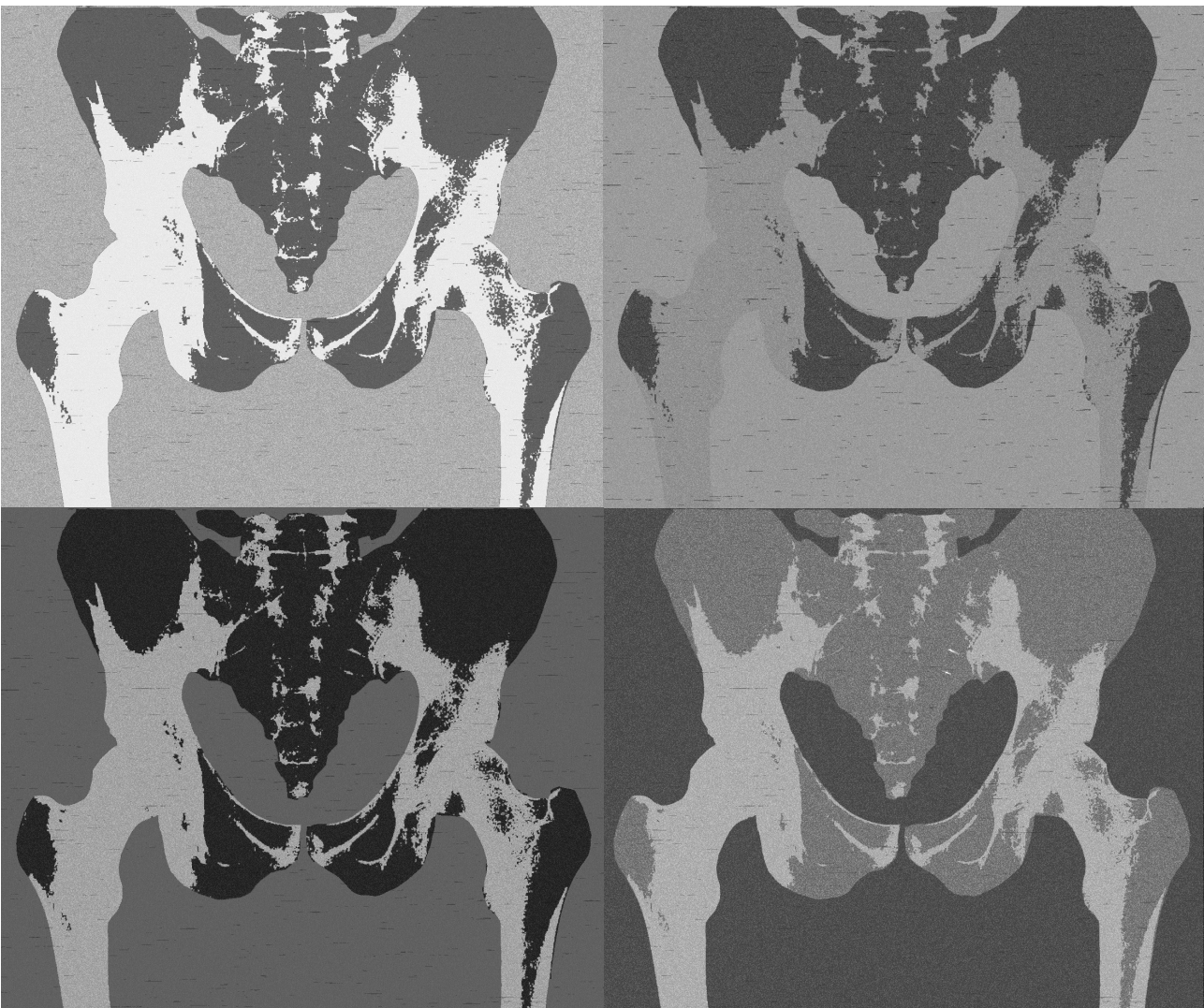
41. de Vries M, van der Horst-Bruinsma I, Nurmohamed M, *et al.* Immunogenicity does not influence treatment with etanercept in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68:531-5.
42. de Vries M, Wolbink G, Stapel S, *et al.* Decreased clinical response to infliximab in ankylosing spondylitis is correlated with anti-infliximab formation. Ann Rheum Dis 2007;66:1252-4.
43. de Vries M, Brouwer E, van der Horst-Bruinsma I, *et al.* Decreased clinical response to adalimumab in ankylosing spondylitis is associated with antibody formation. Ann Rheum Dis 2009;68:1787-8.

11

2010 update of the ASAS/EULAR recommendations for the management of ankylosing spondylitis

Ann Rheum Dis 2011;70:896-904

J. Braun
R. van den Berg
X. Baraliakos
H. Boehm
R. Burgos-Vargas
E. Collantes-Estevez
H. Dagfinrud
B. Dijkmans
M. Dougados
P. Emery
P. Geher
M. Hammoudeh
R. Inman
M. Jongkees
M. Khan
U. Kiltz
T. Kvien
M. Leirisalo-Repo
W. Maksymowych
I. Olivieri
K. Pavelka
J. Sieper
E. Stanislawska-Biernat
D. Wendling
S. Ozgocmen
C. van Drogen
B. van Royen
D. van der Heijde



Abstract

This first update of the ASAS/EULAR recommendations on the management of ankylosing spondylitis (AS) is based on the original paper, a systematic review of existing recommendations and the literature since 2005 and the discussion and agreement among 21 international experts, 2 patients and 2 physiotherapists in a meeting in February 2010. Each original bullet point was discussed in de tail and reworded if necessary. Decisions on new recommendations were made - if necessary after voting. The strength of the recommendations (SOR) was scored on an 11-point numerical rating scale after the meeting by email. These recommendations apply to patients of all ages that fulfill the modified NY criteria for AS, independent of extra-articular manifestations, and they take into account all drug and non-drug interventions related to AS. Four overarching principles were introduced, implying that one bullet has been moved to this section. There are now 11 bullet points including 2 new ones, one related to extra-articular manifestations and one to changes in the disease course. With a mean score of 9.1 (range 8-10) the SOR was generally very good.

Introduction

The European League against Rheumatism (EULAR) has developed management recommendations for various rheumatic conditions in the past decade [1–6] based on standard operating procedures published some years ago [7]. The basis for the methodology is the AGREE instrument [8]. A systematic literature review (SLR) serves as the basis for the expert discussions and the consensus process [9–11]. The Assessments in Ankylosing Spondylitis International Society (ASAS), which published a core set of endpoints for the disease more than 10 years ago [12] has taken the lead in developing recommendations for anti-tumour necrosis factor (TNF) therapy in ankylosing spondylitis (AS) [13], which have already been updated twice [14, 15]. The two organisations jointly developed the first set of recommendations for the management of AS together in 2005.

As this is a requirement of the EULAR standard operating procedures for management recommendations and as the field of spondyloarthritis is moving rapidly, an update of the first recommendations for the management of AS is needed after 5 years.

While the first version of the management recommendations was initially developed without patients, and as discrepancies between patients' and physicians' perspectives are well known [16], on this occasion patients were involved in the project group from the beginning. Moreover, other stakeholders, such as physiotherapists, were also represented in the project group. A patient-specific version of the first recommendations has been developed with the active support of patients of many European and North American countries [17]. The original and the patient version of the recommendations has been evaluated [18, 19] and disseminated in several countries [20–23].

AS is the prototype [24], a subtype, and an outcome of spondyloarthritis, particularly of the axial form of spondyloarthritis. Recent new classification criteria have widened the spectrum of spondyloarthritis by including earlier forms in addition to AS [25, 26]. This project has also led to a separation in the classification to predominantly axial and peripheral forms of spondyloarthritis. The term 'axial spondyloarthritis' covers patients with chronic back pain who have AS, which is defined by the presence of definite structural changes on radiographs in the sacroiliac joints, and patients with early or abortive forms of spondyloarthritis, which is defined by the presence of sacroiliac inflammation as detected by MRI or the presence of HLA-B27 in combination with the presence of features typical of spondyloarthritis [27, 28]. It can be anticipated that future trials will increasingly target axial spondyloarthritis rather than AS. Some trials with that aim have already been performed and some have started. However, as the evidence from such trials is currently limited it has been decided to restrict the recommendations to AS, although the project group unanimously agreed that these recommendations can equally be applied to patients with axial spondyloarthritis.

As the number of clinical trials and publications on AS therapy has steadily increased over the first decade of the millennium, this provided a sound rationale for a SLR.

Methods

ASAS and EULAR agreed in 2009 to collaborate in the development of the first update of the recommendations. To facilitate the process, it was decided that the convenor (JB) and the epidemiologist (DvdH) would maintain the same role that they undertook in the development of the first recommendations.

These original recommendations [1] formed the basis for the update. Two fellows performed the SLR, which needed an update since 2005 when the previous SLR was performed [11]. The international expert group included 21 rheumatologists, two orthopaedic surgeons, four patients (two of them were also rheumatologists) and one physiotherapist - representing 16 countries worldwide. The same group of international AS experts who participated in the development of the first recommendations was invited to participate.

The experts met on 25/26 February 2010 in Zurich. During the meeting, the data from the SLR dating from the previous search in 2005 until December 2009 were presented to the international experts. Each bullet point was discussed in detail until consensus was reached as to whether rewording was necessary. New recommendations were considered if this was proposed by a member of the panel.
Scoring on an 11-point numerical rating scale for the strength of recommendation was done by email by each expert for each bullet point after the meeting.
The methodology and detailed results of the SLR are described elsewhere in two separate papers: one dealing with biological agents and the other with all other management aspects such as non-biological drugs, education and physiotherapy (submitted).

Results

General definitions

The target population was defined as follows: the recommendations were to apply to all patients fulfilling the modified New York criteria for AS, independent of extra-articular manifestations. Patients of all ages, including paediatric patients, were included, and all pharmacological and non-pharmacological interventions for AS were taken into account.
The first discussion addressed whether the terminology of the recommendations should be changed to 'Recommendations for the management of axial spondyloarthritis'. The arguments in favour were mainly that the new classification criteria for axial spondyloarthritis [25, 26] are now available and they should therefore be included in the recommendations. The arguments against this were rather pragmatic, such as 'the world of rheumatology is not yet ready for that change'. Furthermore, there is a paucity of papers in early disease. The group finally decided to stick to the term 'AS' for the time being. However, every expert expressed the opinion that patients with early axial spondyloarthritis who do not yet fulfill the modified New York criteria are part of the same spectrum of disease and that these management recommendations most likely apply equally to those patients. Importantly, this patient population is already well recognised in the last update of the ASAS recommendations for anti-TNF therapy [15]. However, it should also be clearly stated that not all patients who fulfill classification criteria for axial spondyloarthritis will necessarily develop structural damage with radiographic changes in the sacroiliac joints and/or spine, which is presently considered essential in order for patients to fulfill currently used criteria for AS [29, 30]. This is actually similar to patients fulfilling the 2010 criteria for rheumatoid arthritis (RA) versus patients fulfilling the 1987 criteria for RA.
Although there are first hints that TNF blockers may be safer in AS compared with RA [31], a decision was made not to create a unique update on the safety of biological agents in AS/spondyloarthritis, but rather to rely on the extensive work done by Furst *et al.* [32] who have undertaken an annual consensus document on this topic from the 'Targeted therapies' meeting.

Results of the SLR

The detailed results will be published elsewhere (submitted).
However, the information that was obtained from the SLR was taken into account during the discussions of each bullet point.

Results of the discussions

The first change the expert group agreed on was, by analogy with other EULAR recommendations (e.g. management recommendations for RA, [6]), to define overarching principles of management.

Bullet point number 3 in the first published version of the recommendations [1] stating that the optimal management of patients with AS requires a combination of non-pharmacological and pharmacological treatment modalities has now been moved to this section.
Of note, the citations in this section are not the complete results of the SLR and they are not complete. They are just examples given to document the basis of the statements and notations made in the text.
An overview of the new recommendation is given in box 1.
The overarching principles of the management of patients with AS are:

- AS is a potentially severe disease with diverse manifestations, usually requiring multidisciplinary treatment coordinated by the rheumatologist.
- The primary goal of treating the patient with AS is to maximise long term health-related quality of life through control of symptoms and inflammation, prevention of progressive structural damage, preservation/normalisation of function and social participation.
- Treatment of AS should aim at optimal care and must be based on a shared decision between the patient and the rheumatologist.
- The optimal management of patients with AS requires a combination of non-pharmacological and pharmacological treatment modalities.

Comment

Patients with AS present with different disease manifestations [24] and a high proportion may run a severe course of disease [33]. The main health problems of patients with AS have recently been listed as part of an International Classification of Functioning, Disability and Health consensus process [34, 35].
It is important to stress that the rheumatologist is the expert who should take the lead in the management of patients with AS. The major aim for the treatment of rheumatic diseases is the preservation and gain of short and long-term health-related quality of life. The general view is that this is best achieved through control of symptoms and inflammation—with the aim to prevent deformity and disability due to structural damage caused by new bone formation and the decline of function and social participation.
Strength of recommendation: 9.5±0.1.
Thereafter, the bullet points were discussed point by point in considerable detail, and agreement was achieved on 11 points.
The updated recommendations are:

General treatment

The treatment of patients with AS should be individualised according to:

- The current manifestations of the disease (axial, peripheral, entheseal, extra-articular symptoms and signs)
- The level of current symptoms, clinical findings and prognostic indicators
- The general clinical status (age, gender, comorbidities, concomitant medications, psychosocial factors).

Comment

This general bullet point was not changed. It stresses that there may be considerable variation in how AS patients may present to the rheumatologist. The aim of management and appropriate interventions may thus also differ substantially. This implies that these aims must be tailored to the unique features of the particular AS patient.
Strength of recommendation: 9.5±0.1.

Disease monitoring

The disease monitoring of patients with AS should include:

- Patient history (e.g. questionnaires)
- Clinical parameters
- Laboratory tests
- Imaging

All according to the clinical presentation as well as the ASAS core set.

The frequency of monitoring should be decided on an individual basis depending on:

- Course of symptoms
- Severity
- Treatment.

Comment

This bullet point was not changed. It basically leaves the decision as to how frequently patients should be monitored to the rheumatologist in charge of management. This is mainly due to the fact that the course of disease may differ substantially between patients and different aspects, as stated in the bullet point, may need to be considered.

Importantly, experts agreed that, in general, spinal X-rays should not be repeated more frequently than every 2 years unless clearly indicated in individual cases. This recommendation is based on the experience from clinical studies [36, 37].

Strength of recommendation: 9.4±0.2.

Non-pharmacological treatment

- The cornerstone of non-pharmacological treatment of patients with AS is patient education and regular exercise.
- Home exercises are effective. Physical therapy with supervised exercises, land or water based, individually or in a group, should be preferred as these are more effective than home exercises.
- Patient associations and self-help groups may be useful.

For comparison, the old recommendation was: non-pharmacological treatment of AS should include patient education and regular exercise. Individual and group physical therapy should be considered. Patient associations and self-help groups may be useful.

Comment

This bullet point was changed according to the SLR and the recent Cochrane review on the subject [38], and was supported by the view of an experienced physiotherapist (HD) and the participating patients.

Strength of recommendation: 8.8±0.4.

Extra-articular manifestations and comorbidities

- The frequently observed extra-articular manifestations, e.g. psoriasis, uveitis, and chronic inflammatory bowel disease (IBD), should be managed in collaboration with the respective specialists.
- Rheumatologists should be aware of an increased risk of cardiovascular disease and osteoporosis.

Comment

This is a new bullet point, with agreement being achieved after considerable discussion. The main argument was that extraarticular manifestations are rather frequent in AS and

the entire spectrum of spondyloarthritis [39], and that they constitute a frequent challenge in management that clearly requires cooperation between specialities.
On the other hand, there are frequent comorbidities that require the attention of the managing rheumatologist. These include low bone mineral density, osteoporotic fractures [40, 41] and cardiovascular diseases [42, 43], which have been reported to occur in AS and spondyloarthritis at an increased rate compared with the general population.
The rheumatologist is encouraged to identify patients at risk and the potential additional risk factors. At this time, it is difficult to make a clear-cut recommendation on the management of osteopaenia and osteoporosis for patients with AS in the absence of any studies on the subject.
Regarding the management of cardiovascular risk there are recent EULAR recommendations that propose an annual risk assessment related to national guidelines [44]. Although this is mainly intended for patients with RA, these same guidelines should also be considered for patients with AS and psoriatic arthritis. Rheumatologists are referred to local guidelines for the management of cardiovascular risk and, if no local guidelines are available, the management should be carried out according to the systematic coronary risk evaluation (SCORE) function [45] (for overview see Cooney *et al*.) [46]. In addition to appropriate cardiovascular risk management, aggressive suppression of the infl ammatory process is recommended to lower the cardiovascular risk further.
Strength of recommendation: 9.0±0.3.

Non-steroidal anti-inflammatory drugs

- Non-steroidal anti-inflammatory drugs (NSAID), including Coxibs, are recommended as first-line drug treatment for AS patients with pain and stiffness.
- Continuous treatment with NSAID is preferred for patients with persistently active, symptomatic disease.
- Cardiovascular, gastrointestinal and renal risks should be taken into account when prescribing NSAID.

For comparison, the old recommendation was: NSAID are recommended as first-line drug treatment for patients with AS with pain and stiffness. In those with increased gastrointestinal risk, non-selective NSAID plus a gastroprotective agent, or a selective COX-2 inhibitor with or without a gastroprotective agent could be used.

Comment

This bullet point was subject to some minor modifications but the significance of the statement remains unchanged.
The main issues are still that NSAID are recommended as the first-line drug therapy, that NSAID are recommended to be taken continuously in active patients, and that NSAID are considered relatively safe in the population of patients with AS, although the cardiovascular, gastrointestinal and renal risks may be somewhat increased in this population.
The main challenges are that it is unclear whether a cut-off such as a Bath ankylosing spondylitis disease activity index of 4 is valuable in classifying patients as responders or nonresponders with regard to NSAID therapy, whether NSAID should be taken continuously regardless of symptoms by all (even asymptomatic) patients to prevent new bone formation, whether long-term NSAID therapy is safe, whether patients at risk can be readily identified, and how this should be done in clinical practice.
There is evidence that NSAID are efficacious for the relief of pain and stiffness in patients with AS [47] for both short-term and prolonged periods of treatment [48, 49]. The efficacy is, at least partly, dose related [48]. There seems to be no effect on spinal inflammation as assessed by MRI in one small study [50], but continuous therapy may be superior in the prevention of

new bone formation [51]. Coxibs may be safe for short-term therapy even in patients with IBD [52]. One recent step forward for clinical trials in AS has been the ASAS proposal on how information on NSAID intake should be collected in studies [53].
Strength of recommendation: 9.3±0.3.

Analgesics

- Analgesics, such as paracetamol and opioid-(like) drugs, might be considered for residual pain after previously recommended treatments have failed, are contraindicated, and/or poorly tolerated.

Comment

This bullet point has remained unchanged. This topic has been the source of frequent discussion and there are experts who have proposed eliminating this bullet point, but the majority still felt that inclusion of this bullet point was necessary because it was important to draw attention to the possibility that not all back pain in AS may derive from spinal inflammation.
Strength of recommendation: 8.0±0.5.

Glucocorticoids

- Glucocorticoid injections directed to the local site of musculoskeletal inflammation may be considered.
- The use of systemic glucocorticoids for axial disease is not supported by evidence.

Comment

This bullet point has remained unchanged. There have been no new studies and the available literature is still scarce.
Strength of recommendation: 8.9±0.4.

Disease modifying antirheumatic drugs

- There is no evidence for the efficacy of disease-modifying antirheumatic drugs (DMARD), including sulfasalazine and methotrexate, for the treatment of axial disease.
- Sulfasalazine may be considered in patients with peripheral arthritis.

Comment

This bullet point has remained unchanged. After the last Cochrane review [54] there were two new studies on sulfasalazine [55, 56], but the experts did not find that these provided sufficient new information to change this bullet point. The results of the first study, which was performed mainly in patients who had early spondyloarthritis, are conflicting [55], whereas in the head-to-head trial against etanercept there was no placebo group [56]. Overall, a marginal positive effect of sulfasalazine with a rather limited effect size in AS cannot be excluded. Therefore, no strong recommendation can be given to support its use but the rheumatologist may decide on a trial of sulfasalazine for a limited period, usually not more than 4 months, after which further benefit is unlikely. The majority of the studies suggest some efficacy of sulfasalazine in patients with peripheral spondyloarthritis and in the prevention of anterior uveitis. However, etanercept was more efficacious in the active comparator trial [56]. Finally, there is clearly no reason other than economic to recommend the obligatory use of a conventional DMARD in AS before anti-TNF therapy.
The data on methotrexate are still very limited and no positive recommendation can be given on an evidence basis. After the last Cochrane review [57] there was one new open-label

study with a high dose of methotrexate given subcutaneously [58], which again demonstrated no effect on patients with axial disease.
Most rheumatologists will try methotrexate in patients with predominant peripheral spondyloarthritis, but no evidence based recommendation can presently support this. Strength of recommendation: 9.4±0.2.

Anti-TNF therapy

- Anti-TNF therapy should be given to patients with persistently high disease activity despite conventional treatments according to the ASAS recommendations.
- There is no evidence to support the obligatory use of DMARD before or concomitant with anti-TNF therapy in patients with axial disease.
- There is no evidence to support a difference in efficacy of the various TNF inhibitors on the axial and articular/entheseal disease manifestations; but in the presence of IBD a difference in gastrointestinal efficacy needs to be taken into account.
- Switching to a second TNF blocker might be beneficial especially in patients with loss of response.
- There is no evidence to support the use of biological agents other than TNF inhibitors in AS.

For comparison, the old recommendation was: anti-TNF treatment should be given to patients with persistently high disease activity despite conventional treatments according to the ASAS recommendations. There is no evidence to support the obligatory use of DMARD before, or concomitant with, anti-TNF treatment in patients with axial disease.

Comment

This recommendation was substantially changed - based on extensive discussions related to the literature review, as the vast majority of new studies published in the past 5 years were related to anti-TNF therapy. The statement is of course strongly related to the recent update of the ASAS recommendations on anti-TNF therapy in AS. [15]
Since the last systematic review [11] there were many new studies. In addition to infliximab and etanercept, adalimumab [59] and golimumab [60] have also been approved. There are substantial data on patient-reported outcomes [61]. There is evidence that patients with advanced disease [62] also have some benefit, but patients with early [63] and very early [64] disease seem to have even more benefit. The highest remission rate reported is up to 50% after 16 weeks [64] in patients with inflammatory back pain [65] of less than 3 years (mean 15 months) and sacroiliitis on MRI but not on radiographs. Of note, the majority of the patients in these trials did not fulfill the modified New York criteria for AS.
The retention rate of patients with AS after 1 year of anti-TNF therapy was better than for patients with RA in a large registry [66]. There is evidence that the efficacy of anti-TNF therapy lasts over several years [67–69].
Spinal inflammation, as assessed by MRI, improves substantially after anti-TNF therapy [70]. Radiographic progression (mainly new bone formation) does not seem to be inhibited by anti-TNF therapy [71], but there is also no evidence that syndesmophyte formation is accelerated.
The major new aspect of the updated recommendations is the differential effect of anti-TNF therapy when available drugs have similar efficacy on musculoskeletal manifestations but differential efficacy in clinically symptomatic IBD [72]. Here the monoclonal antibodies work better than the fusion protein (infliximab is approved for both Crohn's disease (CD) and ulcerative colitis, adalimumab for CD, no data yet available for golimumab). The differences regarding acute anterior uveitis are less evident [73]. The presence or absence of psoriasis does not seem to make a difference as regards efficacy on musculoskeletal symptoms [74].

There is evidence that anti-TNF agents improve the signs and symptoms of peripheral arthritis and enthesitis [75, 76].
Furthermore, a recommendation for switching is included for the first time since several studies have suggested high success rates [77–79]. It was discussed that antibody formation [80, 81] may be involved in the phenomenon of loss of response (secondary nonresponse) and that such patients seem to have an even higher potential for response to a second TNF blocker than primary non-responders.
The statement that there is no evidence for the efficacy of other biological therapies in AS is also new. It is based on two studies evaluating rituximab and abatacept, which both failed to show convincing response rates in patients who had failed TNF blockers [82, 83]. The response rate to rituximab in TNF-naïve AS patients deserves further study [82].
Some experts stressed the importance of exercise and regular physiotherapy in patients with AS under treatment with TNF blockers, but the literature on this topic is still scarce [84].
Strength of recommendation: 9.4±0.2.

Surgery

- Total hip arthroplasty should be considered in patients with refractory pain or disability and radiographic evidence of structural damage, independent of age.
- Spinal corrective osteotomy may be considered in patients with severe disabling deformity.
- In patients with AS and an acute vertebral fracture a spinal surgeon should be consulted.

For comparison, the old recommendation was: total hip arthroplasty should be considered in patients with refractory pain or disability and radiographic evidence of structural damage, independent of age. Spinal surgery, such as corrective osteotomy and stabilisation procedures, may be of value in selected patients.

Comment

This bullet point was modified based on discussions with the two orthopaedic surgeons in the expert committee. The significance of hip involvement has been confirmed by a recent multinational study [85, 86]. A statement related to that problem therefore remains important, and the first sentence on total hip arthroplasty remains unchanged. Cement is only rarely used in young patients [87], and heterotopic ossification does not seem to be a problem in patients with AS [88]. The recommendation on spinal surgery was intensively discussed and a more detailed statement agreed on.
The second statement addresses an elective surgical procedure in the spine, which was shown to be beneficial for many patients with advanced AS and hyperkyphosis who have lost their horizontal vision ability. This technically challenging operation, which is only performed in experienced centres and is not available in some countries, leads to at least the partial correction of kyphosis. Triangular pieces of bone are removed from selected vertebral bodies (pedicle subtraction osteotomy) before the spine is re-stabilised by metallic bars and screws.
The third statement addresses spinal fractures that may lead to instability of the spine. These are often but not always rather acute clinical situations, which may or may not be associated with neurological symptoms [89–94]. In addition, as mechanical stress may prevent discovertebral spinal lesions from fusion and lead to the development of pseudarthrosis, a spinal surgeon should at least be consulted in patients with symptomatic discovertebral lesions [95].
This has been regarded as so important that an extra bullet point, no 11, was added. Strength of recommendation: 9.2±0.3.

Box 1: First update of the ASAS/EULAR recommendations for the management of AS.

The overarching principles of the management of patients with AS are:

- AS is a potentially severe disease with diverse manifestations, usually requiring multidisciplinary treatment coordinated by the rheumatologist.
- The primary goal of treating the patient with AS is to maximise long term health-related quality of life through control of symptoms and inflammation, prevention of progressive structural damage, preservation/normalisation of function and social participation.
- Treatment of AS should aim at the best care and must be based on a shared decision between the patient and the rheumatologist.
- The optimal management of patients with AS requires a combination of non-pharmacological and pharmacological treatment modalities.

1. General treatment

The treatment of patients with AS should be tailored according to:

- The current manifestations of the disease (axial, peripheral, entheseal, extra-articular symptoms and signs).
- The level of current symptoms, clinical findings, and prognostic indicators.
- The general clinical status (age, gender, comorbidity, concomitant medications, psychosocial factors).

2. Disease monitoring

The disease monitoring of patients with AS should include:

- Patient history (e.g. questionnaires)
- Clinical parameters
- Laboratory tests
- Imaging
  - All according to the clinical presentation as well as the ASAS core set

The frequency of monitoring should be decided on an individual basis depending on:

- Course of symptoms
- Severity
- Treatment

3. Non-pharmacological treatment

- The cornerstone of non-pharmacological treatment of patients with AS is patient education and regular exercise.
- Home exercises are effective. Physical therapy with supervised exercises, land or water based, individually or in a group, should be preferred as these are more effective than home exercises.
- Patient associations and self-help groups may be useful.

4. Extra-articular manifestations and comorbidities

- The frequently observed extra-articular manifestations, for example, psoriasis, uveitis and IBD, should be managed in collaboration with the respective specialists.
- Rheumatologists should be aware of the increased risk of cardiovascular disease and osteoporosis.

5. Non-steroidal anti-inflammatory drugs

- NSAID, including Coxibs, are recommended as first-line drug treatment for AS patients with pain and stiffness.
- Continuous treatment with NSAID is preferred for patients with persistently active, symptomatic disease.
- Cardiovascular, gastrointestinal and renal risks should be taken into account when prescribing NSAID.

6. Analgesics
- Analgesics, such as paracetamol and opioid (like) drugs, might be considered for residual pain after previously recommended treatments have failed, are contraindicated, and/or poorly tolerated.

7. Glucocorticoids
- Corticosteroid injections directed to the local site of musculoskeletal inflammation may be considered.
- The use of systemic glucocorticoids for axial disease is not supported by evidence.

8. Disease-modifying antirheumatic drugs
- There is no evidence for the efficacy of DMARD, including sulfasalazine and methotrexate, for the treatment of axial disease.
- Sulfasalazine may be considered in patients with peripheral arthritis.

9. Anti-TNF therapy
- Anti-TNF therapy should be given to patients with persistently high disease activity despite conventional treatments according to the ASAS recommendations.
- There is no evidence to support the obligatory use of DMARD before or concomitant with anti-TNF therapy in patients with axial disease.
- There is no evidence to support a difference in efficacy of the various TNF inhibitors on the axial and articular/entheseal disease manifestations; but in the presence of IBD a difference in gastrointestinal efficacy needs to be taken into account.
- Switching to a second TNF blocker might be beneficial especially in patients with loss of response.
- There is no evidence to support the use of biological agents other than TNF inhibitors in AS.

10. Surgery
- Total hip arthroplasty should be considered in patients with refractory pain or disability and radiographic evidence of structural damage, independent of age.
- Spinal corrective osteotomy may be considered in patients with severe disabling deformity.
- In patients with AS and an acute vertebral fracture a spinal surgeon should be consulted.

11. Changes in the disease course
- If a significant change in the course of the disease occurs, other causes than inflammation, such as a spinal fracture, should be considered and appropriate evaluation, including imaging, should be performed.

Changes in the disease course

- If a significant change in the course of the disease occurs, causes other than inflammation, such as a spinal fracture, should be considered and appropriate evaluation, including imaging, should be performed.

Comment

This is a new recommendation. The major point is that changes in the course of the disease should be carefully evaluated and MRI performed - especially in situations in which the nature of back pain changes. An experienced spinal surgeon may need to be consulted. It seems important to stress that not all AS patients with spinal fractures have neurological symptoms (and not all need to be operated on).
There are other important differential diagnoses such as spinal infections.
Strength of recommendation: 9.0±0.3.

Discussion

The ASAS/EULAR recommendations were successfully updated. The introduction of overarching principles led to some changes, e.g. one bullet point and one sentence was moved to this section. There are now 11 bullet points including two new points: one for extra-articular manifestations and one for changes in the clinical course of AS.

A patient version of the recommendations will be developed. We encourage translation of these recommendations into various languages in a collaboration between rheumatologists and patients. After presentation at the EULAR 2010 meeting in Rome and publication in the EULAR journal, individual countries can now take on dissemination.

The collaboration between ASAS and EULAR has again been very successful and should be continued for the next update that may be renamed according to the new classification criteria for axial spondyloarthritis. There will be a need for further discussion as to whether the new criteria for peripheral spondyloarthritis [96] should give rise to separate recommendations for these patients.

Although it was decided that these recommendations concentrate on AS, the authors are well aware that the treatment of patients with non-radiographic axial spondyloarthritis is also a very important topic. There are now data of clinical trials available that address this question in a controlled manner [63, 64]. They provide evidence that anti-TNF agents work in early disease in at least the same but probably in an even superior way.

The original publication has already set a standard for the management of patients with AS. As we feel that this update has even improved the original set we are confident that these recommendations will be useful for patients and healthcare workers, including rheumatologists and other physicians treating patients with AS, as well as physiotherapists.

References

1. Zochling J, van der Heijde D, Burgos-Vargas R, *et al.* ASAS/EULAR recommendations for the management of ankylosing spondylitis. Ann Rheum Dis 2006;65:442-52.
2. Zhang W, Doherty M, Bardin T, *et al.*; for the EULAR Standing Committee for International Clinical Studies Including Therapeutics. EULAR evidence based recommendations for gout. Part II: Management. Report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). Ann Rheum Dis 2006;65:1312-24.
3. Mukhtyar C, Guillevin L, Cid M, *et al.*; for the European Vasculitis Study Group. EULAR recommendations for the management of primary small and medium vessel vasculitis. Ann Rheum Dis 2009; 68:310-17.
4. Landewé R, Günther K, Lukas C, *et al.* EULAR/EFORT recommendations for the diagnosis and initial management of patients with acute or recent onset swelling of the knee. Ann Rheum Dis 2010;69:12-19.
5. Kowal-Bielecka O, Landewé R, Avouac J, *et al.*; EUSTAR Co-Authors. EULAR recommendations for the treatment of systemic sclerosis: a report from the EULAR Scleroderma Trials and Research group (EUSTAR). Ann Rheum Dis 2009;68:620-8.
6. Smolen J, Landewé R, Breedveld F, *et al.* EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. Ann Rheum Dis 2010;69:964-75.
7. Dougados M, Betteridge N, Burmester G, *et al.* EULAR standardised operating procedures for the elaboration, evaluation, dissemination, and implementation of recommendations endorsed by the EULAR standing committees. Ann Rheum Dis 2004;63:1172-6.
8. The AGREE Collaboration. Appraisal and Guidelines for Research and Evaluation (AGREE) Instrument, 2001. www.agreecollaboration.org (accessed Feb 2010).
9. Moher D, Tetzlaff J, Tricco A, *et al.* Epidemiology and reporting characteristics of systematic reviews. PLoS Med 2007;4:e78.
10. Shojania K, Sampson M, Ansari M, *et al.* How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med 2007;147:224-33.
11. Zochling J, van der Heijde D, Dougados M, *et al.* Current evidence for the management of ankylosing spondylitis: a systematic literature review for the ASAS/ EULAR management recommendations in ankylosing spondylitis. Ann Rheum Dis 2006;65:423-32.
12. van der Heijde D, Bellamy N, Calin A, *et al.* Preliminary core sets for endpoints in ankylosing spondylitis. Assessments in Ankylosing Spondylitis Working Group. J Rheumatol 1997;24:2225-9.
13. Braun J, Pham T, Sieper J, *et al.*; ASAS Working Group. International ASAS consensus statement for the use of anti-tumour necrosis factor agents in patients with ankylosing spondylitis. Ann Rheum Dis 2003;62:817-24.
14. Braun J, Davis J, Dougados M, *et al.* First update of the international ASAS consensus statement for the use of anti-TNF agents in patients with ankylosing spondylitis. Ann Rheum Dis 2006;65:316-20.
15. van der Heijde D, Sieper J, Maksymowych W, *et al.* (Second) Update of the ASAS recommendations on the use of TNF-blockers in ankylosing spondylitis [abstract]. Arthritis and Rheumatism 2009;60(Suppl10):1790.
16. Hewlett S. Patients and clinicians have different perspectives on outcomes in arthritis. J Rheumatol 2003;30:877-9.
17. Kiltz U, van der Heijde D, Mielants H, *et al.* ASAS/EULAR recommendations for the management of ankylosing spondylitis: the patient version. Ann Rheum Dis 2009;68:1381-6.
18. Gossec L, Dougados M, Phillips C, *et al.* Dissemination and evaluation of the ASAS/EULAR recommendations for the management of ankylosing spondylitis: results of a study among 1507 rheumatologists. Ann Rheum Dis 2008;67:782-8.
19. Braun J, Sieper J. [Dissemination and evaluation of ASAS/EULAR recommendations for the management of ankylosing spondylitis in Europe]. Z Rheumatol 2009;68:417-19.
20. Braun J, Zochling J, Märker-Hermann E, *et al.* [Recommendations for the management of ankylosing spodylitis after ASAS/EULAR. Evaluation in the German language area]. Z Rheumatol 2006;65:728-42.

21. Burgos-Vargas R, Abud-Mendoza C, Díaz-Jouanen E, *et al.* [Treatment guidelines for ankylosing spondylitis and its effect on Mexican rheumatology]. Gac Med Mex 2009;145:41-9.
22. Kiltz U, Feldtkeller E, Braun J. [Patient evaluation of the German version of the ASAS/EULAR recommendations for the management of ankylosing spondylitis]. Z Rheumatol 2008;67:694-700;702.
23. Kiltz U, Feldtkeller E, Braun J. German patient version of the ASAS/EULAR recommendations for the management of ankylosing spondylitis. Z Rheumatol 2010;69:171-4;176-9.
24. Braun J, Sieper J. Ankylosing spondylitis. Lancet 2007;369:1379-90.
25. Rudwaleit M, Landewé 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.
26. Rudwaleit M, van der Heijde D, Landewé 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.
27. Rudwaleit M, Jurik A, Hermann K, *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.
28. 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(Suppl2):ii1-44.
29. Sampaio-Barros P, Bortoluzzo A, Conde R, *et al.* Undifferentiated spondyloarthritis: a longterm followup. J Rheumatol 2010;37:1195-9.
30. Bennett A, McGonagle D, O'Connor P, *et al.* Severity of baseline magnetic resonance imaging—evident sacroiliitis and HLA-B27 status in early infl ammatory back pain predict radiographically evident ankylosing spondylitis at eight years. Arthritis and Rheumatism 2008;58:3413-18.
31. Fouque-Aubert A, Jette-Paulin L, Combescure C, *et al.* Serious infections in patients with ankylosing spondylitis with and without TNF blockers: a systematic review and meta-analysis of randomised placebo-controlled trials. Ann Rheum Dis 2010;69:1756-61.
32. Furst D, Keystone E, Fleischmann R, *et al.* Updated consensus statement on biological agents for the treatment of rheumatic diseases, 2009. Ann Rheum Dis 2010;69(Suppl 1):i2-29.
33. Zink A, Braun J, Listing J, *et al.* Disability and handicap in rheumatoid arthritis and ankylosing spondylitis—results from the German rheumatological database. German Collaborative Arthritis Centers. J Rheumatol 2000;27:613-22.
34. WHO. International classification of functioning, disability and health. Geneva: World Health Organization, 2001.
35. Boonen A, Braun J, van der Horst-Bruinsma I, *et al.* The ASAS/WHO ICF core sets for ankylosing spondylitis: how to classify the impact of AS on functioning and health. Ann Rheum Dis 2010;69:102-7.
36. Spoorenberg A, de Vlam K, van der Linden S, *et al.* Radiological scoring methods in ankylosing spondylitis. Reliability and change over 1 and 2 years. J Rheumatol 2004;31:125-32.
37. van der Heijde D, Landewé R, van der Linden S. How should treatment effect on spinal radiographic progression in patients with ankylosing spondylitis be measured? Arthritis and Rheumatism 2005;52:1979-85.
38. Dagfinrud H, Kvien T, Hagen K. Physiotherapy interventions for ankylosing spondylitis. Cochrane Database Syst Rev 2008;1:CD002822.
39. vander Cruyssen B, Ribbens C, Boonen A, *et al.* The epidemiology of ankylosing spondylitis and the commencement of anti-TNF therapy in daily rheumatology practice. Ann Rheum Dis 2007;66:1072-7.
40. Karberg K, Zochling J, Sieper J, *et al.* Bone loss is detected more frequently in patients with ankylosing spondylitis with syndesmophytes. J Rheumatol 2005;32:1290-8.
41. Vosse D, Landewé R, van der Heijde D, *et al.* Ankylosing spondylitis and the risk of fracture: results from a large primary care-based nested case–control study. Ann Rheum Dis 2009;68:1839-42.

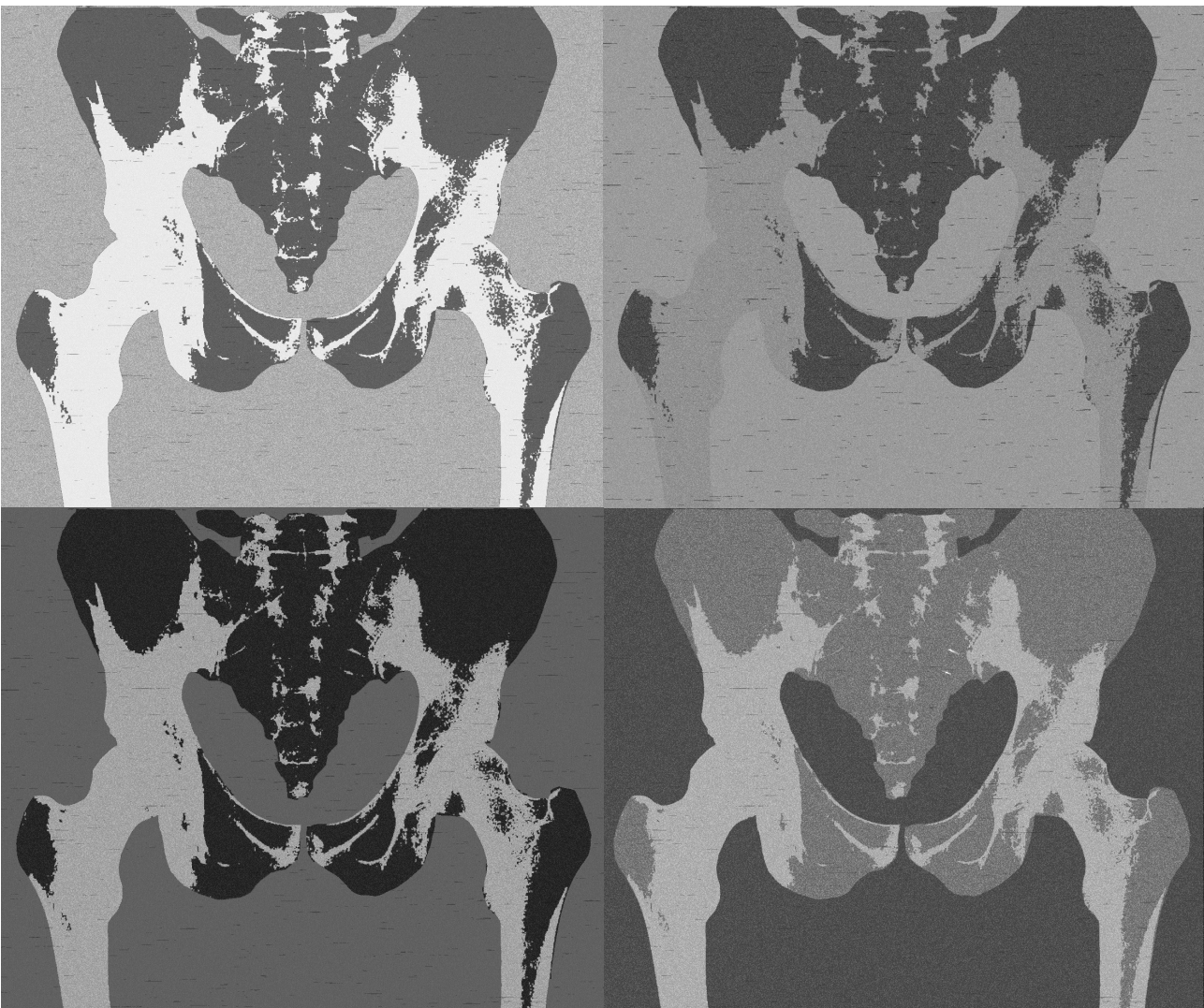
42. Peters M, Visman I, Nielen M, *et al.* Ankylosing spondylitis: a risk factor for myocardial infarction? Ann Rheum Dis 2010;69:579-81.
43. Peters M, van Eijk I, Smulders Y, *et al.* Signs of accelerated preclinical atherosclerosis in patients with ankylosing spondylitis. J Rheumatol 2010;37:161-6.
44. Peters M, Symmons D, McCarey D, *et al.* EULAR evidence-based recommendations for cardiovascular risk management in patients with rheumatoid arthritis and other forms of infl ammatory arthritis. Ann Rheum Dis 2010;69:325-31.
45. Conroy R, Pyörälä K, Fitzgerald A, *et al.* Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. Eur Heart J 2003;24:987-1003.
46. Cooney MT, Dudina AL, Graham IM. Value and limitations of existing scores for the assessment of cardiovascular risk: a review for clinicians. J Am Coll Cardiol 2009;54:1209-27.
47. Escalas C, Trijau S, Dougados M. Evaluation of the treatment effect of NSAIDs/TNF blockers according to different domains in ankylosing spondylitis: results of a metaanalysis. Rheumatology (Oxford) 2010;49:1317-25.
48. van der Heijde D, Baraf H, Ramos-Remus C, *et al.* Evaluation of the efficacy of etoricoxib in ankylosing spondylitis: results of a fifty-two-week, randomized, controlled study. Arthritis and Rheumatism 2005;52:1205-15.
49. Sieper J, Klopsch T, Richter M, *et al.* Comparison of two different dosages of celecoxib with diclofenac for the treatment of active ankylosing spondylitis: results of a 12-week randomised, double-blind, controlled study. Ann Rheum Dis 2008;67:323-9.
50. Jarrett SJ, Sivera F, Cawkwell LS, *et al.* MRI and clinical findings in patients with ankylosing spondylitis eligible for anti-tumour necrosis factor therapy after a short course of etoricoxib. Ann Rheum Dis 2009;68:1466-9.
51. Wanders A, Heijde D, Landewé R, *et al.* Nonsteroidal antiinflammatory drugs reduce radiographic progression in patients with ankylosing spondylitis: a randomized clinical trial. Arthritis and Rheumatism 2005;52:1756-65.
52. El Miedany Y, Youssef S, Ahmed I, *et al.* The gastrointestinal safety and effect on disease activity of etoricoxib, a selective cox-2 inhibitor in inflammatory bowel diseases. Am J Gastroenterol 2006;101:311-17.
53. Dougados M, Simon P, Braun J, *et al.* ASAS recommendations for collecting, analysing and reporting NSAID intake in clinical trials/ epidemiological studies in axial spondyloarthritis. Ann Rheum Dis 2011;70:249-51.
54. Chen J, Liu C. Sulfasalazine for ankylosing spondylitis. Cochrane Database Syst Rev 2005;2:CD004800.
55. Braun J, Zochling J, Baraliakos X, *et al.* Efficacy of sulfasalazine in patients with inflammatory back pain due to undifferentiated spondyloarthritis and early ankylosing spondylitis: a multicentre randomised controlled trial. Ann Rheum Dis 2006;65:1147-53.
56. Braun J, van der Horst-Bruinsma I, Huang F, *et al.* Clinical efficacy and safety of etanercept versus sulfasalazine in ankylosing spondylitis patients: a randomized, double-blind study (ASCEND Trial). Arthritis Rheum 2011;63:1543-51.
57. Chen J, Liu C, Lin J. Methotrexate for ankylosing spondylitis. Cochrane Database Syst Rev 2006;4:CD004524.
58. Haibel H, Brandt H, Song I, *et al.* No efficacy of subcutaneous methotrexate in active ankylosing spondylitis: a 16-week open-label trial. Ann Rheum Dis 2007;66:419-21.
59. van der Heijde D, Kivitz A, Schiff M, *et al.*; ATLAS Study Group. Efficacy and safety of adalimumab in patients with ankylosing spondylitis: results of a multicenter, randomized, double-blind, placebo-controlled trial. Arthritis and Rheumatism 2006;54:2136-46.
60. Inman R, Davis J, Heijde D, *et al.* Efficacy and safety of golimumab in patients with ankylosing spondylitis: results of a randomized, double-blind, placebo-controlled, phase III trial. Arthritis and Rheumatism 2008;58:3402-12.
61. Braun J, McHugh N, Singh A, *et al.* Improvement in patient-reported outcomes for patients with ankylosing spondylitis treated with etanercept 50 mg once-weekly and 25 mg twice-weekly. Rheumatology (Oxford) 2007;46:999-1004.
62. van der Heijde D, Pangan A, Schiff M, *et al.*; ATLAS Study Group. Adalimumab effectively reduces the signs and symptoms of active ankylosing spondylitis in patients with total spinal ankylosis. Ann Rheum Dis 2008;67:1218-21.

63. Haibel H, Rudwaleit M, Listing J, *et al.* Efficacy of adalimumab in the treatment of axial spondylarthritis without radiographically defined sacroiliitis: results of a twelveweek randomized, double-blind, placebo-controlled trial followed by an open-label extension up to week fi fty-two. Arthritis and Rheumatism 2008;58:1981-91.
64. Barkham N, Keen H, Coates L, *et al.* Clinical and imaging efficacy of infliximab in HLA-B27-Positive patients with magnetic resonance imaging-determined early sacroiliitis. Arthritis and Rheumatism 2009;60:946-54.
65. Braun J, Inman R. Clinical significance of inflammatory back pain for diagnosis and screening of patients with axial spondyloarthritis. Ann Rheum Dis 2010;69:1264-8.
66. Heiberg M, Koldingsnes W, Mikkelsen K, *et al.* The comparative one-year performance of anti-tumor necrosis factor alpha drugs in patients with rheumatoid arthritis, psoriatic arthritis, and ankylosing spondylitis: results from a longitudinal, observational, multicenter study. Arthritis and Rheumatism 2008;59:234-40.
67. van der Heijde D, Schiff M, Sieper J, *et al.*; ATLAS Study Group. Adalimumab effectiveness for the treatment of ankylosing spondylitis is maintained for up to 2 years: long-term results from the ATLAS trial. Ann Rheum Dis 2009;68:922-9.
68. Davis J, van der Heijde D, Braun J, *et al.* Effi cacy and safety of up to 192 weeks of etanercept therapy in patients with ankylosing spondylitis. Ann Rheum Dis 2008;67:346-52.
69. Braun J, Baraliakos X, Listing J, *et al.* Persistent clinical effi cacy and safety of anti-tumour necrosis factor alpha therapy with infliximab in patients with ankylosing spondylitis over 5 years: evidence for different types of response. Ann Rheum Dis 2008;67:340-5.
70. Braun J, Landewé R, Hermann K, *et al.*; ASSERT Study Group. Major reduction in spinal inflammation in patients with ankylosing spondylitis after treatment with infliximab: results of a multicenter, randomized, double-blind, placebo-controlled magnetic resonance imaging study. Arthritis and Rheumatism 2006;54:1646-52.
71. van der Heijde D, Landewé R, Einstein S, *et al.* Radiographic progression of ankylosing spondylitis after up to two years of treatment with etanercept. Arthritis and Rheumatism 2008;58:1324-31.
72. Braun J, Baraliakos X, Listing J, *et al.* Differences in the incidence of fl ares or new onset of infl ammatory bowel diseases in patients with ankylosing spondylitis exposed to therapy with anti-tumor necrosis factor alpha agents. Arthritis and Rheumatism 2007;57:639-47.
73. Braun J, Baraliakos X, Listing J, *et al.* Decreased incidence of anterior uveitis in patients with ankylosing spondylitis treated with the anti-tumor necrosis factor agents infliximab and etanercept. Arthritis and Rheumatism 2005;52:2447-51.
74. Braun J, Rudwaleit M, Kary S, *et al.* Clinical manifestations and responsiveness to adalimumab are similar in patients with ankylosing spondylitis with and without concomitant psoriasis. Rheumatology (Oxford) 2010;49:1578-89.
75. Rudwaleit M, van den Bosch F, Kron M, *et al.* Effectiveness and safety of adalimumab in patients with ankylosing spondylitis or psoriatic arthritis and history of anti-tumor necrosis factor therapy. Arthritis Res Ther 2010; 12: R117.
76. Dougados M, Combe B, Braun J, *et al.* A randomised, multicentre, double-blind, placebo-controlled trial of etanercept in adults with refractory heel enthesitis in spondyloarthritis: the HEEL trial. Ann Rheum Dis 2010;69:1430-5.
77. Cantini F, Niccoli L, Benucci M, *et al.* Switching from infl iximab to once-weekly administration of 50 mg etanercept in resistant or intolerant patients with ankylosing spondylitis: results of a fifty-four-week study. Arthritis and Rheumatism 2006;55:812-16.
78. Coates L, Cawkwell L, Ng N, *et al.* Real life experience confirms sustained response to long-term biologics and switching in ankylosing spondylitis. Rheumatology (Oxford) 2008;47:897-900.
79. Pradeep D, Keat A, Gaffney K, *et al.* Switching anti-TNF therapy in ankylosing spondylitis. Rheumatology (Oxford) 2008;47:1726-7.

80. de Vries M, Wolbink G, Stapel S, *et al.* Decreased clinical response to infliximab in ankylosing spondylitis is correlated with anti-infliximab formation. Ann Rheum Dis 2007;66:1252-4.
81. de Vries M, Brouwer E, van der Horst-Bruinsma I, *et al.* Decreased clinical response to adalimumab in ankylosing spondylitis is associated with antibody formation. Ann Rheum Dis 2009;68:1787-8.
82. Song I, Heldmann F, Rudwaleit M, *et al.* Different response to rituximab in tumor necrosis factor blocker-naive patients with active ankylosing spondylitis and in patients in whom tumor necrosis factor blockers have failed: a twenty-four-week clinical trial. Arthritis and Rheumatism 2010;62:1290-7.
83. Song I, Heldmann H, Rudwaleit M, *et al.* Treatment of active ankylosing spondylitis with abatacept – an open label 24-week study. Ann Rheum Dis 2010;69(Suppl3):60.
84. Colina M, Ciancio G, Garavini R, *et al.* Combination treatment with etanercept and an intensive spa rehabilitation program in active ankylosing spondylitis. Int J Immunopathol Pharmacol 2009;22:1125-9.
85. van der Cruyssen B, Muñoz-Gomariz E, Font P, *et al.*; ASPECT-REGISPONSER-RESPONDIA working group. Hip involvement in ankylosing spondylitis: epidemiology and risk factors associated with hip replacement surgery. Rheumatology (Oxford) 2010;49:73-81.
86. Baraliakos X, Braun J. Hip involvement in ankylosing spondylitis: what is the verdict? Rheumatology (Oxford) 2010;49:3-4.
87. Sochart D, Porter M. Long-term results of total hip replacement in young patients who had ankylosing spondylitis. Eighteen to thirty year results with survivorship analysis. J Bone Joint Surg Am 1997;79:1181-9.
88. Bhan S, Eachempati K, Malhotra R. Primary cementless total hip arthroplasty for bony ankylosis in patients with ankylosing spondylitis. J Arthroplasty 2008;23:859-66.
89. Sapkas G, Kateros K, Papadakis S, *et al.* Surgical outcome after spinal fractures in patients with ankylosing spondylitis. BMC Musculoskelet Disord 2009;10:96.
90. Fordham S, Lloyd G. Clinical management of injured patients with ankylosing spondylitis. BMJ 2009;339:b2568.
91. Whang P, Goldberg G, Lawrence J, *et al.* The management of spinal injuries in patients with ankylosing spondylitis or diffuse idiopathic skeletal hyperostosis: a comparison of treatment methods and clinical outcomes. J Spinal Disord Tech 2009;22:77-85.
92. Westerveld L, Verlaan J, Oner F. Spinal fractures in patients with ankylosing spinal disorders: a systematic review of the literature on treatment, neurological status and complications. Eur Spine J 2009;18:145-56.
93. van Royen B, Scheerder F, Jansen E, *et al.* ASKyphoplan: a program for deformity planning in ankylosing spondylitis. Eur Spine J 2007;16:1445-9.
94. Kiaer T, Gehrchen M. Transpedicular closed wedge osteotomy in ankylosing spondylitis: results of surgical treatment and prospective outcome analysis. Eur Spine J 2010;19:57-64.
95. Bron J, de Vries M, Snieders M, *et al.* Discovertebral (Andersson) lesions of the spine in ankylosing spondylitis revisited. Clin Rheumatol 2009;28:883-92.
96. Rudwaleit M, van der Heijde D, Landewé R, *et al.* The Assessment of SpondyloArthritis International Society classifi cation criteria for peripheral spondyloarthritis and for spondyloarthritis in general. Ann Rheum Dis 2011;70:25-31.

12

Summary and general discussion



Summary and general discussion

Although our understanding of the expressions and mechanisms of SpA is far from complete, major advances have been made over the past decades. The need for earlier diagnosis was addressed by including MRI in the diagnostic process as well as by putting more emphasis on HLA-B27 in the diagnostic process. The development of new classification criteria, with the major objective to ensure the identification of non-radiographic types of SpA, is a further aid in classifying patients in early phases of the disease. The recently by ASAS proposed classification criteria sets for SpA included MRI for the first time [1, 2]. With the help of MRI, patients can be classified earlier in the disease stage since MRI can provide evidence of inflammatory sacroiliitis. Furthermore, in 2006 ASAS/EULAR published management recommendations to provide guidance for monitoring and treatment of AS patients. However, the performance of the ASAS classification criteria, including (the role of) imaging (both MRI and radiographs), required thorough evaluation. Moreover, as the number of clinical trials and publications on AS therapy has steadily increased over the first decade of the millennium, the ASAS/EULAR recommendations for the management of AS needed an update. The performed studies described in this thesis have contributed to the continuously developing field of SpA. The main results and conclusions are summarized and discussed in this chapter.

Part I: Early recognition and classification criteria of spondyloarthritis

The performance of the various developed classification criteria for SpA (mNY, Amor, ESSG, ASAS axial SpA, ASAS peripheral SpA, and the CASPAR criteria (for PsA only)) [1-6] was tested in a group of patients with predominantly axial complaints included in the SPondyloArthritis Caught Early (SPACE)-cohort (**chapter 2**) and in a group of patients with predominantly peripheral complaints (**chapter 4**) included in the Leiden Early Arthritis Clinic (EAC)-cohort [7]. In the EAC-cohort, patients diagnosed with PsA and peripheral SpA by the treating rheumatologist, and a control group matched on age, gender and symptom duration were selected and studied. In this group of PsA and peripheral SpA patients, the ASAS peripheral SpA criteria and CASPAR criteria had substantial overlap by classifying the same patients. In the peripheral SpA subgroup, the ASAS peripheral SpA criteria outperformed all other classification criteria, and in the PsA subgroup, the CASPAR criteria outperformed all other criteria. The diagnosis by the rheumatologist served as external standard. However, this setting is neither representative for all peripheral SpA patients nor for the whole concept of SpA since the EAC-cohort does not include patients with dactylitis or enthesitis or patients with predominantly axial complaints. This is reflected in the modest sensitivities of all classification criteria found in this analysis. Thus, in daily practice rheumatologists include a wider group of patients in the SpA disease spectrum than defined in the classification criteria, thereby underscoring the fact that classification criteria are not diagnostic criteria. Yet, it is very reassuring that the specificities of all criteria sets are in accordance with the reported specificities in the original validation cohorts [1-6]. This is especially of importance for the ASAS peripheral SpA criteria as these criteria are quite new and there was a fear that they might be insufficiently specific. Moreover, the results of **chapter 4** were recently confirmed in the ESPERANZA-cohort, which is developed in Spain to facilitate early diagnoses of SpA by creating early SpA units with standard operating procedures, and to improve the knowledge and practical skills of GPs and specialists in the field of SpA. Patients with complaints (IBP or asymmetrical arthritis (preferably in the lower limbs) or spinal/joint pain in combination with one other SpA-feature) were included. Thereby, a slightly different population is created than the population in the EAC-cohort and the ASAS validation cohort, but the sensitivity and specificity of the ASAS peripheral SpA criteria were very similar (sensitivity of 56%, specificity of 85%) [8].

We also tested the existing classification criteria in patients with predominantly axial complaints included in the SPACE-cohort. The SPACE-cohort is an ongoing prospective longitudinal observational cohort including patients with back pain ≥3 months but ≤2 years with the onset of symptoms <45 years, which is extensively described in **chapter 2**. At baseline, all patients in the SPACE-cohort undergo a diagnostic work-up including imaging (MRI and radiographs of the SI-joints and spine), laboratory tests (including HLA-B27 testing) and physical examination and history taking. In **chapter 2** we showed that at baseline almost 60% of the patients in the SPACE-cohort fulfilled any classification criteria set (mNY, Amor, ESSG, ASAS axial SpA criteria) [2-5]; almost 40% fulfilled the ASAS axial SpA criteria [2]. The ASAS axial SpA criteria outperformed all other sets (including modifications of Amor and ESSG by adding MRI) with respect to sensitivity, specificity, positive and negative likelihood ratio (LR+ and LR-). Again, the diagnosis by the rheumatologist served as external standard. As the percentage of patients with axSpA according to the ASAS axSpA criteria in the SPACE-cohort appears to be high, it could be argued that our observed prevalence of axSpA is influenced by referral bias; e.g. that due to increased awareness among referring physicians about the SPACE cohort over time, patients from areas other than the Leiden area are referred to the LUMC or that only patients with a high suspicion of axSpA are referred.

However, the percentage of axSpA among all referred patients over the years was similar, and the percentage of referrals from outside the Leiden area was also similar over time. Moreover, 33 of the 157 patients (21.0%) included at baseline had none or only one less specific SpA feature. This indicates, but does not prove, that there is no referral bias, thereby suggesting that the observed prevalence of axSpA could be generalized to primary care (**chapter 2**). Moreover, very similar results regarding the performance of the ASAS axial SpA criteria are recently found in another study, also including patients with back pain ≥3 months, onset <45 years, conducted by Moltó *et al.* among rheumatologists working in office-based and hospital-based practices in France [9].

As there are indications that not all rheumatologists as well as registration authorities (U.S. Food and Drug Administration (FDA)) [10, 11] do appreciate the validity of the clinical arm of the ASAS axial SpA criteria as similar to the imaging arm, we compared patients fulfilling the clinical arm to patients fulfilling the imaging arm (both patients with radiographic sacroiliitis and patients with inflammatory sacroiliitis on MRI) in **chapter 2**. Noteworthy, patients in both arms were remarkably similar with respect to the presence of most SpA-features and level of disease activity (BASDAI and ASDAS). Similar comparisons were made in the DESIR-cohort and in the ABILITY-1 trial. The latter is a randomized controlled trial performed in patients with nr-axSpA (fulfilling the ASAS axSpA criteria but not fulfilling the mNY criteria) to evaluate the efficacy and safety of adalimumab in those patients [12]. The results of these studies were comparable to the results we found in the SPACE-cohort that patients fulfilling the clinical arm and patients fulfilling the imaging arm are very similar [12, 13]. Nevertheless, the level of confidence about the diagnosis indicated by the rheumatologist in the SPACE-cohort is lower in patients fulfilling the clinical arm than in patients fulfilling the imaging arm (**chapter 2**). This seems to indicate that rheumatologists heavily base their diagnosis on positive imaging. This concept is confirmed in the study by Moltó *et al*. mentioned above, pointing out that (radiographic) sacroiliitis has the highest LR+ on the diagnosis of SpA according to rheumatologists [9].

For further analyses in the SPACE-cohort, patients are classified as no-SpA or axial SpA based on the ASAS axSpA criteria (the best performing classification criteria) as classification criteria are by definition exactly defined and therefore reproducible while the diagnosis of the rheumatologist is not. Of the axSpA patients, approximately 80% have IBP, and the other way around, 56.7% of the patients without SpA have IBP, thereby showing that - at least in the SPACE-cohort - IBP is not a very useful feature to discriminate between axSpA patients and patients without SpA (**chapter 2**). These results are in line with results reported

before [14, 15]. Nevertheless, IBP history taking is cheap and can be useful in the diagnostic process in combination with other SpA-features (the more clinical items suggestive of SpA, the more likely the diagnosis) [16]. Moreover, some items of IBP, like age at onset ≤35 years, are more important than other items by having a higher calculated LR+ on the diagnosis of axial SpA [17]. However, the interpretation of (items) of IBP differ from person to person, reflected in low agreement on whether a patient is suffering from IBP or not [14, 18].

Nonetheless, in the previously developed ESSG criteria but also in the diagnostic Berlin algorithm, IBP was used as (one of the) entry criteria [4, 19]. As the Berlin algorithm is the only available diagnostic tool, this challenged us to propose two modifications of the Berlin algorithm, presented in **chapter 3**. In modification 1, the first step of the algorithm – fulfillment of the ASAS IBP criteria – was slightly changed. Instead of ≥4 out of the 5 criteria, patients need to fulfill ≥3 out 5 criteria. Modification 2 slightly changed the structure and the set of SpA-features by deleting IBP as obligatory entry criterion and adding it as a SpA-feature. This resulted in three entry groups based on the requirement of ≥4, 2-3 and 0-1 SpA-features. The performance of the (modified) algorithms, was tested against fulfillment of the ASAS axSpA criteria, the disease probability based on the likelihood ratio product [19, 20] and the diagnosis by the rheumatologist as external standard due to the lack of a true gold standard. Modification 1 resulted in a major increase in sensitivity, at the cost of little specificity. With modification 2, the number of missed axSpA diagnoses by the algorithm even further decreased. Additional adjustments that might improve the diagnostic algorithm even more could be contemplated; rheumatologists could consider performing an MRI in HLA-B27 negative patients who do have 2-3 other SpA-features, especially male patients [21, 22]. As it is now stated in the algorithm, those patients leave the algorithm ('consider other diagnosis'), however, if the MRI is positive they would fulfill the ASAS axSpA criteria. Moreover, as this algorithm is developed to guide rheumatologists in the diagnostic process, it will often be applied in patients with relatively short symptom duration. The usefulness of performing conventional radiographs of the SI-joints as a first step could therefore be challenged, reflected by the high number of negative radiographs in both the SPACE-cohort and the ASAS-cohort. Nevertheless, modification 2 of the algorithm might be a useful tool for rheumatologists in daily practice. The Dutch Society for Rheumatology (Nederlandse Vereniging voor Reumatologie (NVR)) recently published new guidelines for the diagnosis and treatment of axSpA in which the modified algorithm is included [23].

Although the tools available to rheumatologists for classification and diagnostic purposes improved a lot over the last years, one of the unmet needs is the referral of the most appropriate patients by physicians and health professionals to the rheumatologist. Referring physicians have only limited knowledge of manifestations belonging to SpA [24], illustrated by the poor agreement regarding the evaluation of IBP by referring physicians and rheumatologists (κ=0.04 to κ=0.20) [14, 18]. Therefore, it is a challenge for referring physicians to recognize patients with a suspicion of having SpA who should be referred to a rheumatologist. Several (complex) referral strategies have been developed in order to early identify patients with possible SpA. All strategies performed well in research settings with instructed GPs, yielding 24% to 45% SpA patients [14, 25-28]. However, the limited knowledge of referring physicians poses a challenge on successful implementation of referral strategies in the common daily primary care setting [24]. Hence, it might be useful to consider an easy referral structure instead of complex strategies, by just referring patients with back pain ≥3 months with the onset of symptoms <45 years, like the eligibility criteria of the ASAS axSpA criteria. At least in the SPACE-cohort - with the additional restriction of a maximum symptom duration of ≤2 years - and in the study by Moltó *at al*. these criteria yield high percentages of SpA (41.4% and 35.1%, respectively) (**chapter 2**) [9]. However, it remains to be seen whether this would be successful in other centers or in countries with other healthcare systems, and therefore more research is needed.

Another important related question is which patients are erroneously not referred to the rheumatologist by applying these referral strategies [14, 25-28]. We tried to answer this question by testing the performance of various referral strategies in the SPACE-cohort, even though the SPACE-cohort might not be ideal to sort out this question as patients already had been referred [29, 30]. Remarkably, almost all non-referred patients that fulfilled the ASAS axSpA criteria had (radiographic) sacroiliitis [29, 30].

Although major improvements have been achieved in classifying and diagnosing patients, more research and education is needed, starting with warranted improvements in referring the right patients to the rheumatologist. This could be achieved by (further) educating referring physicians about SpA-features. Furthermore, both referring physicians and rheumatologists could be trained in acknowledging that axSpA should not be ruled out if IBP is absent. If performance of the current strategies appears to be insufficient - even after educating referring physicians - eventually the development of new referral strategies could be considered. These strategies offer the possibility of referring patients with positive imaging without the necessity of performing imaging in daily practice (a proxy for positive imaging is needed), since with the current strategies precisely imaging positive patients are often not referred.

Moreover, only diagnostic tool for rheumatologists, the ASAS modified Berlin algorithm, is not validated in other cohorts than the two validation cohorts yet. In addition to this necessary validation, other improvements of the algorithm could be considered. For example, it could be investigated whether it is useful to make a distinction in the group of HLA-B27 negative patients with 0-1 SpA-features to decide on whether or not performing MRI. Patients with 1 feature could fulfill the ASAS axSpA criteria if imaging is positive while the patients without any SpA-features will never fulfill the ASAS axSpA criteria. Moreover, performing conventional radiographs of the SI-joints after medical history taking and physical examination - instead of before - could be considered.

Moreover, long-term follow-up studies are required in order to study outcomes in patients fulfilling the clinical arm and patients fulfilling the imaging arm of the ASAS axSpA criteria, and to compare the long-term outcomes of patients in both groups. This will help in understanding the disease and in concluding on whether it was the right decision to include the clinical arm in the ASAS axSpA criteria. The clinical arm was included as it gave the best balanced sensitivity and specificity compared to including the imaging arm only [2]. Moreover, more knowledge will assist in considering potential adjustments of the ASAS axSpA criteria like weighting the various SpA-features (as in the ASAS modified Berlin algorithm) since some SpA-features, such as a 'positive family history for SpA' and 'HLA-B27 positivity' are more strongly correlated than others.

Part II: The role of imaging in the early diagnosis of spondyloarthritis

MRI has proven its usefulness in diagnosing and classifying SpA patients over the last years. Nevertheless, the newly acquired prominent role of MRI as well as the role of conventional radiographs are currently under debate. The discussion regarding the role of radiographic sacroiliitis has its origin in the knowledge that it is challenging to recognize radiographic sacroiliitis. The undulating articular surface and the complex anatomy of the SI-joints by a 2-dimensional imaging technique can result in misinterpretations [31, 32]. Recently, the poor reliability of evaluating conventional radiographs was emphasized in post-hoc analyses on the data of the ABILITY-1 and RAPID-axSpA pivotal clinical trials for the registration of TNF-blockers in patients with nr-axSpA [12,33-35]. The analysis in ABILITY-I pointed out that 37% of the patients classified as nr-axSpA by local readers were reclassified as AS by fulfilling the mNY criteria according to central readers [12, 34]. In the RAPID-axSpA trial a similar analysis resulted in reclassification of 36% of the patients; 26% of the nr-axSpA patients according to local readers were reclassified as fulfilling the mNY criteria according to central readers,

and 10% of the mNY fulfilling patients by local readers were reclassified as nr-axSpA based on the central reading [33, 35].

We performed similar analyses in the DESIR-cohort (**chapter 5** and **chapter 6**), in which imaging evaluations by both local readers and central readers are available. In daily practice the diagnosis of AS is based on the judgment of local radiologists and/or rheumatologists, while in cohorts and clinical trials the radiographs are usually judged by one or more trained central readers. The comparison described in **chapter 5** on the presence/absence of radiographic sacroiliitis by local readers versus central readers revealed that the agreement was only moderate (κ=0.55). The local readers primarily overrated sacroiliitis in comparison with central readers as external standard, resulting in an unacceptably high percentage (41.5%) of false-positive diagnoses of AS in daily practice. Only a small minority of patients with a classification of AS according to central readers is not recognized in daily clinical practice (7.5%). Even in patients with bilateral involvement and patients with at least one fused SI-joint major discrepancies are seen between local readers and central readers. Moreover, interreader agreement between the two central readers was also only moderate (κ=0.54), indicating that evaluating SI-joints on radiographs is very difficult and that training does not improve the agreement substantially. But where misclassification by local readers almost exclusively consisted of overclassification of positive cases, the disagreement between the two central readers was more balanced in two directions.

Thus far, there is no data on this aspect for sacroiliitis on MRI, except for the data of the DESIR-cohort presented in **chapter 6**. In contrast to the moderate agreement regarding radiographic sacroiliitis, agreement regarding sacroiliitis on MRI between the two central readers as well as between the local readers and the central readers is substantial (κ=0.73 and κ=0.70, respectively). Potentially 163/582 patients (28.0%) in whom the MRI and/or radiograph reading was different between the local readers and central readers, could have been classified differently according to the ASAS axSpA criteria. Yet, only 46/582 patients (7.9%) were classified differently. These results point out the robustness of the ASAS axSpA classification criteria to differences in reading of the images. This is mainly due to the clinical arm; patients fulfilling the clinical arm will always fulfil the clinical arm, regardless of the imaging results, as HLA-B27 status will not change.

Given the only moderate reliability in conventional radiograph reading and the substantial reliability in MRI reading, some experts in the field argue that the option of leaving out conventional radiographs completely and conducting MRI only should be considered. As the current definition of a positive MRI is based on the presence of inflammatory lesions only, this discussion is becoming even more interesting if structural lesions (erosions, ankylosis and sclerosis) on MRI are taken into account as well. To be able to evaluate the potential additive value of adding structural lesions to the definition of a positive MRI, it is first important to know whether structural lesions can be detected reliably on MRI. Therefore, the performance of MRI in the detection of structural lesions in the SI-joints was tested against the conventional radiographs as gold standard. Agreement varied from κ=0.11 to κ=0.15 for erosions, from κ=0.16 to κ=0.46 for sclerosis, and from κ=0.08 to κ=0.85 for ankylosis (partial or total) in the GESPIC-cohort and SPACE-cohort [36, 37]. Overall, agreement is poor; more erosions and less sclerotic lesions are detected on MRI compared to conventional radiographs [36, 37]. These comparisons should be extended and repeated with an alternative external standard such as CT.

The incorporation of various combinations of structural lesions and fatty depositions on the definition of a positive MRI is presently being investigated in both the SPACE-cohort and the DESIR-cohort, but should be investigated and validated in other cohorts as well, especially in patients with longer symptom duration. Before we can conclude on the role of conventional radiographs and structural lesions on MRI, we will have to wait for the results of these studies. In the meantime, we investigated in the DESIR-cohort the possible consequences

of this proposal of leaving out conventional radiographs and using the current definition of a positive MRI based on inflammatory lesions only (**chapter 6**). Taking into account the complete ASAS axSpA criteria including the clinical arm, this would result in only 11 to 14 missed patients (1.9-2.4%using either the local reading or the central reading) as those patients only fulfill the imaging arm by having radiographic sacroiliitis only (and not inflammatory sacroiliitis on MRI). However, it should be stated immediately that this is in an early cohort and this may be different in patients with more advanced disease.

MRI is also used to quantify inflammation in the SI-joints and spine, for example by using the SPARCC-score. We tested the metric properties of the SPARCC-score of the sacroiliac joints in the SPACE-cohort in **chapter 7.** We found out that a SPARCC-score of 2 as cut-off value is the best equivalent of the ASAS definition of a positive MRI. This cut-off value can be used (in clinical trials) to create a dichotomous MRI variable of potential prognostic interest. Additionally, we calculated smallest detectable changes (SDCs), which in this study were close enough to the proposed minimally important change (MIC) of 2.5 SPARCC-units to add credibility to a cut-off level of 2.5 units representing a true change rather than only measurement error. A large proportion of the SPARCC-score changes seen in the patients in the SPACE-cohort could be considered as noise as these changes are smaller than the calculated SDCs (62.9% and 45% (3 months, campaign 1 and 2) and 39.1% (1 year in campaign 2)). Surprisingly, true (>SDC) changes in SPARCC-score over time (both increases and decreases) were frequently observed while patients are on stable treatment. This observation strongly suggests that MRI-activity fluctuates over time.

Other intriguing matter is the hypothesis of inflammation being the inciting cause of structural lesions including ossification. Prospective long-term follow-up data is necessary to study the possible relationship between inflammation and structural lesions in more detail [38, 39], in both the SI-joints and spine. To get more insight in why some patients do develop spinal lesions and others do not, information is needed on the prevalence of spinal lesions (inflammatory and structural), especially in patients without (radiographic) sacroiliitis. Several studies already addressed this question, but the results are inconclusive as the prevalence of spinal lesions varies with age and disease duration [40-45]. In addition, it is questioned whether the results of MRI of the spine should be taken into account in the diagnostic and/or classification process, which could be of particular interest in patients without (radiographic) sacroiliitis. To address these research questions as well as other questions, long-term follow-up data is currently being collected in, among others, the SPACE-cohort and DESIR-cohort.

Part III: Treatment of spondyloarthritis and ankylosing spondylitis

ASAS together with EULAR published recommendations for the management of AS in 2006. As the number of clinical trials and publications on AS therapy is increasing, ASAS developed an update of the first recommendations for the management of AS. These recommendations are described in **chapter 11**, based on systematic literature reviews (**chapter 9** and **chapter 10**) as recommended by the EULAR standard operating procedures for management recommendations [46]. In the title, ASAS has restricted the recommendations to AS since the evidence from trials in axSpA patients is currently limited and this was an update of the previous recommendations on AS (and not axSpA). Nevertheless, the project group unanimously agreed that these recommendations could equally be applied to patients with axSpA. First because AS is part of the total group of axSpA, and second because all available data indicated that efficacy was at least as good in patients with nr-axSpA as in patients with AS. And indeed, this has been confirmed in all trials that have been published since. As described in **chapter 11**, ASAS recommend tailored treatment, taking into account all aspects of the disease including peripheral and extra-articular manifestations, level of disease activity, gender, and comorbidities etc. Disease should be

monitored regularly, according to the clinical presentation as well as the ASAS core set for assessment in clinical practice [47]. Treatment should consist of non-pharmacological and pharmacological treatment. Non-pharmacological treatment is the cornerstone, comprising patient education and regular exercise. The review on non-pharmacological treatment and non-biological drugs described in **chapter 9** pointed out that home exercises have positive effects on physical function (BASFI), patient reported disease activity (BASDAI), pain and spinal mobility, but that physical therapy with supervised exercises, either land or water based, either individually or in a group, are more effective than home exercises. This is adopted in the ASAS recommendations (**chapter 11**).

NSAIDs, including coxibs, are recommended as first line drug to relief pain and stiffness for both short-term and prolonged periods of treatment (**chapter 9** and **chapter 11**), and these should be taken in an anti-inflammatory dose [48]. ASAS recommends patients with persistently active, symptomatic disease to use NSAIDs continuously (**chapter 11**) as continuous therapy may be superior to on-demand therapy on the prevention of new bone formation [49]. After the update of the ASAS recommendations was published, a post-hoc analysis was conducted in this randomized trial comparing continuous to on-demand NSAID treatment, revealing that solely patients with elevated acute phase reactants will benefit from continuous treatment with NSAIDs [50]. Additionally, a study was recently performed in the German Spondyloarthritis Inception Cohort (GESPIC), investigating the influence of NSAIDs intake on radiographic spinal progression over 2 years in both AS and nr-axSpA patients. The results showed that a high NSAIDs intake is associated with retarded radiographic spinal progression in AS patients while this effect was less evident in nr-axSpA patients, probably due to a low grade of new bone formation in the spine at this stage [51].

Analgesics might be considered for residual pain after previously recommended treatments have failed, are contraindicated and/or poorly tolerated. Glucocorticoid injections directed to the local site of musculoskeletal inflammation may be considered, but systemic glucocorticoid use for axial disease is not supported by evidence. There is no evidence for the efficacy of DMARDs, including sulfasalazine and methotrexate, for the treatment of axial disease, however, sulfasalazine may be considered in patients with peripheral disease (**chapter 11**).

The results of the systematic literature review on biologics are described in **chapter 10**. TNF-α inhibitors should be given to patients with persistently high disease activity despite conventional treatments according to the ASAS recommendations [52]. Overall, all TNF-α inhibitors available for AS have proved to be effective on BASDAI, BASFI and BASMI, both in AS patients with established disease as well as in nr-axSpA patients, especially in patients with elevated CRP and/or inflammation on MRI. In the presence of IBD, a difference in gastrointestinal efficacy needs to be taken into account. There is no evidence to support the obligatory use of DMARDs before or concomitant with TNF-α inhibitors in patients with axial disease. Switching to a second TNF-α inhibitors might be beneficial, especially in patients that lost response (**chapter 10** and **chapter 11**).

Moreover, ASAS recommend considering total hip arthroplasty in case of refractory pain or disability and radiographic evidence of structural damage, independent of age. In case of severe disabling deformity, spinal corrective osteotomy may be considered (**chapter 9** and **chapter 11**).

Despite the fact that there is conclusive evidence that TNF-α inhibitors can dramatically improve disease activity (including inflammation on MRI) and functional capacity, its use is associated with high costs and is not suitable for all patients in terms of safety and increased risk of infections. Therefore, ASAS developed recommendations for the use of TNF-α inhibitors in 2003 [53]. Those recommendations have been updated twice already [52, 54]. Moreover, many countries developed national guidelines for the use of TNF-α inhibitors, either or not based on the ASAS recommendations. In **chapter 8**, the national guidelines of 23 countries worldwide

were compared, revealing that despite some differences, there is general consensus about the use of TNF-α inhibitors in AS. In addition, there is evidence that patients with nr-axSpA show good responses to TNF-α inhibitors, although the number of trials is limited and the sample sizes in those trials are relatively small [12, 33, 42, 43, 55]. Furthermore, only patients with high disease activity and/or elevated acute phase reactants and/or a positive MRI were included in those trials [11, 31, 40, 41, 51]. High disease activity as measured by the ASDAS, elevated acute phase reactants and a positive MRI are all identified as predictors for good response to treatment with TNF-α inhibitors [56]. Recently, adalimumab and certolizumab are approved by the European Commission for the treatment of adults with severe nr-axSpA, who have had an inadequate response to, or are intolerant to NSAIDs, but only in those nr-axSpA patients that show objective signs of inflammation by elevated CRP and/or MRI [57, 58]. Still, in many countries, patients with active, severe axSpA refractory to NSAIDs are only eligible for treatment with TNF-α inhibitors if imaging shows signs of sacroiliitis. However, the only difference between nr-axSpA and axSpA/AS is the presence of (radiographic) sacroilitis, which is proven to be challenging to reliably evaluate (**chapter 5** and **chapter 6**) [32]. Therefore, more research is warranted on the (long-term) effects of TNF-α inhibitors in early nr-axSpA patients, including patients without a positive MRI. In addition, long-term outcomes of head-to-head comparisons of different treatments are needed, focusing on the development of structural lesions. For example, head-to-head comparisons of the different available TNF-α inhibitors could be studied, with or without concomitant use of, amongst others, NSAIDs. In the same manner, TNF-α inhibitors could be compared directly to other types of drugs, like NSAIDs and bisphosphonates [59]. Moreover, the effect of various treatment strategies, like ASDAS-steered treatment, could be investigated. Furthermore, as there is evidence that a recent onset of symptoms is associated with higher response rates, and, importantly, a greater likelihood of a very good response, the existence of a 'window of opportunity' could be investigated [56, 60]. If a 'window of opportunity' exists, it would be favorable in achieving clinical and biological benefits as well as preventing structural damage, especially in recent-onset, active axSpA patients with MRI or laboratory signs of inflammation [60].

In conclusion, patients with SpA can be recognized earlier with the recent developments, thereby offering better treatment options and thus better outcomes. However, in order to further improve care of SpA patients, we cannot afford to stand still, but we have to keep on moving, just like patients with SpA.

References

1. Rudwaleit M, van der Heijde D, Landewé R, *et al*. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. Ann Rheum Dis 2011;70:25-31.
2. Rudwaleit M, van der Heijde D, Landewé 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.
3. Amor B, Dougados M, Mijiyawa M. Classification Criteria of Spondyloarthropathies. Revue du Rhumatisme 1990;57:85-9.
4. Dougados M, van der linden S, Juhlin R, *et al*. The European-Spondylarthropathy Study Group Preliminary Criteria for the Classification of Spondylarthropathy. Arthritis and Rheumatism 1991;34:1218-27.
5. van der Linden S, Valkenburg H, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361-8.
6. Taylor W, Gladman D, Helliwell P, *et al*. Classification criteria for psoriatic arthritis - Development of new criteria from a large international study. Arthritis and Rheumatism 2006;54:2665-73.
7. de Rooy D, van der Linden M, Knevel R, *et al*. Predicting arthritis outcomes--what can be learned from the Leiden Early Arthritis Clinic? Rheumatology (Oxford) 2011;50:93-100.
8. Tomero E, Mulero J, de *et al*. Performance of the Assessment of Spondyloarthritis International Society criteria for the classification of spondyloarthritis in early spondyloarthritis clinics participating in the ESPERANZA programme. Rheumatology (Oxford) 2014;53:353-60.
9. Molto A, Paternotte S, Comet D, *et al*. Performances of the Assessment of SpondyloArthritis International Society axial spondyloarthritis criteria for diagnostic and classification purposes in patients visiting a rheumatologist because of chronic back pain: results from a multicenter, cross-sectional study. Arthritis Care Res (Hoboken) 2013;65:1472-81.
10. http://www.medscape.com/viewarticle/808337. (4 June 2014, date last accessed).
11. http://www.fda.gov/AdvisoryCommittees/Calendar/ucm355358.htm. (4 June 2014, date last accessed).
12. Sieper J, van der Heijde D, Dougados M, *et al*. Efficacy and safety of adalimumab in patients with non-radiographic axial spondyloarthritis: results of a randomised placebo-controlled trial (ABILITY-1). Ann Rheum Dis 2013;72:815-22.
13. Molto A, Paternotte S, van der Heijde D, *et al*. Evaluation of the validity of the different arms of the ASAS set of criteria for axial spondyloarthritis and description of the different imaging abnormalities suggestive of spondyloarthritis: data from the DESIR cohort. Ann Rheum Dis [online 3-1-2014] doi: 10.1136/annrheumdis-2013-204262.
14. Poddubnyy D, Vahldiek J, Spiller I, *et al*. Evaluation of 2 screening strategies for early identification of patients with axial spondyloarthritis in primary care. J Rheumatol 2011;38:2452-60.
15. 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 (Oxford) 2009;48:404-9.
16. Burgos-Vargas R, Braun J. Inflammatory back pain. Rheum Dis Clin North Am 2012;38: 487-99.
17. Braun A, Saracbasi E, Grifka J, *et al*. Identifying patients with axial spondyloarthritis in primary care: how useful are items indicative of inflammatory back pain? Ann Rheum Dis 2011;70:1782-7.
18. Lopez-Gonzalez R, Hernandez-Sanz A, Almodovar-Gonzalez R, *et al*. Are spondyloarthropathies adequately referred from primary care to specialized care? Reumatol Clin 2013;9:90-3.
19. Rudwaleit M, van der Heijde D, Khan M, *et al*. How to diagnose axial spondyloarthritis early. Ann Rheum Dis 2004;63:535-43.
20. Rudwaleit M, Feldtkeller E, Sieper J. Easy assessment of axial spondyloarthritis (early ankylosing spondylitis) at the bedside. Ann Rheum Dis 2006;65:1251-2.

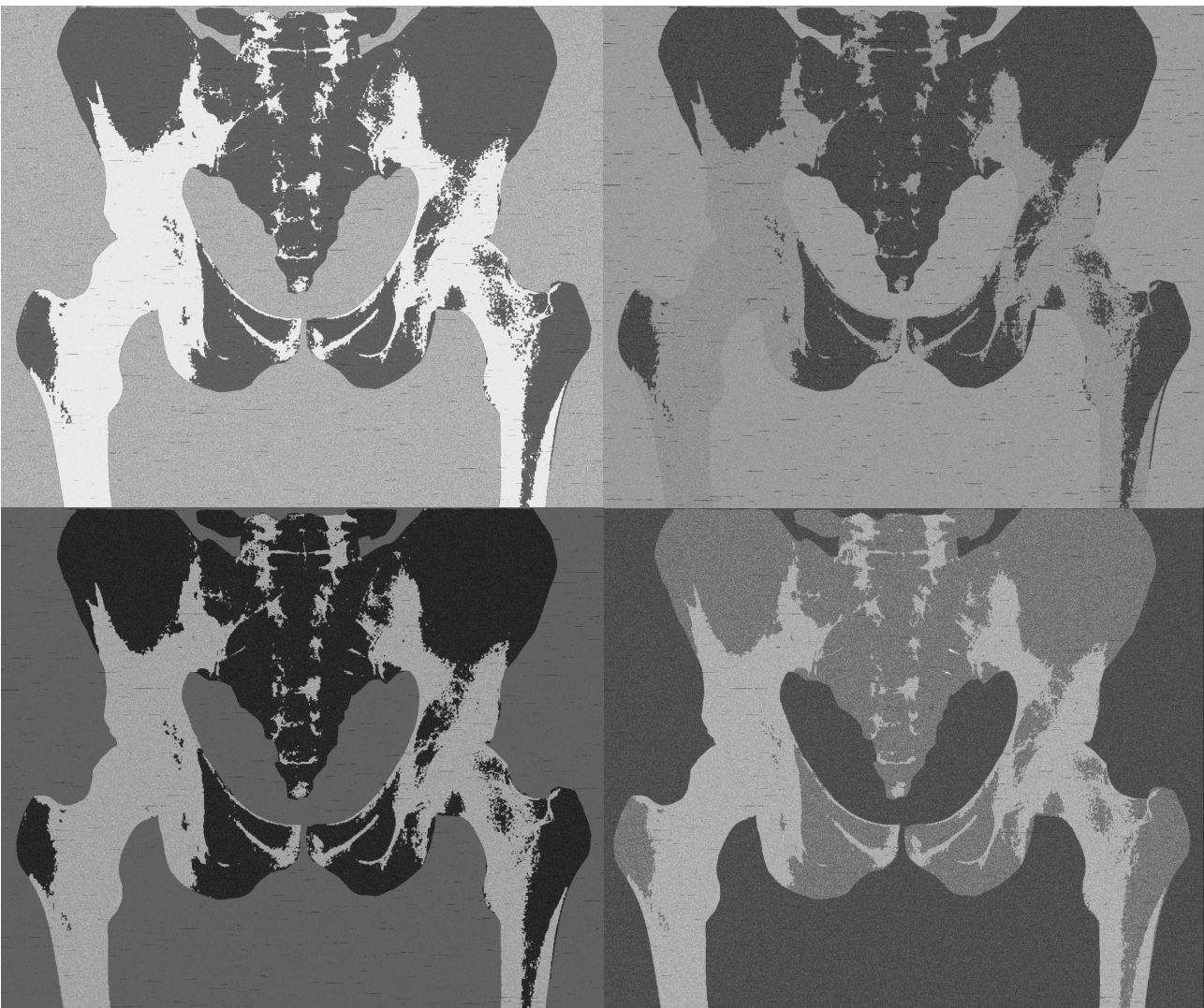
21. van Onna M, Jurik A, van der Heijde D, *et al*. HLA-B27 and gender independently determine the likelihood of a positive MRI of the sacroiliac joints in patients with early inflammatory back pain: a 2-year MRI follow-up study. Ann Rheum Dis 2011;70:1981-5.
22. van den Berg R, de Hooge M, Navarro-Compán V, *et al*. Is it useful to repeat an MRI in the diagnostic work-up for spondyloarthritis? Ann Rheum Dis 2012;71(Suppl 3):711.
23. van der Horst-Bruinsma I, Oostveen J, van Denderen J, *et al*. Richtlijn voor de diagnostiek en behandeling van Axiale Spondyloartritis. 31-1-2014.
24. van Onna M, Gorter S, van Meerendonk A, *et al*. General Practitioners' Perceptions of Their Ability to Identify and Refer Patients with Suspected Axial Spondyloarthritis: A Qualitative Study. J Rheumatol 2014;41:897-901.
25. Sieper J, Srinivasan S, Zamani O, *et al*. Comparison of two referral strategies for diagnosis of axial spondyloarthritis: the Recognising and Diagnosing Ankylosing Spondylitis Reliably (RADAR) study. Ann Rheum Dis 2013;72:1621-7.
26. Brandt H, Spiller I, Song I, *et al*. Performance of referral recommendations in patients with chronic back pain and suspected axial spondyloarthritis. Ann Rheum Dis 2007;66:1479-84.
27. Braun A, Gnann H, Saracbasi E, *et al*. Optimizing the identification of patients with axial spondyloarthritis in primary care--the case for a two-step strategy combining the most relevant clinical items with HLA-B27. Rheumatology (Oxford) 2013;52:1418-24.
28. van Hoeven L, Luime J, Han H, *et al*. Identifying axial spondyloarthritis in Dutch primary care patients, ages 20-45 years, with chronic low back pain. Arthritis Care Res (Hoboken) 2014;66:446-53.
29. Abawi O, van den Berg R, van der Heijde D, *et al*. Evaluation of the two-step referral strategy for axial spondyloarthritis in the SPondyloArthritis Caught Early cohort. Arthritis and Rheumatism 2013; 65(Suppl10):S1039.
30. Abawi O, van den Berg R, van der Heijde D, *et al*. Evaluation of referral models for axial spondyloarthritis in primary care in the SPondyloArthritis Caught Early cohort. Ann Rheum Dis 2014;73(Suppl2):426.
31. Spoorenberg A, de Vlam K, van der Linden S, *et al*. Radiological scoring methods in ankylosing spondylitis. Reliability and change over 1 and 2 years. J Rheumatol 2004;31:125-32.
32. van Tubergen A, Heuft-Dorenbosch L, Schulpen G, *et al*. Radiographic assessment of sacroiliitis by radiologists and rheumatologists: does training improve quality? Ann Rheum Dis 2003;62:519-25.
33. Landewé R, Braun J, Deodhar A, *et al*. Efficacy of certolizumab pegol on signs and symptoms of axial spondyloarthritis including ankylosing spondylitis: 24-week results of a double-blind randomised placebo-controlled Phase 3 study. Ann Rheum Dis 2014;73:39-47.
34. http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM366517.pdf . (24 December 2013, date last accessed).
35. http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM366519.pdf. (24 December 2013, date last accessed).
36. Poddubnyy D, Gaydukova I, Hermann K, *et al*. Magnetic resonance imaging compared to conventional radiographs for detection of chronic structural changes in sacroiliac joints in axial spondyloarthritis. J Rheumatol 2013;40:1557-65.
37. van den Berg R, de Hooge M, Navarro-Compán V, *et al*. Performance of MRI compared to conventional radiographs in the detection of structural lesions in the sacroiliac joints in patients with recent onset axial spondyloarthritis. Ann Rheum Dis 2013;72(Suppl3):658.
38. Maksymowych W, Morency N, Conner-Spady B, *et al*. Suppression of inflammation and effects on new bone formation in ankylosing spondylitis: evidence for a window of opportunity in disease modification. Ann Rheum Dis 2013;72:23-8.

39. Sieper J, Appel H, Braun J, *et al*. Critical appraisal of assessment of structural damage in ankylosing spondylitis: implications for treatment outcomes. Arthritis Rheum 2008;58:649-56.
40. de Hooge M, van den Berg R, Reijnierse M, *et al*. Prevalence of inflammatory lesions MRI-spine in patients with chronic back pain of <2 years duration included in the SPACE-cohort. Ann Rheum Dis 2013;72(Suppl3):660.
41. de Hooge M, van den Berg R, Navarro-Compán V, *et al*. Prevalence of structural lesions MRI-spine in patients with (possible) axial spondyloarthritis (axSpA) included in the SPACE-cohort. Ann Rheum Dis 2013;72(Suppl3):145.
42. Barkham N, Keen H, Coates L, *et al*. Clinical and imaging efficacy of infliximab in HLA-B27-Positive patients with magnetic resonance imaging-determined early sacroiliitis. Arthritis Rheum 2009;60:946-54.
43. Song I, Hermann K, Haibel H, *et al*. Effects of etanercept versus sulfasalazine in early axial spondyloarthritis on active inflammatory lesions as detected by whole-body MRI (ESTHER): a 48-week randomised controlled trial. Ann Rheum Dis 2011;70:590-6.
44. van der Heijde D, Sieper J, Maksymowych W, *et al*. Spinal inflammation in the absence of sacroiliac joint inflammation on magnetic resonance imaging in patients with active nonradiographic axial spondyloarthritis. Arthritis Rheumatol 2014;66:667-73.
45. Weber U, Zubler V, Zhao Z, *et al*. Does spinal MRI add incremental diagnostic value to MRI of the sacroiliac joints alone in patients with non-radiographic axial spondyloarthritis? Ann Rheum Dis [22-1-2014] doi: 10.1136/annrheumdis-2013-203887
46. Dougados M, Betteridge N, Burmester G, *et al*. EULAR standardised operating procedures for the elaboration, evaluation, dissemination, and implementation of recommendations endorsed by the EULAR standing committees. Ann Rheum Dis 2004;63:1172-6.
47. 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 (Suppl2):ii1-44.
48. Dougados M, Simon P, Braun J, *et al*. ASAS recommendations for collecting, analysing and reporting NSAID intake in clinical trials/epidemiological studies in axial spondyloarthritis. Ann Rheum Dis 2011;70:249-51.
49. Wanders A, van der Heijde D, Landewé R, *et al*. Nonsteroidal antiinflammatory drugs reduce radiographic progression in patients with ankylosing spondylitis: a randomized clinical trial. Arthritis Rheum 2005;52:1756-65.
50. Kroon F, Landewé R, Dougados M, *et al*. Continuous NSAID use reverts the effects of inflammation on radiographic progression in patients with ankylosing spondylitis. Ann Rheum Dis 2012;71:1623-9.
51. Poddubnyy D, Rudwaleit M, Haibel H, *et al*. Effect of non-steroidal anti-inflammatory drugs on radiographic spinal progression in patients with axial spondyloarthritis: results from the German Spondyloarthritis Inception Cohort. Ann Rheum Dis 2012;71:1616-22.
52. van der Heijde D, Sieper J, Maksymowych W, *et al*. 2010 Update of the international ASAS recommendations for the use of anti-TNF agents in patients with axial spondyloarthritis. Ann Rheum Dis 2011;70:905-8.
53. Braun J, Pham T, Sieper J, *et al*. International ASAS consensus statement for the use of anti-tumour necrosis factor agents in patients with ankylosing spondylitis. Ann Rheum Dis 2003;62:817-24.
54. Braun J, Davis J, Dougados M, *et al*. First update of the international ASAS consensus statement for the use of anti-TNF agents in patients with ankylosing spondylitis. Ann Rheum Dis 2006;65:316-20.
55. Haibel H, Rudwaleit M, Listing J, *et al*. Efficacy of adalimumab in the treatment of axial spondylarthritis without radiographically defined sacroiliitis. Arthritis and Rheumatism 2008;58:1981-91.
56. Arends S, van der veer E, V, Kallenberg C, *et al*. Baseline predictors of response to TNF-alpha blocking therapy in ankylosing spondylitis. Curr Opin Rheumatol 2012;24:290-8.
57. http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/000481/human_med_000822.jsp&mid=WC0b01ac058001d124. (30 May 2014, date last accessed).

58. http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion/human/001037/WC500150029.pdf . (5 June 2014, date last accessed).
59. Viapiana O, Gatti D, Idolazzi L, *et al*. Bisphosphonates vs infliximab in ankylosing spondylitis treatment. Rheumatology (Oxford) 2014;53:90-4.
60. Claudepierre P. Spondyloarthritis: A window of opportunity? Joint Bone Spine 2014;81:197-9.

13

Nederlandse samenvatting



Nederlandse samenvatting

Inleiding

Ankyloserende spondylitis (AS) of spondylitis ankylopoëtica, ook bekend als de ziekte van Bechterew, is een chronische reumatische ziekte, waarbij met name de gewrichten in de wervelkolom, en die van het bekken (de sacroiliacale (SI) gewrichten) ontstoken kunnen raken. Deze ontstekingen veroorzaken pijn en stijfheid en kunnen leiden tot ernstige verbeningen van de wervelkolom wat het dagelijks functioneren steeds meer belemmert. De klachten ontstaan meestal tussen het 20e en 40e levensjaar bij zowel mannen als vrouwen en rond de 45 jaar is meer dan 95% van de patiënten symptomatisch. Het ontstaan van AS wordt door meerdere factoren bepaald, waarbij genetische en omgevingsfactoren een rol spelen, maar hoe de ziekte precies ontstaat is nog onduidelijk. Het is in ieder geval bekend dat er een sterke associatie met HLA-B27 (een genetische factor) is; in 80-95% van de AS patiënten is HLA-B27 in het bloed gevonden. Hoe vaak AS voorkomt in een bevolkingsgroep hangt sterk samen met hoeveel mensen HLA-B27 in hun bloed hebben. In centraal Europa heeft ongeveer 0.1-0.7% AS (6-9% is positief getest voor HLA-B27) en ongeveer 6% van de Haida indianen in Noord-Amerika heeft AS (50% is HLA-B27 positief).

Naast klachten in de wervelkolom en het bekken, kunnen er ook klachten optreden in de andere gewrichten, zoals knieën, enkels en handen, die veroorzaakt worden door ontstekingen (perifere artritis, dactylitis (zogenaamde worstvormige teen of vinger) en enthesitis (ontsteking van de peesaanhechting aan het bot)) en er kunnen ook ontstekingen ontstaan buiten de gewrichten (extra-articulaire manifestaties), bijvoorbeeld in het oog (ontsteking van de voorste oogkamer (uveïtis anterior)), de huid (psoriasis) en de darmen, zoals de ziekte van Crohn en colitis ulcerosa (inflammatoire darm ziektes (IBD)). Wanneer informatie uit de verschillende onderzoeken bij elkaar wordt genomen, dan blijkt dat ongeveer 26% van de AS patiënten tenminste één keer een uveïtis anterior krijgt op enig moment tijdens het ziektebeloop. Ongeveer 9% van de patiënten heeft psoriasis en ongeveer 7% heeft IBD. Het aantal patiënten dat een perifere artritis krijgt tijdens het ziektebeloop variëert van 14% tot 47%, voor dactylitis varieert dit van 1.9 tot 3.1%, en voor enthesitis van 10% tot 49%.

Er bestaan geen geschikte criteria waarmee de diagnose bij een patiënt gesteld kan worden, maar er zijn wel classificatiecriteria voor AS die ontwikkeld zijn voor wetenschappelijk onderzoek: de gemodificeerde New York (mNY) criteria. Volgens de mNY criteria kan een patiënt geclassificeerd worden als AS wanneer de patiënt voldoet aan minstens een van de drie klinische criteria: 1) lage rugpijn en stijfheid >3 maanden die verbetert met bewegen maar niet met rust, en/of 2) beperkte beweeglijkheid van de onderrug in het sagittale en het frontale vlak (respectievelijk zijwaarts buigen en voor- achterwaarts buigen), en/ of 3) beperkte beweeglijkheid van de borstkas in vergelijking met normaalwaarden van leeftijdgenoten van hetzelfde geslacht. Daarnaast moeten duidelijke tekenen van ontsteking van de SI-gewricht(en) te zien zijn op röntgenfoto's van het bekken (radiografische sacroiliitis; tenminste graad 2 aan beide SI-gewrichten of graad 3-4 aan één SI-gewricht). Helaas duurt het vaak 6-8 jaar vanaf het moment van ontstaan van klachten totdat radiografische sacroiliitis op röntgenfoto's zichtbaar is. Op röntgenfoto's kunnen veranderingen aan het bot waargenomen worden, waaronder extra botvorming. Men gaat er vanuit dat de veranderingen die te zien zijn op röntgenfoto's eigenlijk consequenties van de ontsteking zijn. Echter, de ontstekingen zelf zijn op een röntgenfoto niet zichtbaar. Een röntgenfoto kan dus niet altijd gebruikt worden om een vroege vorm van AS op te sporen. Het onderliggende mechanisme, hoe ontsteking exact kan leiden tot excessieve nieuwvorming van bot, is nog niet volledig bekend.

AS is de meest uitgesproken vorm van spondyloartritis (SpA), een groep gerelateerde reumatische aandoeningen met gemeenschappelijke kenmerken, zoals de al eerder genoemde ontstekingen in de wervelkolom en het bekken. Maar niet bij elke patiënt met SpA staat de rugpijn en/of stijfheid van de rug op de voorgrond. Bij sommige patiënten met SpA voeren perifere of extra-articulaire klachten de boventoon. Daarom zijn er in de jaren '90 classificatiecriteria ontwikkeld, waarmee naast AS patiënten, ook patiënten met SpA geclassificeerd kunnen worden: de Amor en ESSG (European SpondyloArthropathy Study Group) criteria. Net als in de mNY criteria, zit in de Amor en ESSG criteria het criterium radiografische sacroiliitis. In tegenstelling tot de mNY criteria waar het hebben van radiografische sacroiliitis een voorwaarde is om aan de criteria te kunnen voldoen, wordt het in de Amor en ESSG criteria meegenomen als één van de SpA-kenmerken die mogelijk aanwezig kunnen zijn. Alle tot nu toe genoemde klachten en kenmerken (sacroiliitis, uveïtis anterior, IBD, HLA-B27 positiviteit, etc) worden tot de SpA-kenmerken gerekend. Voor het voldoen aan de ESSG criteria is de aanwezigheid van perifere artritis en/of inflammatoire rugpijn (inflammatory back pain (IBP)) vereist. IBP kenmerkt zich onder andere doordat de pijn voornamelijk in rust optreedt en juist afneemt door beweging. Patiënten die naast IBP en/of perifere artritis nog één ander SpA-kenmerk hebben kunnen volgens de ESSG criteria als SpA geclassificeerd worden. Bij de Amor criteria is er geen sprake van een bepaald vereist criterium, maar wordt het aantal aanwezige SpA-kenmerken geteld, welke allemaal één, twee of drie punten toegewezen hebben gekregen. Wanneer de patiënt minimaal 6 punten heeft, kan de patiënt worden geclassificeerd als SpA.

Recent heeft de ASAS, een internationale groep van experts op het gebied van SpA, ook twee sets classificatiecriteria voor SpA ontwikkeld. Één set kan toegepast worden in patiënten met voornamelijk klachten buiten de rug (perifere klachten), de andere set in patiënten met voornamelijk klachten in de rug (axiale klachten). In de axiale SpA criteria set speelt voor het eerst MRI van de SI-gewrichten een belangrijke rol. Ook ligt er veel nadruk op de aanwezigheid van HLA-B27. Dit, omdat gebleken is dat niet elke SpA patiënt radiografische sacroiliitis ontwikkelt en het belangrijkste doel bij de ontwikkeling van de ASAS axiale SpA criteria dus was om ook niet-radiografische SpA te kunnen identificeren. De axiale SpA criteria set is bedoeld voor patiënten met rugpijn die 3 maanden of langer bestaat en die ontstaan is vóór het 45e levensjaar. Als bij de patiënt sacroiliitis op röntgenfoto's of op MRI wordt vastgesteld en er nog minstens één ander SpA-kenmerk aanwezig is, of als er naast de aanwezigheid van HLA-B27 nog minimaal twee andere SpA-kenmerken aanwezig zijn, kan de patiënt worden geclassificeerd als axiale SpA.

De ASAS perifere criteria set kan worden toegepast in patiënten met perifere artritis, enthesitis en/of dactylitis. Als de patiënt daarnaast ook nog psoriasis, IBD, HLA-B27, (radiografische) sacroiliitis of een acute uveïtis anterior heeft, of wanneer de patiënt vier weken voorafgaand aan de perifere artritis/enthesitis/dactylitis diarree of een infectie van de baarmoederhals of van de plasbuis heeft gehad, dan kan de patiënt als perifere SpA worden geclassificeerd. Een patiënt met perifere artritis, enthesitis of dactylitis kan ook aan de ASAS perifere criteria voldoen als er tenminste nog twee andere van de volgende SpA-kenmerken aanwezig zijn: 1) perifere artritis (nu of in het verleden), 2) enthesitis (nu of in het verleden), 3) dactylitis (nu of in het verleden), 4) IBP of 5) een eerste- of tweedegraads familielid met SpA (AS of reactieve artritis) of SpA-kenmerken (psoriasis, acute uveïtis anterior en/of IBD).

Nu sacroiliitis op MRI zo'n prominente rol heeft gekregen in classificatiecriteria heeft de ASAS ook aanbevelingen voor het maken en interpreteren van een MRI ontwikkeld. Zonder alle technische details te bespreken, komt het er op neer dat ASAS aanbeveelt om altijd twee verschillende MRI opnamen te maken; één opname waarop onder andere actieve ontstekingen van het beenmerg (bone marrow edema; BME) goed te zien zijn (STIR opname)

en één opname waarop schade aan het bot, zoals erosies (gaatjes in gewrichtsoppervlak van de botten in de SI-gewrichten) en sclerose (het compacter worden van het bot) en excessieve nieuwvorming van bot goed te zien zijn (T1 TSE fatsat opname). De huidige definitie van een positieve MRI berust alleen op de aanwezigheid van BME laesies zoals te zien op een STIR opname en dus wordt de schade aan het bot niet in deze definitie meegenomen.
Het gebrek aan diagnostische criteria in combinatie met vele verschillende klachten die SpA kan geven (heterogene klinische presentatie), kan het soms lastig maken voor reumatologen om SpA te herkennen, zeker in de afwezigheid van (radiografische) sacroiliitis. Vaak wordt de diagnose gesteld op basis van de aanwezigheid van een combinatie van kenmerken; de bevindingen van de anamnese, het lichamelijk onderzoek, de beeldvorming en de laboratoriumtesten. De reumatoloog kan gebruik maken van een diagnostisch instrument, het Berlijn algoritme, waarin alle relevante SpA-kenmerken, zoals de extra-articulaire manifestaties en sacroiliitis op MRI, in acht worden genomen. Met behulp van het Berlijn algoritme kan voor elke individuele patiënt met IBP via verschillende stappen de waarschijnlijkheid op het hebben van SpA worden uitgerekend, wat kan helpen bij het stellen van de diagnose axiale SpA.
Met behulp van MRI, waarop sacroiliitis dus in een eerder stadium te herkennen is, is het nu ook mogelijk om de diagnose SpA eerder te stellen. Het is belangrijk dat de diagnose SpA zo vroeg mogelijk wordt gesteld, omdat een vroege diagnose onnodig verder onderzoek overbodig maakt en ook voorkomt dat de patiënt ongeschikte behandelingen krijgt. Daarnaast kan dan vroeg in de ziekte een geschikte behandeling worden gegeven, wat belangrijk is omdat patiënten met een korte ziektegeschiedenis beter lijken te reageren op behandeling dan patiënten met een lange(re) ziektegeschiedenis.
Momenteel rust de behandeling van SpA en AS op een niet-medicamenteuze en een medicamenteuze pijler. De niet-medicamenteuze behandeling bestaat uit regelmatig bewegen, al dan niet onder begeleiding van een fysiotherapeut, voorlichting en steun van patiëntenverenigingen en zelfhulpgroepen. Het is van belang om gedurende het hele leven te blijven bewegen - in de vorm van gerichte oefeningen - om de mobiliteit optimaal en de conditie op peil te houden. Bovendien kan een sterk spierkorset bijdragen aan het voorkómen van houdingsafwijkingen. Daarnaast heeft bewegen een gunstig effect op pijn, lichamelijk functioneren en op het ervaren van ziekteactiviteit. De basis van de medicamenteuze behandeling bestaat uit ontstekingsremmende pijnstillers, de zogenaamde NSAID's (non-steroidal anti-inflammatory drugs). Deze geven een snelle verlichting van pijn en stijfheidsklachten indien ze in de juiste dosering gebruikt worden. Wanneer de behandeling met NSAID's onvoldoende effectief is, bestaat sinds een paar jaar de mogelijkheid om te behandelen met TNF-α remmers. Deze TNF-α remmers kunnen een sterk en snel positief effect hebben op onder andere rugklachten, algemeen dagelijks functioneren, mobiliteit, perifere artritis, enthesitis en op MRI zichtbare ontsteking. Echter, het nadeel van het gebruik van TNF-α remmers is dat de behandeling erg duur is en dat een verhoogde kans op infecties bestaat.

Dit proefschrift

Ondanks dat er de afgelopen 10-15 jaar grote stappen in de goede richting zijn gezet, zijn er nog steeds uitdagingen in het verbeteren van de herkenning en het behandelen van (vroege) SpA patiënten.
De onderzoeken beschreven in de eerste hoofdstukken van dit proefschrift zijn grotendeels uitgevoerd in het SPondyloArthritis Caught Early (SPACE)-cohort en het DEvenir des Spondylarthropathies Indifférenciées Récentes (DESIR)-cohort. Een cohort is een groep personen die gedurende een bepaalde periode een zelfde gebeurtenis heeft meegemaakt

en gedurende een bepaalde periode in een onderzoek wordt gevolgd. In dit geval is deze gebeurtenis het hebben van rugpijn, en in beide cohorten worden dan ook patiënten met rugpijn ingesloten om onder andere te kunnen bestuderen hoe patiënten met axiale SpA het beste onderscheiden kunnen worden van patiënten zonder SpA, welke factoren voorspellend zijn voor SpA en welke factoren voorspellend zijn voor een progressief beloop van de ziekte. Om deze vragen te kunnen beantwoorden wordt informatie verzameld over de aan- en afwezigheid van alle SpA-kenmerken. Maar er bestaan ook verschillen tussen beide cohorten. Het DESIR-cohort is een puur Frans cohort met 25 participerende centra uit heel Frankrijk, terwijl het SPACE-cohort - wat begonnen is als een Nederlands cohort - een internationaal cohort is met participerende centra in Nederland, Noorwegen, Italië en Zweden.

De **hoofdstukken 2-4** vormen samen deel I van dit proefschrift met de focus op de vroege herkenning van (axiale) SpA en classificatiecriteria. In **hoofdstuk 2** hebben we in het SPACE-cohort gekeken welke van de bestaande criteria sets het hoogste aantal patiënten (met en zonder SpA) op de juiste manier kan classificeren. Onze speciale interesse ging hierbij uit naar de ASAS axiale SpA classificatiecriteria, omdat deze relatief nieuw zijn. Soortgelijke vergelijkingen hebben we ook gemaakt in het Leiden Early Arthritis Clinic (EAC)-cohort in **hoofdstuk 4** om de ASAS perifere SpA classificatiecriteria te testen. Voor het EAC-cohort worden sinds 1993 patiënten verwezen naar de polikliniek reumatologie van het LUMC door huisartsen uit de omgeving omdat er een verdenking op artritis bestaat. Wanneer bij de reumatoloog blijkt dat de patiënt daadwerkelijk een artritis heeft met een symptoomduur korter dan 2 jaar, dan wordt de patiënt in het EAC-cohort opgenomen. In **hoofdstuk 3** hebben we gekeken hoeveel juiste diagnoses het originele Berlijn algoritme oplevert in onder andere het SPACE-cohort en hebben we ook gekeken of we het Berlijn algoritme nog konden verbeteren.

Deel II van dit proefschrift gaat over de rol van beeldvormende technieken (röntgenfoto's en MRI) in de vroege diagnose van axiale SpA. Omdat uit eerder onderzoek is gebleken dat het lastig is om röntgenfoto's van de SI-gewrichten goed te beoordelen, hebben we in **hoofdstuk 5** gekeken wat er gebeurt met de classificatie van een patiënt volgens de mNY criteria wanneer de röntgenfoto door verschillende beoordelaars geïnterpreteerd wordt. In **hoofdstuk 6** hebben we een soortgelijk onderzoek uitgevoerd, maar dan hebben we gekeken of de classificatie van een patiënt volgens de ASAS axiale SpA criteria veranderde wanneer naast de beoordelingen van de röntgenfoto's ook MRI's van de SI-gewrichten door verschillende beoordelaars geïnterpreteerd worden.

Tot nu toe is er alleen gesproken over een afwijkende ('positieve') en niet-afwijkende ('negatieve') MRI van de SI-gewrichten, maar het is ook mogelijk om in meer detail naar de hoeveelheid ontsteking op een MRI van de SI-gewrichten te kijken, bijvoorbeeld met behulp van de SPondyloArthritis Research Consortium of Canada (SPARCC)-methode. Het bereik van de SPARCC-score loopt van 0 tot 72 en heeft daarmee een hoge gevoeligheid voor veranderingen in ontsteking over de tijd. De SPARCC-methode wordt daarom vaak gebruikt om de effecten van behandeling in klinische onderzoeken te beoordelen. De meetkundige eigenschappen van de SPARCC-methode zijn echter nog niet uitgebreid onderzocht, en daarom hebben wij die geëvalueerd in **hoofdstuk 7**.

Deel III van dit proefschrift gaat over het behandelen van axiale SpA en AS. Sinds de verschijning van de eerste versie van de ASAS/EULAR aanbevelingen voor behandeling van AS in het algemeen en de eerste update van de ASAS/EULAR aanbevelingen voor behandeling van AS met TNF-α remmers in 2006 zijn er veel onderzoeken gedaan en publicaties verschenen over de effecten van (nieuwe) behandelingen voor AS en SpA, en dus waren de ASAS/EULAR aanbevelingen toe aan een update. In **hoofdstuk 8** hebben we

een overzicht gemaakt van de nationale richtlijnen van 23 landen over de behandeling van AS met TNF-α remmers en hebben we gekeken in hoeverre deze nationale richtlijnen overeenkomen met en verschillen van de ASAS/EULAR aanbevelingen. Deze informatie is in aanmerking genomen bij de update van de ASAS/EULAR aanbevelingen voor behandeling van AS met TNF-α remmers. De update van de ASAS/EULAR management aanbevelingen staan beschreven in **hoofdstuk 11**. Twee literatuuroverzichten beschreven in **hoofdstuk 9** en **hoofdstuk 10** vormden de basis voor deze update. Een overzicht van de huidige literatuur over de niet-medicamenteuze behandeling van AS en de behandeling met medicamenten zoals NSAID's staat beschreven in **hoofdstuk 9**, gevolgd door een overzicht van de huidige literatuur over de behandeling met onder andere TNF-α remmers in **hoofdstuk 10**.
De in dit proefschrift uitgevoerde onderzoeken hebben bijgedragen aan het zich continu ontwikkelende veld van SpA, en de belangrijkste resultaten en conclusies zijn hieronder per deel samengevat en bediscussieerd.

Deel I: Vroege herkenning van (axiale) SpA en classificatie criteria

In een groep patiënten met voornamelijk axiale klachten in het SPACE-cohort (**hoofdstuk 2**) en in een groep patiënten met voornamelijk perifere klachten in het EAC-cohort (**hoofdstuk 4**), hebben we getest welke van de bestaande classificatiecriteria (mNY, ESSG, Amor, ASAS en CASPAR criteria) het hoogste aantal patiënten op de juiste manier kan classificeren. De CASPAR criteria (ClASsification criteria for Psoriatic ARthritis) zijn uitsluitend voor het classificeren van artritis geassocieerd met psoriasis (artritis psoriatica). De methodes die we in beide hoofdstukken gebruikt hebben om dit te testen zijn zeer vergelijkbaar; in beide cohorten gold de diagnose van de reumatoloog als de standaard. Alle patiënten met de diagnose artritis psoriatica en de diagnose perifere SpA (ruim 10% van het gehele EAC-cohort) en een controle-groep met gelijke leeftijd, geslacht en symptoomduur uit het EAC-cohort zijn bestudeerd in **hoofdstuk 4**. De ASAS perifere SpA criteria en de CASPAR criteria classificeren grotendeels dezelfde artritis psoriatica patiënten, maar de CASPAR criteria classificeerden verreweg de meeste patiënten in de artritis psoriatica groep op de juiste manier. In de perifere SpA groep classificeerden de ASAS perifere SpA criteria verreweg de meeste patiënten op de juiste manier. Aangezien het doel van classificatiecriteria is om een homogene groep patiënten aan te kunnen duiden, terwijl de klinische presentatie van SpA zo heterogeen is, is het te verwachten dat niet alle patiënten die door de reumatoloog als perifere SpA gediagnosticeerd zijn, worden opgepikt door de classificatiecriteria. Van de perifere SpA patiënten wordt 48.7% ook als perifere SpA geclassificeerd volgens de ASAS perifere SpA criteria (sensitiviteit). Een belangrijke bevinding is dat de ASAS perifere SpA en de CASPAR criteriasets de patiënten in de controlegroepen in het overgrote merendeel terecht classificeren als niet hebbende de aandoening (89.8% en 95.6% specificiteit voor respectievelijk de ASAS perifere SpA en CASPAR criteria). Dit is met name van belang voor de ASAS perifere SpA criteria, omdat deze criteria nog vrij nieuw zijn en de angst bestond dat ze misschien te veel patiënten ten onrechte classificeerden als perifere SpA.
De resultaten gevonden en beschreven in **hoofdstuk 4** zijn recent ook bevestigd in het ESPERANZA-cohort. Dit Spaanse cohort is ontwikkeld om de vroege herkenning en diagnostisering van SpA te faciliteren en om de kennis en praktische vaardigheden van huisartsen en specialisten in SpA te verbeteren. Patiënten met klachten (IBP of asymmetrische artritis (bij voorkeur van de onderste extremiteiten), of rugpijn/gewrichtspijn in combinatie met één ander SpA-kenmerk) kunnen worden geïncludeerd in dit cohort. Door deze inclusiecriteria te hanteren, wordt in het ESPERANZA-cohort dus een iets andere groep patiënten gevormd dan in het EAC-cohort, maar de gevonden sensitiviteit (56%) en specificiteit (85%) van de ASAS perifere SpA criteria zijn zeer vergelijkbaar met de resultaten

gevonden in het EAC-cohort.
We hebben de bestaande classificatiecriteria ook getest in een setting met patiënten met voornamelijk axiale klachten, zoals geïncludeerd in het SPACE-cohort (**hoofdstuk 2**). In Nederland heeft ongeveer 20.8% van de bevolking last van rugpijn, waarvan het merendeel geduid wordt als aspecifieke rugklachten. Vóór aanvang van het SPACE-cohort was men dan ook bang dat de polikliniek reumatologie overladen zou worden met patiënten met aspecifieke rugpijn, maar in **hoofdstuk 2** hebben we aangetoond dat deze angst ongegrond is gebleken. Tijdens de eerste visite wordt bij alle doorverwezen patiënten een anamnese afgenomen, lichamelijk onderzoek uitgevoerd, worden laboratoriumtesten uitgevoerd (inclusief het testen op de aanwezigheid van HLA-B27) en worden een MRI en röntgenfoto's van de SI-gewrichten en de rug gemaakt. Dit heeft de mogelijkheid geboden om te bepalen hoeveel patiënten er ten tijde van de eerste visite aan welke classificatiecriteria voldoen (mNY, Amor, ESSG, ASAS axial SpA criteria). Bijna 60% van de patiënten in het SPACE-cohort voldoet aan één of meerdere classificatiecriteria; bijna 40% voldoet aan de ASAS axiale SpA criteria (**hoofdstuk 2**). De ASAS axiale SpA criteria presteren in het SPACE-cohort het beste van alle classificatiecriteria, doordat de ASAS axiale SpA criteria het hoogste aantal patiënten op de juiste manier classificeren. Bijna vijfentachtig procent van de axiale SpA patiënten (volgens de reumatoloog) werden ook als axiale SpA geclassificeerd door de ASAS criteria en bijna vijfennegentig procent van de patiënten zonder axiale SpA werden ook niet geclassificeerd door de ASAS criteria.
Het percentage patiënten in het SPACE-cohort dat aan de ASAS axiale SpA criteria voldoet (38.2%) lijkt op het eerste gezicht misschien wat hoog, wat de indruk kan wekken dat er sprake is van toegenomen bewustzijn over het SPACE-cohort onder verwijzende artsen, waardoor alleen patiënten met een hoge verdenking op SpA verwezen zouden worden ('referral bias'). Daarom hebben we per jaar dat het SPACE-cohort loopt, gekeken of het percentage patiënten dat aan de ASAS criteria voldoet verschilt, maar dit blijkt niet zo te zijn. Ook het percentage patiënten verwezen naar het LUMC van buiten de regio Leiden was ongeveer gelijk over de jaren. Bovendien zijn er ook 33/157 patiënten (21.0%) verwezen zonder SpA-kenmerken of met slechts 1 weinig specifiek SpA-kenmerk. Hoewel nooit uit te sluiten, maakt dit 'referral bias' minder waarschijnlijk. Bovendien zijn er onlangs zeer vergelijkbare resultaten gevonden in een ander onderzoek waarin ook patiënten met rugpijn die 3 maanden of langer bestaat en ontstaan is voor het 45e levensjaar zijn geïncludeerd, gerekruteerd onder reumatologen in Frankrijk.
Patiënten kunnen op twee manieren aan de ASAS axSpA criteria voldoen: via de klinische arm (aanwezigheid van HLA-B27 plus 2 andere SpA-kenmerken) en via de imaging arm (sacroiliitis plus 1 ander SpA-kenmerk). Aangezien er aanwijzingen zijn dat reumatologen en registratieautoriteiten (U.S. Food and Drug Administration (FDA)) de klinische arm minder accepteren als voldoende specifiek voor classificatie dan de imaging arm, hebben we in het SPACE-cohort patiënten in beide armen met elkaar vergeleken (**hoofdstuk 2**). Patiënten in beide armen bleken opvallend gelijkwaardig wat betreft ziekteactiviteit (BASDAI (Bath Ankylosing Spondylitis Disease Activity Index) en ASDAS (Ankylosing Spondylitis Disease Activity Score)) en de aanwezigheid van SpA-kenmerken. Dergelijke vergelijkingen zijn ook gemaakt in het DESIR-cohort en in het ABILITY-1 onderzoek. Het ABILITY-1 onderzoek is een gerandomiseerd placebogecontroleerd onderzoek naar de effecten en veiligheid van adalimumab - één van de veelgebruikte TNF-α remmers - in patiënten die niet aan de mNY criteria, maar wel aan de ASAS axiale SpA criteria voldoen (non-radiografische axiale SpA). Ook in deze onderzoeken is gevonden dat patiënten in beide armen van de ASAS axiale SpA criteria zeer vergelijkbaar zijn. Wel blijkt het zo te zijn dat reumatologen zekerder zijn van hun diagnose zijn als er sacroiliitis op MRI en/of röntgenfoto's te zien is (**hoofdstuk**

2). Dit lijkt erop te wijzen dat reumatologen hun diagnose zwaar laten afhangen van de aan- of afwezigheid van (radiografische) sacroiliitis. Dit is recent ook bevestigd in een onderzoek uitgevoerd in Frankrijk (hierboven reeds genoemd) waarbij aangetoond is dat (radiografische) sacroiliitis het hoogste voorspellend vermogen heeft voor de diagnose SpA volgens de reumatoloog.
Verder hebben we in het SPACE-cohort gevonden dat ongeveer 80% van de patiënten met axiale SpA (volgens de ASAS axiale SpA criteria) IBP heeft, en andersom, dat bijna 60% van de patiënten zonder SpA IBP heeft (**hoofdstuk 2**). Deze resultaten tonen aan dat IBP geen goed onderscheid kan maken tussen patiënten met en zonder axiale SpA in het SPACE-cohort en we bevestigen hiermee eerder gevonden resultaten. Maar omdat het uitvragen van IBP geen kosten met zich mee brengt, kan IBP toch van nut zijn in het diagnostisch proces, zolang het maar in combinatie met andere SpA-kenmerken wordt toegepast (hoe meer SpA-kenmerken aanwezig, des te suggestiever voor SpA). Bovendien lijkt er bewijs te zijn dat bepaalde items van IBP meer voorspellend voor SpA zijn dan andere items. Maar de interpretatie van IBP (items) verschilt van persoon tot persoon, zoals blijkt uit de lage overeenstemming tussen huisartsen en reumatologen over de aanwezigheid van IBP. Dit alles tezamen maakt de rol van IBP in het diagnostisch proces onduidelijk.
IBP wordt gebruikt als (deel van het) toelatingscriterium in de ESSG criteria en het diagnostische Berlijn algoritme. Gezien het feit dat slechts 70-80% van de patiënten met axiale SpA typische IBP klachten heeft, betekent dit dat 20-30% van de patiënten met axiale SpA niet door het Berlijn algoritme kunnen worden opgepikt. Dit heeft ons gestimuleerd om verbeteringen aan het algoritme aan te brengen.
We hebben 2 modificaties van het algoritme voorgesteld in **hoofdstuk 3**. In de eerste modificatie stellen we voor dat patiënten met ≥3 van de 5 (in plaats van ≥4 van de 5) items van IBP het algoritme in mogen stromen. In de tweede modificatie verandert de gehele structuur van het algoritme daar we voorstellen om IBP als toelatingscriterium te verwijderen en toe te voegen aan de lijst van SpA-kenmerken. Bovendien mogen nu alle patiënten met rugpijn ≥3 maanden die ontstaan is vóór het 45e levensjaar het algoritme instromen. Dit resulteert in drie groepen met respectievelijk ≥4, 2-3 en 0-1 SpA-kenmerken. Alle drie de algoritmes hebben we getest in het SPACE-cohort en het ASAS-cohort met de classificatie volgens de ASAS axiale SpA criteria, de post-test waarschijnlijkheid op het hebben van SpA en de diagnose van de reumatoloog als externe standaards. Net zoals in het SPACE-cohort worden in het ASAS-cohort ook patiënten met rugpijn ≥3 maanden ontstaan vóór het 45e levensjaar geïncludeerd, maar patiënten moeten daarnaast ook een sterke verdenking op SpA hebben (zonder reeds een definitieve diagnose op het moment van verwijzen). In beide cohorten pikt modificatie 1 meer juiste SpA patiënten op dan het originele algoritme (hogere sensitiviteit), ten koste van slechts een klein beetje verlies aan specificiteit. De sensitiviteit van modificatie 2 is nog iets hoger, en ook de specificiteit is iets hoger dan de specificiteit van modificatie 1. De resultaten zijn aan de ASAS-leden gepresenteerd tijdens een vergadering in januari 2012 in Amsterdam. De leden stemden voor modificatie 2 als het nieuwe algoritme.
Maar er valt nog na te denken over verdere verbeteringen. Reumatologen zouden bijvoorbeeld kunnen overwegen een MRI te maken in HLA-B27-negatieve (mannelijke) patiënten die 2-3 andere SpA-kenmerken hebben. In modificatie 2 verlaten deze patiënten het algoritme, maar als de MRI positief zou zijn, dan voldoen deze patiënten aan de ASAS axiale SpA criteria. Daarnaast is het waardevol om te onderzoeken of het zinnig is om binnen de groep HLA-B27-negatieve patiënten met 0 of 1 SpA-kenmerk onderscheid te maken tussen de patiënten zonder SpA-kenmerken en de patiënten met 1 SpA-kenmerk om te beslissen of er nog een MRI gemaakt moet worden of niet. In het huidige algoritme wordt geadviseerd

om bij al deze patiënten een MRI te maken, maar alleen de patiënten met 1 kenmerk zouden aan de ASAS axiale SpA criteria kunnen voldoen als de MRI positief blijkt te zijn, terwijl de patiënten zonder kenmerken niet aan de ASAS axiale SpA criteria zullen kunnen voldoen. Bovendien valt het nut van het maken van een röntgenfoto van de SI-gewrichten als eerste stap te betwijfelen, aangezien dit algoritme vaak toegepast zal worden in nieuw te diagnosticeren patiënten met een korte klachtengeschiedenis. Dit wordt gereflecteerd door het hoge aantal normale röntgenfoto's in het SPACE- en het ASAS-cohort. Er zou daarom overwogen kunnen worden of het zinniger is om het maken van de röntgenfoto's van de SI-gewrichten één van de vervolgstappen te laten zijn, nadat de informatie uit de anamnese en het lichamelijk onderzoek bekend is, zoals ook in de dagelijkse praktijk gebruikelijk is. Desalniettemin kan modificatie 2 in de huidige vorm al van nut zijn voor reumatologen in de dagelijkste praktijk. De Nederlandse Vereniging voor Reumatologie (NVR) heeft recent richtlijnen voor de diagnose en de behandeling van axiale SpA gepubliceerd, waarin dit gemodificeerde algoritme is opgenomen.

Alhoewel er de afgelopen jaren enorme vooruitgang is geboekt in het diagnosticeren en classificeren van SpA patiënten, worden nog niet altijd de juiste patiënten naar de reumatoloog doorverwezen. Dit komt waarschijnlijk doordat verwijzers slechts geringe kennis hebben van de manifestaties die bij SpA horen, zoals geïllustreerd wordt door de slechte overeenstemming tussen verwijzers en reumatologen over de aanwezigheid van IBP in het onderzoek van een Duitse en een Spaanse onderzoeksgroep. Het blijkt een uitdaging voor verwijzers om patiënten met een verdenking op SpA, die naar de reumatoloog verwezen zouden moeten worden, te herkennen. Daarom zijn er in de loop van de tijd verscheidene (complexe) verwijsstrategieën ontwikkeld met het doel patiënten met een verdenking op SpA in een vroeg stadium te identificeren en naar de reumatoloog te verwijzen. Alle strategieën doen het goed in een onderzoeksomgeving waarin de verwijzers geïnstrueerd zijn (24-45% SpA patiënten onder alle verwezen patiënten), maar de implementatie van deze (complexe) verwijsstrategieën in de dagelijkse eerstelijnszorg blijkt lastig. Daarom is het wellicht nuttig om een zeer gemakkelijke verwijsstrategie te overwegen, zoals het verwijzen van alle patiënten met rugpijn ≥3 maanden die ontstaan is voor het 45e levensjaar - de toelatingscriteria van de ASAS axiale SpA criteria. In het SPACE-cohort, waarin de additionele restrictie van een maximum van 2 jaar klachten wordt toegepast en in een onderzoek uit Frankrijk leveren deze criteria hoge percentages SpA patiënten op (respectievelijk 41.4% (**hoofdstuk 2**) en 35.1%). Maar het moet nog blijken of een dergelijke strategie ook in andere centra en in andere landen zal werken omdat de gezondheidszorgsystemen overal verschillend zijn. Meer onderzoek hiernaar is dus noodzakelijk.

Bovendien zijn studies met een lange periode van follow-up nodig om de uitkomsten van patiënten in de klinische arm en patiënten in de imaging arm van de ASAS axiale SpA criteria te bestuderen en te vergelijken. Dit zal bijdragen aan het begrip van het ziektebeeld. Daarnaast zal meer kennis van uitkomsten op lange termijn ook helpen in het overwegen van mogelijke aanpassingen aan de ASAS axiale SpA criteria zelf, bijvoorbeeld het wegen van de verschillende SpA-kenmerken omdat sommige SpA-kenmerken, zoals een positieve familieanamnese voor SpA en HLA-B27 positiviteit, onderling sterker gecorreleerd zijn dan andere SpA-kenmerken.

Deel II: De rol van beeldvormende technieken in de vroege diagnose van spondyloartritis

MRI heeft de afgelopen jaren zijn nut bewezen in het diagnosticeren en classificeren van patiënten met SpA. Desondanks staat de nieuwverworven prominente rol van MRI, maar ook de rol van conventionele röntgenfoto's, ter discussie. Dit, omdat bekend is dat het een uitdaging is om betrouwbaar beelden van de SI-gewrichten te beoordelen, met name

röntgenfoto's. De SI-gewrichten hebben een complexe anatomie, waarbij het golvende gewrichtsoppervlak wordt afgebeeld op een 2-dimensionaal beeld en dit kan leiden tot verkeerde interpretaties.
Waar in de dagelijkse praktijk de diagnose AS gebaseerd is op de beoordeling van een röntgenfoto door een lokale radioloog en/of reumatoloog, worden beelden in klinische onderzoeken en cohorten meestal beoordeeld door één of meer getrainde 'centrale' lezers. Recent is de variatie in het interpreteren van röntgenfoto's bevestigd in een post-hoc analyse van de data van het ABILITY-1 onderzoek en het RAPID-axSpA onderzoek voor de registratie van TNF-α remmers in patiënten met niet-radiografische axiale SpA. De analyses van de ABILITY-1 data laten zien dat 37% van de patiënten die door lokale lezers geclassificeerd was als niet-radiografische axiale SpA door centrale lezers als AS (mNY criteria positief) werd geclassificeerd. In het RAPID-axSpA onderzoek laat een soortgelijke analyse zien dat 36% van de patiënten anders geclassificeerd werd; 26% van de door lokale lezers als niet-radiografische axiale SpA geclassificeerde patiënten werd door de centrale lezers als AS (mNY positief) geclassificeerd, en 10% van de door lokale lezers als AS (mNY positief) geclassificeerde patiënten werd als niet-radiografische axiale SpA geclassificeerd door centrale lezers.
Daar in het DESIR-cohort de beoordelingen van MRI en röntgenfoto's van zowel lokale als centrale lezers beschikbaar zijn, bood dit cohort de unieke gelegenheid om soortgelijke analyses uit te voeren. De vergelijking van de beoordeling van de aan- of afwezigheid van radiografische sacroiliitis staat beschreven in **hoofdstuk 5** en laat zien dat de overeenstemming tussen lokale lezers en centrale lezers (als externe standaard) slechts matig is. Lokale lezers zien veel vaker sacroiliitis dan centrale lezers; 41.5% van de patiënten die volgens lokale lezers als AS zijn geclassificeerd, kon niet worden bevestigd door de centrale lezers. Zelfs wanneer we een heel strikte definitie van radiografische sacroiliitis voor de lokale lezers hanteerden, kon 29.4% van de patiënten, die volgens de lokale lezers als AS zijn geclassificeerd, niet worden bevestigd door de centrale lezers. Dit percentage liep zelfs op tot 65.1% wanneer we daarnaast ook nog eens een strikte definitie voor de centrale lezers hanteerden. Hier moet wel bij gezegd worden dat de lokale lezers op een iets andere manier de röntgenfoto's hebben beoordeeld dan de centrale lezers en dat het werkelijke percentage fout-positieve classificaties door lokale lezers dus ergens tussen de 29.4% en 65.1% ligt.
Daarnaast werd 7.5% van de patiënten die volgens de centrale lezers als AS zijn geclassificeerd, niet herkend door lokale lezers. Dit percentage liep op tot 11.7% met de striktere definitie van sacroiliitis voor lokale lezers. Zelfs in patiënten bij wie ankylose (benige brug tussen beide gewrichtsvlakken van het SI-gewricht) te zien is of bij wie beide SI-gewrichten zijn aangedaan, waren er grote verschillen in de beoordeling tussen lokale en centrale lezers.
Maar ook de overeenstemming over de aan- of afwezigheid van radiografische sacroiliitis tussen de twee centrale lezers onderling is slechts matig. Dit benadrukt nogmaals de moeilijkheid van het beoordelen van de SI-gewrichten op röntgenfoto's en laat tegelijkertijd zien dat training de beoordeling niet substantieel lijkt te verbeteren. Maar waar de discrepantie tussen de centrale lezers in beide richtingen te zien is, bestaat de misclassificatie door lokale lezers bijna volledig uit fout-positieve classificaties, wat resulteert in een onacceptabel hoog percentage onterechte diagnoses van AS in de dagelijkse praktijk. Of deze patiënten wel non-radiografische axiale SpA hadden, hebben we verder onderzocht in **hoofdstuk 6**. Daar hebben we soortgelijk onderzoek ook voor MRI uitgevoerd in het DESIR-cohort.
In tegenstelling tot de slechts matige overeenstemming over radiografische sacroiliitis, bestaat er een aanzienlijke overeenstemming tussen lokale lezers en centrale lezers en

ook tussen de twee centrale lezers onderling over sacroiliitis op MRI. Vervolgens hebben we gekeken wat er gebeurt met de classificatie van patiënten volgens de ASAS axiale SpA criteria wanneer de MRI en/of röntgenfoto mogelijk anders geïnterpreteerd wordt door verschillende lezers.
In 163/582 patiënten (28.0%), bij wie de beoordeling van de MRI en/of röntgenfoto door lokale lezers verschilde van de beoordeling door centrale lezers, zou potentieel de classificatie volgens de ASAS axiale SpA criteria ook verschillend kunnen zijn. In slechts 46/582 patiënten (7.9%) was dit ook daadwerkelijk het geval. Dit toont aan dat de ASAS axiale SpA criteria vrij ongevoelig zijn voor verschillen in beoordeling van MRI en/of röntgenfoto. Dit wordt met name veroorzaakt door de aanwezigheid van de klinische arm. Patiënten die namelijk aan deze klinische arm voldoen, zullen dat altijd blijven doen omdat de HLA-B27-status nooit zal veranderen, ongeacht de beoordelingen van de MRI en röntgenfoto en er zijn veel patiënten die zowel aan de imaging als aan de klinische arm voldoen.
Gezien de matige betrouwbaarheid van het beoordelen van conventionele röntgenfoto's en de aanzienlijke betrouwbaarheid van het beoordelen van MRI's, beargumenteren sommige experts dat het misschien wel een optie is om helemaal geen röntgenfoto's meer te maken en alleen te volstaan met MRI's. In het DESIR-cohort hebben we gekeken wat de mogelijke consequenties zijn van het compleet weglaten van röntgenfoto's (**hoofdstuk 6**). Wanneer we kijken naar de volledige ASAS axiale SpA criteria - inclusief de klinische arm - dan zouden 11 tot 14 patiënten (1.9-2.4%; gebaseerd op de beoordeling van lokale of centrale lezers) gemist worden die op basis van radiografische sacroiliitis (en dus geen sacroiliitis op MRI hebben) alleen aan de imaging arm voldoen (en dus niet aan de klinische arm). De resultaten zouden er misschien anders uit hebben gezien als een dergelijke analyse niet in een cohort als DESIR zou zijn uitgevoerd waar patiënten met kort bestaande klachten zijn geïncludeerd, maar in een cohort met patiënten met lange ziektegeschiedenis.
De huidige definitie van een positieve MRI is momenteel alleen gebaseerd op de aanwezigheid van actieve ontstekingen, die niet te zien zijn op röntgenfoto's. Op röntgenfoto's is juist alleen schade aan het bot te zien. Dus de discussie over het weglaten van röntgenfoto's wordt pas echt interessant als de aanwezigheid van schade aan het bot (erosies, ankylose en sclerose) een rol zou kunnen spelen in de definitie van een positieve MRI. Voordat bepaald kan worden of dergelijke structurele afwijkingen ook daadwerkelijk van toegevoegde waarde kunnen zijn op de definitie van een positieve MRI, moet er eerst gekeken worden of deze afwijkingen betrouwbaar afgebeeld en beoordeeld kunnen worden op MRI. Daarom is in het SPACE-cohort en het German Spondyloarthritis Inception Cohort (GESPIC)-cohort getest wat het vermogen van MRI is om structurele afwijkingen in de SI-gewrichten te detecteren. Het GESPIC-cohort bestaat uit AS patiënten met een klachtengeschiedenis ≤10 jaar en non-radiografische axiale SpA patiënten met een klachtengeschiedenis ≤5 jaar. Om het vermogen te testen van MRI om structurele afwijkingen te detecteren, zijn de structurele afwijkingen die te zien zijn op MRI vergeleken met de structurele afwijkingen die te zijn op röntgenfoto's, die tevens als gouden standaard dienden. Over het algemeen is de overeenstemming tussen MRI en röntgenfoto's echter slecht. Om deze resultaten te bevestigen of te weerleggen zouden deze vergelijkingen in andere cohorten moeten worden herhaald en zouden ze uitgebreid kunnen worden met een alternatieve externe standaard zoals CT.
Momenteel wordt in het DESIR-cohort en het SPACE-cohort onderzocht of en zo ja, in welke combinatie(s) structurele afwijkingen aan het bot zouden kunnen worden toegevoegd aan de definitie van een positieve MRI. Dergelijke analyses zouden ook kunnen worden uitgevoerd in andere cohorten waarin patiënten met een langere klachtengeschiedenis zijn geïncludeerd. Totdat de resultaten van deze onderzoeken bekend zijn, kunnen er nog geen

definitieve conclusies getrokken worden over de rol die röntgenfoto's en de mogelijke rol die structurele afwijkingen op MRI zouden kunnen gaan spelen.
MRI wordt ook gebruikt om de hoeveelheid ontsteking in de SI-gewrichten en rug te kwantificeren, bijvoorbeeld met behulp van de SPARCC-methode. De SPARCC-methode heeft een hoge gevoeligheid voor veranderingen in ontsteking over de tijd, maar het is niet bekend of alle veranderingen die gemeten kunnen worden met de SPARCC-methode ook daadwerkelijk klinisch relevante veranderingen zijn, of dat het misschien om meetfouten gaat. De onderzoeksgroep die de SPARCC-methode ontwikkeld heeft, heeft een voorstel gedaan voor een klinisch relevante verandering in de uitkomst (minimally important change (MIC)) van 2.5 SPARCC-punten. Maar de kleinst detecteerbare, voor meetfouten gecorrigeerde verandering (smallest detectable change (SDC)) is nog nooit berekend en daarom is het niet duidelijk of een klinisch relevante verandering te onderscheiden valt van een meetfout. Daarom hebben we in **hoofdstuk 7** de SDC van de SPARCC-score berekend in het SPACE-cohort over periodes van 3 maanden en 1 jaar. De gevonden SDC's zijn 3.4 en 2.1 SPARCC-punten over een periode van 3 maanden, en 2.4 SPARCC-punten over een periode van 1 jaar en liggen daarmee dus dicht bij de voorgestelde MIC. Een verandering kleiner dan de SDC kan beschouwd worden als meetfout of 'ruis'. Een groot deel van de veranderingen in SPARCC-score in de patiënten in het SPACE-cohort over de periodes van 3 maanden en 1 jaar kan als meetfout beschouwd worden, omdat deze kleiner zijn dan de berekende SDC's (45-62.9% over 3 maanden tijd en 39.1% over 1 jaar tijd). Het is opmerkelijk dat het merendeel van de daadwerkelijke veranderingen (de veranderingen die dus groter zijn dan de berekende SDC's), zowel toenames als afnames, worden geobserveerd in patiënten bij wie de behandeling niet veranderd is. Deze observatie suggereert sterk dat ontstekingen op MRI spontaan kunnen fluctueren over de tijd. Dit is belangrijke informatie om rekening mee te houden bij het bestuderen en interpreteren van de resultaten van klinische onderzoeken en bij het verder uitpluizen van de hypothese dat ontsteking zou aanzetten tot het vormen van structurele afwijkingen.
Daarnaast hebben we gekeken welke afkapwaarde van de SPARCC-score het beste equivalent van een positieve MRI volgens de ASAS definitie is. Dit blijkt een afkapwaarde van ≥2 SPARCC-punten te zijn. Aangezien de SPARCC-methode vaak in klinische onderzoeken gebruikt wordt om de effecten van behandeling met geneesmiddelen te meten, kan deze afkapwaarde gebruikt worden om patiënten in subgroepen in te delen met en zonder een positieve MRI. Dit kan bijvoorbeeld gebruikt worden als een voorspellende factor voor de respons op behandeling.

Deel III: Behandeling van axiale SpA en AS

De in 2006 door de ASAS/EULAR gepubliceerde aanbevelingen voor de behandeling van AS waren toe aan een update, aangezien sindsdien het aantal klinische onderzoeken en publicaties over (nieuwe) behandelingen van AS enorm gegroeid is. De update van deze aanbevelingen uit 2010 staat beschreven in **hoofdstuk 11**. Systematische literatuuroverzichten liggen hieraan ten grondslag, zoals geadviseerd wordt door de EULAR. Deze literatuuroverzichten staan beschreven in **hoofdstuk 9** en **hoofdstuk 10**.
In de titel van de update van de aanbevelingen wordt alleen AS genoemd en niet axiale SpA omdat het om een update van de bestaande aanbevelingen voor AS ging en omdat het bewijs van trials in non-radiografische axiale SpA nog beperkt was ten tijde van de update. Desalniettemin was de projectgroep unaniem van mening dat deze aanbevelingen ook in axiale SpA kunnen worden toegepast, omdat AS een deel van de totale axiale SpA groep is en omdat alle beschikbare data aantonen dat effectiviteit van behandeling minstens zo goed is in patiënten met niet-radiografische axiale SpA als in patiënten met AS. Dit laatste is

ook bevestigd in alle trials in niet-radiografische axiale SpA die sinds de verschijning van de aanbevelingen gepubliceerd zijn.
ASAS adviseert elke patiënt een behandeling op maat te geven, waarbij rekening moet worden gehouden met alle aspecten van de ziekte, inclusief extra-articulaire manifestaties, ziekteactiviteit en geslacht etc (**hoofdstuk 11**). De ziekteactiviteit en klinische presentatie van patiënten met AS zou regelmatig gecontroleerd moeten worden volgens standaardparameters zoals door ASAS aanbevolen. De behandeling zou moeten bestaan uit niet-medicamenteuze en medicamenteuze componenten. Niet-medicamenteuze behandeling vormt één van de twee pijlers waarop de behandeling van AS momenteel berust en bestaat uit voorlichting en regelmatig bewegen. Het literatuuroverzicht over niet-medicamenteuze behandeling beschreven in **hoofdstuk 9** laat zien dat thuisoefeningen een goed effect hebben op fysiek functioneren (Bath Ankylosing Spondylitis Functional Index; BASFI), patiënt gerapporteerde ziekteactiviteit (BASDAI), pijn en mobiliteit van de rug, maar dat groeps- en individuele fysiotherapie onder toezicht met oefeningen in het water of droog, effectiever zijn dan thuisoefeningen. Dit is ook zo opgenomen in de ASAS aanbevelingen (**hoofdstuk 11**).
Ontstekingsremmende pijnstillers (NSAID's, inclusief coxibs) in een maximale (ontstekingsremmende) dosis behoren tot de standaardbehandeling en bestrijden effectief pijn en stijfheid op korte en langere termijn (**hoofdstuk 9** en **hoofdstuk 11**). De ASAS adviseert dat patiënten met aanhoudend actieve en symptomatische ziekte NSAID's continu zouden moeten gebruiken. Een belangrijk argument daarvoor komt uit één onderzoek waarin is aangetoond dat continu gebruik van NSAID's verbening beter remt dan wanneer NSAID's slechts sporadisch ('zo nodig') worden gebruikt. Nadat de update van de aanbevelingen was gepubliceerd, is een post-hoc analyse uitgevoerd in dit onderzoek, waaruit is gebleken dat het continue gebruik van NSAID's alleen superieur is aan het sporadisch gebruik van NSAID's in patiënten met verhoogde acute-fase-eiwitten. Daarnaast is recent ook een onderzoek uitgevoerd in het GESPIC-cohort waarbij de invloed van NSAID's op radiografische progressie in de rug over een periode van 2 jaar in AS en niet-radiografische SpA patiënten is bestudeerd. De resultaten van dit onderzoek tonen aan dat inname van hoge doses NSAID's geassocieerd is met minder radiografische progressie in AS patiënten.
Er is geen bewijs dat sulfasalazine en methotrexaat effect hebben op klachten in de rug en het bekken. Echter, sulfasalazine zou overwogen kunnen worden in patiënten met perifere betrokkenheid (**hoofdstuk 11**). Patiënten met aanhoudend hoge ziekteactiviteit, ondanks conventionele behandeling volgens de ASAS aanbevelingen, kunnen in aanmerking komen voor behandeling met TNF-α remmers. De resultaten van het literatuuroverzicht over behandeling met TNF-α remmers is beschreven in **hoofdstuk 10**. Alle TNF-α remmers die voor AS beschikbaar zijn blijken effectief te zijn in het verbeteren van BASDAI, BASFI, en BASMI (Bath Ankylosing Spondylitis Metrology Index) in patiënten met langdurig bestaande AS, maar ook in patiënten met niet-radiografische axiale SpA. Niet-radiografische axiale SpA patiënten blijken met name goed op behandeling met TNF- α remmers te reageren als er sprake is van verhoogde acute-fase-eiwitten in het bloed en/of ontsteking op MRI. Als de patiënt ook IBD heeft, zou het verschil in effectiviteit van de verscheidene TNF-α remmers op gastro-intestinale klachten meegewogen moeten worden. Verplicht gebruik van methotrexaat of sulfasalazine vóór de start van TNF-α remmers of het gelijktijdig gebruik van methotrexaat/sulfasalazine en TNF-α remmers in patiënten met axiale SpA wordt niet door bewijs ondersteund. Overstappen naar een tweede TNF-α remmer kan een gunstig effect hebben, zeker in patiënten die voorheen wel een respons hadden op TNF-α remmers maar deze verloren zijn (**hoofdstuk 10** en **hoofdstuk 11**).
Ondanks dat er eenduidig bewijs is dat TNF-α remmers de ziekteactiviteit, inclusief ontsteking

op MRI en het dagelijks functioneren in patiënten met AS drastisch kunnen verbeteren, kan en hoeft niet elke AS patiënt hiermee behandeld te worden. Daarnaast zijn aan het gebruik van TNF-α remmers hoge kosten verbonden. Daarom heeft de ASAS in 2003 aanbevelingen voor het gebruik van TNF-α remmers gepubliceerd. Sindsdien zijn deze aanbevelingen al twee keer vernieuwd en inmiddels hebben veel landen nationale richtlijnen voor het gebruik van TNF-α remmers ontwikkeld, al dan niet op basis van deze ASAS aanbevelingen.
In **hoofdstuk 8** hebben we de nationale richtlijnen van 23 landen met elkaar en met de ASAS aanbevelingen uit 2006 vergeleken. Deze vergelijking toont aan dat er, ondanks enkele verschillen, over het algemeen consensus bestaat wat betreft het gebruik van TNF-α remmers in AS patiënten. Inmiddels zijn er ook voorspellers voor een goede reactie op TNF-α remmers geïdentificeerd. Deze voorspellers zijn onder andere een hoge ziekteactiviteit gemeten met de ASDAS (Ankylosing Spondylitis Disease activity score), verhoogde acute-fase-eiwitten en een positieve MRI. Behalve de goedkeuring van verscheidene TNF-α remmers voor de behandeling van AS zijn adalimumab en certolizumab ook recent door de Europese Commissie goedgekeurd voor de behandeling van volwassenen met ernstige niet-radiografische axiale SpA, die inadequaat reageren op NSAID's of intolerant zijn voor NSAID's. Maar patiënten met niet-radiografische axiale SpA komen alleen in aanmerking voor behandeling met TNF-α remmers als er objectieve kenmerken van ontsteking aanwezig zijn (verhoogde acute-fase-eiwitten en/of positieve MRI). Toch is het in veel landen nog steeds zo dat patiënten met actieve, ernstige axiale SpA, die geen baat hebben bij behandeling met NSAID's, alleen in aanmerking komen voor behandeling met TNF-α remmers als er sacroiliitis op de röntgenfoto te zien is. Maar het enige verschil tussen niet-radiografische axiale SpA en AS is de aanwezigheid van radiografische sacroiliitis, waarvan nu juist is aangetoond dat het lastig is om die betrouwbaar te interpreteren (**hoofdstuk 5** en **hoofdstuk 6**). Er is daarom meer onderzoek nodig naar de (langetermijn)effecten van TNF-α remmers in vroege niet-radiografische axiale SpA, inclusief patiënten zonder ontsteking op MRI.
Verder zouden de langetermijnuitkomsten van verschillende typen behandelingen (NSAID's versus TNF-α remmers etc) met elkaar vergeleken moeten worden waarbij met name gekeken zou moeten worden of er verschillen zijn in het remmen of misschien zelfs tegengaan van structurele afwijkingen, wat tot op heden (nog) onvoldoende onderzocht is. Ook zouden de effecten van verschillende behandelstrategieën (bijvoorbeeld al dan niet ASDAS-gestuurd behandelen) vergeleken kunnen worden. Ook het bestaan van een 'window of opportunity' zou in detail kunnen worden bestudeerd aangezien er aanwijzingen zijn voor het bestaan ervan. Een 'window of opportunity' wordt gezien als een korte periode waarin het starten van een behandeling een beter effect heeft op zowel korte als lange termijn. Gedacht wordt dat dit 'window of opportunity' zich vroeg in het ziektebeloop opent (en sluit) aangezien er is gebleken dat recent ontstane klachten namelijk geassocieerd zijn met betere reacties en, belangrijker nog, een grotere kans op een goede respons. Dus als er daadwerkelijk een 'window of opportunity' bestaat dan zou dit klinische en biologische voordelen hebben, en zou structurele schade voorkomen kunnen worden, vooral in de actieve axiale SpA patiënten bij wie de ziekte recent is ontstaan en die tekenen van ontsteking op MRI of in lab waarden laten zien.
Al met al kunnen we stellen dat we door de recente ontwikkelingen beter in staat zijn patiënten met SpA tijdig te herkennen, de behandelmogelijkheden verbeterd zijn en er ook een gunstiger beloop van de ziekte met minder schade en beter functioneren mogelijk is. Maar om de zorg voor patiënten met SpA nog verder te optimaliseren, kunnen we het ons niet permitteren stil te blijven staan, maar moeten we in beweging blijven, net als onze SpA patiënten.

List of abbreviations

ACR	American College of Rheumatology
AS	Ankylosing spondylitis
ASAS	Assessment of SpondyloArthritis international Society
ASDAS	Ankylosing Spondylitis Disease Activity Score
ASQoL	Ankylosing Spondylitis Quality of Life
axSpA	axial SpondyloArthritis
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASFI	Bath Ankylosing Spondylitis Functional Index
BASMI	Bath Ankylosing Spondylitis Metrology Index
BME	Bone Marrow Edema
CASPAR	ClASsification criteria for Psoriatic ARthritis
Cox2	Cyclo-oxygenase-2
CRF	Case Record Form
CRP	C-reactive Protein
CT	Computed Tomography
DESIR	DEvenir des Spondylarthropathies Indifférenciées Récentes
DMARD	Disease Modifying Anti-Rheumatic Drug
EAC	Early Arthritis Clinic
ESPAC	Early SPondyloArthritis Clinic
ESR	Erythrocyte Sedimentation Rate
ESSG	European SpondyloArthropathy Study Group
EULAR	European League Against Rheumatism
Gd	Gadolinium
GESPIC	German Spondyloarthritis Inception Cohort
GP	General practitioner
HLA-B27	Human Leucocyte Antigen B27
IBD	Inflammatory Bowel Disease
IBP	Inflammatory Back Pain
LR-	negative Likelihood Ratio
LR+	positive Likelihood Ratio
LUMC	Leiden University Medical Center
MIC	Minimally Important Change
mNY	modified New York
MRI	Magnetic Resonance Imaging
mSASSS	modified Stoke Ankylosing Spondylitis Spinal Score
nr-axSpA	non-radiographic axial SpondyloArthritis
NSAID	Non-Steroidal Anti-Inflammatory Drug
NVR	Nederlandse vereniging voor Reumatologie
OA	OsteoArthritis
PsA	Psoriatic Arthritis
pSpA	peripheral SpondyloArthritis
RA	Rheumatoid Arthritis
SDC	Smallest Detectable Change
SI	sacroiliac
SLR	Systematic Literature Review
SPACE	SPondyloArthritis Caught Early
SPARCC	SPondyloArthritis Research Consortium of Canada
TNF	Tumor Necrosis Factor
UA	Undifferentiated Arthritis
VAS	Visual Analogue Scale

List of publications

1. **van den Berg R**, de Groot S, Swart K, van der Woude L. Physical capacity after 7 weeks of low-intensity wheelchair training. Disabil Rehabil 2010;32:2244-52.

2. **van den Berg R**, van der Heijde D. Vroegdiagnostiek en behandeling inflammatoire rugpijn. Modern Medicine 2010;34:329-32 *[Dutch].*

3. **van den Berg R**, van der Heijde D. How should we diagnose spondyloarthritis according to the ASAS classification criteria: a guide for practicing physicians. Pol Arch Med Wewn 2010;120:452-7.

4. Braun J, **van den Berg R**, Baraliakos X, Boehm H, Burgos-Vargas R, Collantes-Estevez E, Dagfinrud H, Dijkmans B, Dougados M, Emery P, Geher P, Hammoudeh M, Inman R, Jongkees M, Khan M, Kiltz U, Kvien T, Leirisalo-Repo M, Maksymowych W, Olivieri I, Pavelka K, Sieper J, Stanislawska-Biernat E, Wendling D, Özgocmen S, van Drogen C, van Royen B, van der Heijde D. 2010 update of the ASAS/EULAR recommendations for the management of ankylosing spondylitis. Ann Rheum Dis 2011;70:896-904.

5. **van den Berg R**, Stansilawska-Biernat E, van der Heijde D. Comparison of recommendations for the use of anti-tumour necrosis factor therapy in ankylosing spondylitis in 23 countries worldwide. Rheumatology (Oxford) 2011;50:2270-7.

6. van Gaalen F, **van den Berg R**, de Hooge M, van der Heijde D. Vroege herkenning van axiale spondyloartritis met behulp van MRI: het SPACE-project. NTvR 2012;1:31-4 *[Dutch].*

7. **van den Berg R**, Baraliakos X, Braun J, van der Heijde D. First update of the current evidence for the management of ankylosing spondylitis with non-pharmacological treatment and non-biological drugs: a systematic literature review for the ASAS/EULAR management recommendations in ankylosing spondylitis. Rheumatology (Oxford) 2012;51:1388-96.

8. Baraliakos X, **van den Berg R**, Braun J, van der Heijde D. Update of the literature review on the treatment with biologics as basis for the first update of the ASAS/EULAR management recommendations of ankylosing spondylitis. Rheumatology (Oxford) 2012;51:1378-87.

9. **van den Berg R**, van Gaalen F, van der Helm-van Mil A, Huizinga T, van der Heijde D. Performance of classification criteria for peripheral spondyloarthritis and psoriatic arthritis in the Leiden Early Arthritis Cohort. Ann Rheum Dis 2012;71:1366–1369.

10. **van den Berg R**, de Hooge M, Rudwaleit M, Sieper J, van Gaalen F, Reijnierse M, Landewé R, Huizinga T, van der Heijde D. ASAS modification of the Berlin Algorithm for diagnosing axial spondyloarthritis: Results from the SPondyloArthritis Caught Early (SPACE)-cohort and from the Assessment of SpondyloArthritis international Society (ASAS)-cohort. Ann Rheum Dis 2013;72:1646-53.

11. de Hooge M, **van den Berg R**, Navarro-Compán V, van Gaalen F, van der Heijde D, Huizinga T, Reijnierse M. Magnetic resonance imaging of the sacroiliac joints in the early detection of spondyloarthritis: No added value of gadolinium compared to short tau inversion recovery (STIR) sequence. Rheumatology (Oxford) 2013;52:1220-4.

12. **van den Berg R**, de Hooge M, van Gaalen F, Reijnierse M, Huizinga T, van der Heijde D. Percentage of patients with spondyloarthritis in patients referred because of chronic back pain and performance of classification criteria: experience from the SpondyloArthritis Caught Early (SPACE)-cohort. Rheumatology (Oxford) 2013;52:1492-9.

13. **van den Berg R**, Lenczner G, Feydy A, van der Heijde D, Reijnierse M, Saraux A, Rahmouni A, Dougados M, Claudepierre P. Reading of sacroiliac joints on plain pelvic radiographs: agreement between clinical practice and trained central reading. Results of the DESIR-cohort. Arthritis Rheumatol [online 6-6-2014] doi: 10.1002/art.38738.

14. **van den Berg R**, Lenczner G, Thévenin F, Claudepierre P, Feydy A, Reijnierse M, Saraux A, Rahmouni A, Dougados M, van der Heijde D. Classification of axial SpA based on positive imaging (radiographs and/or MRI of the sacroiliac joints) by local rheumatologists or radiologists versus central trained readers in the DESIR-cohort. Ann Rheum Dis [online 24-6-2014] doi: 10.1136/annrheumdis-2014-205432.

15. van Gaalen F, **van den Berg R**, Verhoog I, Schonkeren J, van der Helm-van Mil A, Huizinga T, van der Heijde D. Is HLA-B27 increased in patients diagnosed with undifferentiated arthritis? Results from the Leiden Early Arthritis Cohort. J Rheumatol [online 15-8-2014] doi:10.3899/jrheum.131462.

16. **van den Berg R**, van der Heijde D, Landewé R, van Lambalgen K, Huizinga T, on behalf of the Merit foundation. The METEOR initiative: The way forward for optimal, worldwide data integration to improve care for RA patients. Clin Exp Rheum accepted for publication.

17. **van den Berg R**, de Hooge M, Bakker P, van Gaalen F, Navarro-Compán V, Fagerli K, Landewé R, van Oosterhout M, Ramonda R, Reijnierse M, van der Heijde D. Metric properties of the SPARCC-score of the sacroiliac joints – Data from baseline, 3 and 12 months follow-up in the SPACE-cohort. Submitted.

Curriculum vitae

Rosaline van den Berg werd geboren op 30 januari 1984 in Alkmaar. Na het behalen van haar gymnasium diploma aan het Christelijk Gymnasium in Utrecht in 2002, verhuisde zij naar Maastricht om aan de Universiteit Maastricht de studie Gezondheidswetenschappen te volgen. In 2003 is zij daarnaast gestart met de opleiding Fysiotherapie aan de Hogeschool Zuyd in Heerlen en in 2006 rondde zij deze opleiding succesvol af.

Nadat zij in 2006 ook haar bachelor Gezondheidswetenschappen in de richting Bewegingswetenschappen behaalde, besloot zij Maastricht te verlaten om verder te studeren in de richting Revalidatie en Fysiotherapie binnen de studie Bewegingswetenschappen aan de Vrije Universiteit te Amsterdam, waar zij sinds 2007 ook woont. Tijdens deze master Bewegingswetenschappen heeft zij gewerkt als fysiotherapeute bij het Instituut voor Fysiotherapie in Bunnik.

In 2008 behaalde ze haar master Bewegingswetenschappen, en haar masterthesis heeft geresulteerd in een wetenschappelijke publicatie.

In 2009 is zij gestart met haar promotietraject op de afdeling reumatologie van het Leids Universitair Medisch Centrum, waar zij sinds 2013 met veel plezier een staffunctie bekleedt binnen het METEOR project.

Dankwoord

De rit zit er op, het boekje is af. Ik kijk terug op een prachtige periode, die mede mogelijk is gemaakt door vele mensen. Zonder iemand tekort te doen, wil ik graag een aantal mensen in het bijzonder bedanken, te beginnen met mijn promotor. Prof.dr. D.M.F.M. van der Heijde, Désirée, wat heb ik geboft dat je mij onder je hoede wilde nemen! Ik heb zó ontzettend veel van jou geleerd; niet alleen vakinhoudelijke kennis, maar ook sociale vaardigheden, hoe te handelen in ingrijpende situaties en nog veel meer. Je maakt altijd tijd voor me en bent er met een luisterend oor en adviezen hoe het probleem aan te pakken. Je hebt het beste in mij naar boven gehaald; dankjewel!
Prof.dr. T.W.J. Huizinga, Tom, wat fijn dat je mij de kans hebt gegeven om op de afdeling reumatologie aan de slag te gaan. Bedankt voor je vertrouwen in mij; dat ik dit geweldige werk-avontuur mag voortzetten op jouw afdeling.
Dr. F.A. van Gaalen, Floris, onze samenwerking begon met 'het EAC-stuk' en uiteindelijk ben je mijn co-promotor geworden. Ik heb enorm veel geleerd van jouw recht-door-zee benadering en je was mijn steun en toeverlaat als ik vragen had over patiënten in SPACE of SIAS. Bedankt voor je nimmer aflatende en ontzettend aanstekelijke enthousiasme.
Dr. M. Reijnierse, Monique, als co-promotor van de radiologie heb je mij de vele kneepjes in het beoordelen van röntgenfoto's en MRI's geleerd, waarvoor ik je dankbaar ben.

Naast Désirée en Tom wil ik ook graag de rest van het bestuur van METEOR - prof. dr. R.B.M. Landewé, Robert, ir. K. van Lambalgen, Karel, en mr. drs. P. Bilars, Paul - bedanken voor de kans die jullie mij geboden hebben om onderzoek te blijven doen in de reumatologie. Ik kijk uit naar onze verdere samenwerking.
Anne Marie, ik ben je dankbaar dat ik het stokje van jou heb kunnen overnemen. Jan, dankjewel dat je mij hebt ingewijd in de METEOR database.

Prof.dr. M. Dougados, Maxime, and prof.dr. P. Claudepierre, Pascal, I enjoyed working with you in the DESIR-cohort, and it is my great pleasure to continue working with you.
Prof.dr. J. Braun, Jürgen, and dr. X. Baraliakos, Xenofon, thank you for the stimulating cooperation while working on the update of the ASAS management recommendations.

SPACE begon klein, maar inmiddels loopt het cohort al ruim 5,5 jaar en is er een heel team nodig om het cohort en alles wat erbij komt kijken te runnen. Mijn dank is groot aan alle helpende handen van het (poli-)secretariaat en het datamanagement, met speciale dank aan Jozé, Hughine en Marie-Jose. Zonder jullie hulp was het opzetten van SPACE een stuk moeizamer gegaan!
Manouk, jij was de eerste die het SPACE-team kwam versterken. Ik heb enorm genoten van alle leuke activiteiten die we samen (ook buiten het werk om) hebben ondernomen! Pauline, toen jij ook het SPACE-team kwam versterken liet jij mij inzien dat ik al best veel geleerd had en ik de 'sjaars' best wel wat bij kon brengen. Je bent een waardevolle aanwinst, ook op persoonlijk vlak. Dank voor je vriendschap! Zineb en Miranda, jullie zijn de nieuwste aanwinsten voor het SPACE-team. Stukje bij beetje draag ik steeds meer van de SPACE taken over, maar ik hoop dat we samen nog veel kunnen gaan 'SPACE-n'.
Freek en Lianne, jullie zijn vanuit de radiologie en de MDL ook in meer of mindere mate betrokken bij SPACE, en ook bij SIAS (Freek). Bedankt voor jullie inspirerende samenwerking over de afdelingsgrenzen heen.

Mijn dank gaat ook uit naar alle stafleden en arts-assistenten die patiënten hebben geïncludeerd in SPACE en/of SIAS.
Daarnaast wil ik graag alle - wisselende - collega's van kamer C1-45 en C1-46 bedanken voor alle fijne en ook minder fijne momenten die we samen gedeeld hebben. Rachel, mijn

overbuurvrouw aan ons eiland, jij hebt mij wegwijs gemaakt op de afdeling. Linda, samen in de trein van en naar 'ons dorp', samen hebben we als niet-dokters ons plekje veroverd tussen de dokters op de afdeling. Badelog, wat was het fijn om te weten dat ik niet de enige was met bepaalde frustraties; ik kon altijd mijn hart bij jou luchten. Lotte, jij was er altijd met een luisterend oor en geweldige tips om alles weer goed te 'bakken'. Sweet Victoria, it was a great pleasure to share room C1-46 with you. You are my happy, bright and joking support. I learned a lot from you. Muchas gracias por tu amistad y apoyo.

Sofia en Iris, mijn lieve paranimfen, ik vind het een grote eer dat jullie mij terzijde willen staan. We hebben samen het 'leed' van onze promotietrajecten gedeeld en nu kunnen we samen het 'lief' van mijn promotietraject op deze bijzondere wijze delen.

Al mijn lieve vriendjes en vriendinnetjes van de Hippo's, SKV Amsterdam en Triaz, wil ik graag bedanken voor jullie onvoorwaardelijke vriendschap. Jullie gezelschap was een welkome afleiding van het werken aan dit proefschrift. Jeanine, lief Zeeuws meisje, dankjewel dat ik altijd mijn hart bij je mag uitstorten. Lieve Mandy, je bent altijd geïnteresseerd, je onthoudt altijd alles en je hebt altijd geweldige adviezen, ook in moeilijke tijden.

Liselotte, lieve grote zus, jij bent mijn maatje door dik en dun. Ik geniet van onze wekelijkse theemomentjes als 'collega's' in het LUMC. Zullen we nog vele theemomentjes inlassen?
Lieve papa en mama, dank voor jullie onvoorwaardelijke steun bij alles wat ik doe. Misschien zijn jullie af en toe bevreesd geweest dat er niets van mij terecht zou komen aangezien ik nooit echt een hoogvlieger was, maar ik hoop dat ik jullie nu heb laten zien dat ik gewoon een 'laatbloeier' ben en op mijn pootjes terecht gekomen ben, zolang ik maar iets kan doen wat ik écht leuk vind.
Tot slot wil ik jou, lieve Rikkert, bedanken voor wie je bent. Jij hebt ervoor gezorgd dat de laatste loodjes van het maken van dit proefschrift helemaal niet het zwaarst waren. Jij verrijkt mijn leven en ik kijk uit naar onze toekomst samen.