

Cover Page



Universiteit Leiden



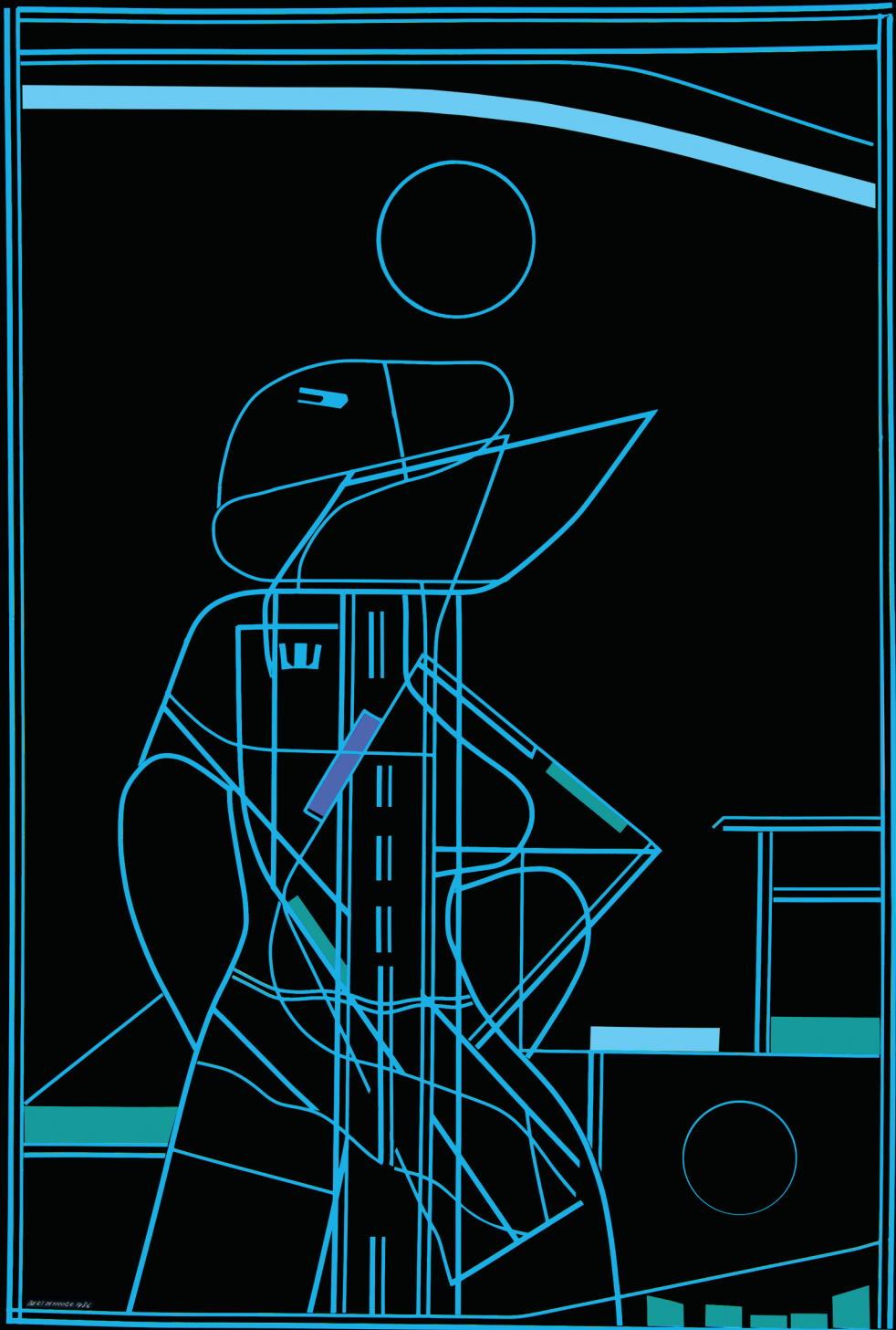
The handle <http://hdl.handle.net/1887/43533> holds various files of this Leiden University dissertation.

Author: Hooge, M.S.M. de

Title: Imaging in axial spondyloarthritis

Issue Date: 2016-10-13

Imaging in axial Spondyloarthritis



Manouk de Hooge

Stellingen behorend bij het proefschrift:

Imaging in axial Spondyloarthritis

1. De ASAS classificatie criteria voor axiale spondyloarthritis zijn ook valide in een setting waarbij de a priori kans op axSpA lager is dan in het cohort waarbinnen de criteria ontwikkeld zijn (dit proefschrift)
2. Het verwijderen van inflammatoire rugpijn (IBP) als verplicht kenmerk verbetert het Berlijn Algoritme als hulpmiddel voor reumatologen in het diagnostisch proces van vroege axSpA (dit proefschrift)
3. De T1 post gadolinium sequentie heeft geen toegevoegde waarde bij het beoordelen van MRI's om sacroiliitis vast te stellen in het kader van de ASAS axiale spondyloarthritis criteria
4. 'The rule of five', die refereert naar minstens vijf vet laesies en/of erosies op MRI van de SI-gewrichten, minstens vijf inflammatoire laesies en minstens vijf vet laesies op de MRI van de wervelkolom, is nuttig om onderscheid te maken tussen patiënten met en zonder axiale spondyloarthritis (dit proefschrift)
5. Afwijkingen op de röntgenfoto van de wervelkolom, kenmerkend voor axiale spondyloarthritis, worden door radiologen en reumatologen in de lokale ziekenhuizen overschat ten opzichte van getrainde beoordelaars (dit proefschrift)
6. Radiologen en reumatologen in de lokale ziekenhuizen zijn onzeker over de aan- of afwezigheid van inflammatoire MRI laesies in de wervelkolom
7. Inflammatoire SpA afwijkingen in de SI-gewrichten zijn geassocieerd met bilpijn
8. MRI heeft in belangrijke mate bijgedragen aan de mogelijkheid niet-radiologische axiale spondyloarthritis vroeg op te sporen
9. Het belangrijkste is dat je in elke ronde het beste van jezelf hebt gegeven. ('L'importante è avere dato il meglio di sé in ogni singolo giro') (Marco Simoncelli, jong overleden MotoGP coureur en wereldkampioen wegrace 250cc)
10. When you aim for perfection, you'll discover it's a moving target (Fernando Alonso, tweevoudig wereldkampioen Formule 1)
11. Van een stimulusrijke naar een stimulusarme omgeving verhuizen vergt moed, maar kan leiden tot meer zelfinzicht en eindigen in persoonlijke groei

Imaging in axial Spondyloarthritis

**Manouk de Hooge
2016**

ISBN: 978-94-6332-050-4

Copyright © 2016 by Manouk de Hooge.

All rights reserved. No part of this book may be reproduced in any form without written permission of the author or, when appropriate, of the publishers of the publications.

Printing of thesis was financially supported by Pfizer B.V., UCB B.V., ABBVIE B.V., which are gratefully acknowledged.

Cover design: Bert (Lambertus Cornelis Hendrik) de Hooge & Menno Simons

Printing: GVO drukkers & vormgevers B.V.

Imaging in axial Spondyloarthritis

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 donderdag 13 oktober 2016
klokke 11:15 uur

door

Manouk Sumudu Madushani de Hooge

geboren te Negombo (Sri Lanka)
in 1986

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

Co-promotores: Dr. F. Van Gaalen
Dr. M. Reijnierse

Promotiecommissie: Prof. dr. T.W.J. Huizinga
Prof. dr. G. Kloppenburg
Prof. dr. J.L. Bloem
Prof. dr. A.E.R.C.H. Boonen
Maastricht Universitair Medisch Centrum in Maastricht
Prof. dr. F van den Bosch
Universitair Ziekenhuis Gent in Gent

Content

Chapter 1 General introduction

Compatibility axSpA criteria sets

Chapter 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.

Chapter 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.

Imaging

Chapter 4 Magnetic resonance imaging of the sacroiliac joints in the early detection of spondyloarthritis: no added value of gadolinium compared with short tau inversion recovery sequence.

Chapter 5 Patients with chronic back pain of short duration from the SPACE cohort: which MRI structural lesions in the sacroiliac joints and inflammatory and structural lesions in the spine are most specific for axial spondyloarthritis?

Chapter 6 Reliability of mSASSS scoring in everyday practice in DESIR cohort study centres: cross-sectional study of agreement with trained readers.

Chapter 7 Assessment of typical SpA lesions on MRI of the spine: do local readers and central readers agree in the DESIR cohort at baseline?

Chapter 8	Is site of back pain related to location of MRI lesion in patients with chronic back pain included in the SPACE cohort?
Chapter 9	Summary and conclusions
Chapter 10	Nederlandse samenvatting
	List of publications
	Curriculum Vitae
	Dankwoord

1

General introduction

Introduction

Spondyloarthritis (SpA) is a common inflammatory rheumatic disease with a heterogeneous presentation. The prevalence of patients with SpA is estimated at 1% of the general population¹⁻⁴. There are two major SpA subgroups; one subgroup with predominantly inflammation in peripheral joints (peripheral SpA) and one subgroup of patients with predominantly inflammation of the axial skeleton including the spine and sacroiliac (SI) joints (axial spondyloarthritis (axSpA))⁵. In this thesis the focus is on the axSpA subgroup.

AxSpA is characterized by several features (referred to as SpA features), which mainly are related to inflammation. Typically, multiple features are present in a patient with axSpA. The composition of features is not set and can differ between patients. SpA features can be assessed using patient history, physical examination, laboratory testing and imaging. Clinically, a common complaint in patients with axSpA is inflammatory back pain (IBP). IBP is a symptom complex rather than a condition. On physical examination, peripheral arthritis, heel pain suggestive of enthesitis and dactylitis are also frequently present in patients with axSpA⁶. Extra-articular manifestations include anterior uveitis, inflammatory bowel disease (IBD) and psoriasis, and may occur in a substantial portion of the patients⁷⁻⁹. In addition, a good response -reduced back pain- to non-steroidal anti-inflammatory drugs (NSAIDs) and a positive family history for SpA or related diseases are clinical characteristics of axSpA patients and can be derived from patient history¹⁰. Presence of human leukocyte antigen B27 (HLA-B27), elevated C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR) are laboratory features associated with axSpA¹¹⁻¹⁴. Imaging features are those related to sacroiliitis on radiographs and magnetic resonance imaging (MRI) (detailed description is given later)¹⁵⁻¹⁷.

Recognising axSpA in an early stage is challenging, because patients can present with any combination and number of SpA features¹⁸. Furthermore, there are no diagnostic criteria available. Diagnosing axSpA largely depends on the combination of SpA features and the exclusion of other diagnoses. Therefore, a systematic approach to diagnosis and clinical experience of the treating physician are of importance to get a unique and complete clinical view on the patient^{10,18}. Several tools that can guide physicians in their diagnostic process of (early) axSpA have been developed and may help increase the confidence of the physicians' diagnosis.

Classification criteria and tools helpful for diagnosing axSpA

Ideally, a physicians' diagnosis is based on the (presumed) etiopathogenesis and clinical presentation (including history and results of lab- and imaging studies) of the patient. Though a diagnosis is customized to the patient and therefore of individual nature, generalising patients by means of classification is important to show the similarities between axSpA patients and to create groups for research that are as well-defined as possible.

Likely, the most well-known classification criteria in the field of axSpA are the modified New York (mNY) criteria, developed in 1984. With these criteria the most typical form of axSpA, ankylosing spondylitis (AS), can be classified. This criteria set combines clinical symptoms with radiographic sacroiliitis. In the mNY criteria radiographic sacroiliitis is defined as sacroiliitis grade ≥ 2 bilaterally or grade 3-4 unilaterally. Patients could be classified as 'definite AS' when in addition to radiographic sacroiliitis they have at least one of the following clinical criteria: 1) low back pain and stiffness for >3 months with improvement during exercise but not during rest, 2) limited motion of the lumbar spine in sagittal as well as frontal planes or 3) limited chest expansion compared to age- and sex-related normal values¹⁵.

Conventional imaging of the pelvis has long been the imaging modality of choice in the clinical assessment of patients with axSpA^{15,19}. However, since radiographs can only detect structural changes, initial imaging (radiography) at the onset of patients' complaints is often normal. In the literature a diagnostic delay up to 6 and 8 years is described before patients developed sacroiliitis which is detected on radiographs^{13,20}. On top of that, it is not likely that all axSpA patients with predominant axial involvement will develop radiographic sacroiliitis^{21,22}. Thus, axSpA can be subdivided into two groups; radiographic axSpA (r-axSpA) -which is synonymous with AS- and nonradiographic axSpA (nr-axSpA). The difference between these two groups is based on the presence or absence of abnormal SI joints on radiographs¹⁰.

In the early nineties the Amor and European Spondyloarthropathy Study Group (ESSG) criteria were developed attempting to classify patients with axial involvement including those without definite radiographic changes. At that time the term 'nr-axSpA patients' was not yet commonly used. Later on this term became in vogue and it describes the same patients without definite radiographic changes that are also included in the Amor and ESSG criteria. In both criteria sets radiographic sacroiliitis is included but not mandatory and -in contrast to the mNY- radiographic sacroiliitis is not an entry criterion. As a result both criteria sets cover radiographic and non-radiographic axSpA.

In the Amor criteria there are no entry criteria at all. The Amor criteria are based on a point system where a score from 1 to 3 is assigned to characteristics associated with axSpA and at least a total of 6 points are necessary to classify patients²³. In the ESSG criteria the priority was given to clinical characteristics -IBP or synovitis (asymmetric or predominantly present in the lower limbs)- as entry criterion over radiological variables²⁴. Although the ESSG criteria were much cited they did not perform that well in daily practice. Some studies show that not even half of the patients that were classified by the ESSG criteria at baseline were diagnosed as SpA patient by a rheumatologist after five years^{25,26}.

The Amor and ESSG criteria focused on the total, heterogeneous spectrum of SpA, including patients without radiographic sacroiliitis, and the absence of radiographic sacroiliitis no longer implied absence of the disease. It was suggested that inflammation itself induces structural changes, even though the exact course of immunopathological events remains unclear²⁷⁻²⁹.

Coinciding with the increasing interest in nr-axSpA patients, MRI has been of upcoming focus in the field of axSpA. In contrast to conventional imaging, MRI is an imaging modality that can readily detect active disease indicated by inflammatory lesions, in addition to structural lesions. MRI appears to be a promising technique for detection of axSpA in an early phase. In the Berlin Algorithm, published in 2004, MRI was included as imaging modality. The Berlin algorithm is a tool to guide rheumatologists in their diagnostic process of (early) axSpA. The algorithm recommends in which patients it is reasonable to perform HLA-B27 testing and an MRI in order to increase the confidence of the physician's diagnosis. This instrument was developed based on the sensitivity and specificity of SpA features reported in the literature¹⁰. With the Berlin algorithm an approach was developed that potentially limits unnecessary diagnostic tests, which thereby also decreases costs, time and hassle for both patients and society.

Although the Berlin Algorithm included the use of MRI in the diagnostic process of the physician, it was only until 2009 that actual classification criteria include MRI. In 2009, an international group of experts in the field of SpA developed the Assessment of Spondyloarthritis international Society (ASAS) axSpA classification criteria³⁰. With the expert physician's diagnosis as gold standard the criteria set performed well with a sensitivity of 82.9% and specificity of 84.4%. The ASAS group specified that the classification criteria should be applied only in patients with (almost) daily back pain for three or more months, with an onset before the age of 45 years and in whom the origin of back pain was unidentified^{30,31}. In addition, the ASAS group stated that since

the study in which the criteria set was developed, was conducted in a rheumatology setting with a high prevalence of axSpA patients, the ASAS criteria should be used as classification criteria rather than as a diagnostic tool.

With the presentation of the ASAS criteria, MRI was officially introduced in the process of classifying axSpA patients, since there are two ways -referred to as 'arms'- to fulfil the ASAS axSpA classification criteria. One arm is generally referred to as the 'clinical arm' and the other arm is referred to as the 'imaging arm'. Patients who are HLA-B27 positive in combination with two additional SpA features are classified as axSpA according to the 'clinical-arm' of the ASAS axSpA classification criteria. The individual SpA features besides HLA-B27 positivity are IBP, peripheral arthritis, heel pain (enthesitis), dactylitis, anterior uveitis, IBD, psoriasis, good response to NSAIDs and a positive family history for SpA or related diseases and elevated CRP or ESR levels. IBP should be highlighted because in the ASAS classification criteria IBP is defined according to the ASAS IBP criteria³² and all SpA features have clear definitions before they can be counted as such³⁰.

Patients fulfilling the criteria via the imaging arm have 'sacroiliitis on imaging' with one additional SpA feature. 'Sacroiliitis on imaging' does not only refer to radiographic sacroiliitis -according to the mNY criteria- but can also indicate the presence of inflammation on MRI of the SI joints (MRI-SI), in the form of 'a positive MRI-SI'³³. When specific criteria are met, a MRI-SI in which inflammation is shown can be referred to as 'a positive MRI'³¹. Subchondral bone marrow edema seen on MRI is highly suggestive of sacroiliitis and the sole presence of synovitis, capsulitis or enthesitis (without concomitant subchondral bone marrow oedema/osteitis) is insufficient for a positive MRI-SI. Furthermore, one inflammatory lesion should be visible on at least two consecutive MRI slices or more lesions should be present on a single slice to mark the MRI-SI as positive. The ASAS group also provides technical recommendations and scanning parameters considering the obtainment of MRI-SI. The images obtained with the Short Tau Inversion Recovery (STIR) sequence are considered sufficient to detect inflammatory lesions, apart from synovitis. In addition, T1-weighted turbo spin-echo sequence should be taken into consideration when judging the presence of inflammation. Synovitis can be detected when using T1 fat-saturated post gadolinium (T1/Gd) images³¹. It should be noted that the use of a paramagnetic contrast agent as gadolinium has specific disadvantages. The administration is intravenously and therefore uncomfortable for patients. Also exposure to this contrast agent is not without risks, in particular in patients with decreased kidney function³⁴.

In the ASAS criteria MRI is appraised equal to radiographs. This shows the growing importance of MRI in the field of axSpA. However, it is unclear if the current form of the definition of ‘positive imaging’ in the ASAS criteria is the optimal form. In addition to inflammatory lesions, structural lesions can also be detected on MRI. Despite the possibility to detect structural lesions on MRI, only inflammatory MRI-SI lesions were included in the ASAS criteria pursuant to the lack of evidence on the utility of structural lesions on MRI-SI in the diagnostic process of (early) axSpA.

MRI is also a useful imaging modality to visualise inflammatory and structural lesions (fatty lesions, erosions and syndesmophytes) of the spine of patients with axSpA³⁵. Information on the diagnostic utility of lesions seen on the MRI of the spine (MRI-spine) is limited. However, such lesions, typical for axSpA, could have substantial clinical and maybe even prognostic relevance.

The information above has resulted in the following research questions elucidated in this thesis:

1. To review different criteria sets for axSpA. Are we doing well or is improvement needed?
2. To find the additional value and utility of structural lesions on MRI-SI and the use of MRI-spine for classifying axSpA patients.
3. What is the reliability of the assessment of spinal axSpA lesions?

To answer these questions, we used data from two cohorts in which patients with recent onset of chronic back pain at a reasonably young age were included: The DEvenir des Spondylarthropathies Indifférenciées Récentes (DESIR) and Spondyloarthritis Caught Early (SPACE) cohorts. Both cohorts have similarities but also have a few important differences.

In the **DESIR cohort** patients between 18 and 50 years old were included from 25 participating centers across France when patients had inflammatory back pain (thoracic, lumbar or buttock region) according to the Calin³⁶ or Berlin criteria³⁷ for ≥ 3 months but < 3 years. In addition, the symptoms of the patient should be indicated as moderately to highly suggestive of SpA according to the local rheumatologist (a score of ≥ 5 on a numerical rating scale from 0 to 10). Specific exclusion criteria as well as a detailed description of the DESIR cohort have been published previously³⁸.

The **SPACE cohort** was launched in January 2009 at the Leiden University Medical Center (LUMC). It is an ongoing project in which patients who are referred to a rheumatologist are included when they are ≥ 16 years, have (almost) daily chronic back pain for ≥ 3 months but ≤ 2 years, with an onset of back pain < 45 years. The origin of

back pain ought to be unknown. Reasons to exclude patients, besides not meeting the criteria above, are 1) the presence of other painful conditions not associated with axSpA that could interfere with the assessment or evaluation of disease activity and 2) any reason that would invalidate the informed consent or limit the ability of the participant to comply with the protocol requirements. Five European centres located in the Netherlands (Leiden, Gouda and Amsterdam), Norway (Oslo) and Italy (Padova) included patients in the SPACE cohort.

Patients included in either the DESIR or SPACE cohort underwent a diagnostic workup, according to a predefined protocol of the corresponding cohort, ensuring the obtainment of all SpA features³⁰ as well as other characteristics.

Compatibility axSpA criteria sets

With the development of the ASAS classification criteria, the question arose how this criteria set relates to other criteria sets developed to cover the entire SpA spectrum. The ASAS criteria include two ‘arms’ through which patients can be classified as having axSpA³⁰. Since one arm includes sacroiliitis on imaging and the other is limited to only clinical features in addition to HLA-B27 positivity, there was uncertainty about the similarity of the patients’ disease phenotype in the two arms. Besides a detailed description of the SPACE cohort, the performance of the ASAS criteria and the resemblance between patients classified via the clinical- and imaging-arm are described in **chapter 2**.

IBP is one of the characteristics of axSpA and is considered so important that it is recommended as referral parameter in primary care³⁹. Even though IBP is a strong characteristic of axSpA, there are also patients with IBP who do not have axSpA, as well as axSpA patients without IBP^{32,40}. This leads to the question if the Berlin Algorithm in its original form, in which IBP is an entry criterion, is optimal in guiding rheumatologists in diagnosing axSpA patients. In **chapter 3** the performance of two separate modifications of the Berlin Algorithm are compared to the original algorithm.

Imaging

It is questionable whether there is added value to the use of T1/Gd over STIR images in detecting inflammatory lesions on the MRI-SI, since there might be considerable overlap between the T1/Gd and STIR sequences⁴¹. In **chapter 4** both sequences are compared to each other and the presumed additional value of T1/Gd in detecting inflammatory lesions is investigated in patients included in the SPACE cohort.

Even though several lesions associated with axSpA can be detected on both MRI-SI as well as MRI-spine, the only type of lesions included in the ASAS classification criteria for axSpA are inflammatory lesions. Are structural MRI-SI lesions and spinal MRI lesions only seen in axSpA patients or also in patients with back pain who do not fulfil the ASAS criteria? In **chapter 5** the prevalence of MRI-SI and MRI-spine lesions is described. In addition, the type and quantity of lesions that are most specific for axSpA are provided.

The DESIR cohort has a unique design characteristic where both a locally based specialist as well as central readers give a quantified score of lesions typical for axSpA seen on imaging. This provides the opportunity to investigate whether the scores given by a local reader -as is done in clinical practice- match the scores of central readers with specific training in recognizing lesions typical for axSpA.

Syndesmophytes are typical bony spurs in patients with axSpA that can be detected on radiographs of the spine. The quantification of these abnormalities can be used to assess the severity and progression of structural damage. The most reliable, sensitive and therefore recommended tool to assess radiographic changes in the spine is the modified Stoke Ankylosing Spondylitis Spine Score (mSASSS)^{42,43}. With this scoring method the spinal damage is quantified, by evaluating the anterior vertebral corners of the cervical and lumbar part of the spine on the presence of erosions, squaring, sclerosis and bridging or non-bridging syndesmophytes. **Chapter 6** focuses on the radiographs of the spine where the mSASSS assigned by local readers is compared to the mSASSS assigned by central readers. The agreement on the presence of MRI lesions in the spine between local and central readers is discussed in **chapter 7**.

Chapter 8 evaluates the association between the localization of axial pain and the site of MRI lesions. Lesions on MRI can be seen in patients with axSpA but are also present as a consequence of degenerative changes. It is not clear if axSpA lesions and/or degenerative lesions on MRI at a certain location result in pain at that same location. Therefore, we obtained information on both axSpA lesions and degenerative lesions and related this to pain at various sites in the spine and buttocks. When doing this, it is important to correct for the dependency within a single patient. So, in **chapter 8** the association between location of pain and MRI lesions was investigated in chronic back pain patients from the SPACE cohort, taking into consideration the correlation within a person by performing generalized estimated equation analyses.

In **chapter 9**, this thesis concludes with a summary and general discussion of the results and the influence they might have on future perspectives and analyses. A Dutch summary of this thesis is available in **chapter 10**.

References

- 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.
- Saraux A, Guillemin F, Guggenbuhl P, et al. Prevalence of spondyloarthropathies in France: 2001. *Ann Rheum Dis* 2005;64:1431-5.
- Adomavičiute D, Pileckyte M, Baranauskaite A, et al. Prevalence survey of rheumatoid arthritis and spondyloarthropathy in Lithuania. *Scand J Rheumatol* 2008;37:113-9.
- Helmick CG, Felson DT, Lawrence RC, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part I. *Arthritis Rheum* 2008;58:15-25.
- Dougados M, Baeten D. Spondyloarthritis. *Lancet* 2011;377:2127-37.
- Rudwaleit M. New approaches to diagnosis and classification of axial and peripheral spondyloarthritis. *Curr Opin Rheumatol* 2010;22:375-380.
- Rudwaleit M, Haibel H, Baraliakos X, et al. The early disease stage in axial spondylarthritis: results from the German Spondylarthritis Inception Cohort. *Arthritis Rheum* 2009;60:717-27.
- 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.
- Elewaut D, Matucci-Cerinic M. Treatment of ankylosing spondylitis and extra-articular manifestations in everyday rheumatology practice. *Rheumatology (Oxford)* 2009;48:1029-35.
- Rudwaleit M, van der Heijde D, Khan MA, et al. How to diagnose axial spondyloarthritis early. *Ann Rheum Dis* 2004;63:535-43.
- 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 Rheumatol* 1999;26:980-4.
- Dougados M, Gueguen A, Nakache JP, et al. Clinical relevance of C-reactive protein in axial involvement of ankylosing spondylitis. *J Rheumatol* 1999;26:971-4.
- Feldtkeller E, Khan MA, 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.
- Reveille JD. Genetics of spondylarthritis—beyond the MHC. *Nat Rev Rheumatol* 2012;8:296-304.
- 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.
- van Onna M, Jurik AG, 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.
- Heuft-Dorenbosch L, Landewé R, Weijers R, et al. Combining information obtained from magnetic resonance imaging and conventional radiographs to detect sacroiliitis in patients with recent onset inflammatory back pain. *Ann Rheum Dis* 2006;65:804-8.
- van Tubergen A, Weber U. Diagnosis and classification in spondyloarthritis: identifying a chameleon. *Nat Rev Rheumatol* 2012;8:253-61.
- Oostveen J, Prevo R, den Boer J, et al. Early detection of sacroiliitis on magnetic resonance imaging and subsequent development of sacroiliitis on plain radiography. A prospective, longitudinal study. *J Rheumatol* 1999;26:1953-8.
- Khan MA. Ankylosing spondylitis: introductory comments on its diagnosis and treatment. *Ann Rheum Dis* 2002;61:iii3-iii7.
- Rudwaleit M, Khan MA, Sieper J. The challenge of diagnosis and classification in

- early ankylosing spondylitis: Do we need new criteria? *Arthritis Rheum* 2005;52:1000-8.
22. Bennett AN, McGonagle D, O'Connor P, et al. Severity of baseline magnetic resonance imaging-evident sacroiliitis and HLA-B27 status in early inflammatory back pain predict radiographically evident ankylosing spondylitis at eight years. *Arthritis Rheum* 2008;58:3413-8.
 23. Amor B, Dougados M, Mijiyawa M. Critères de classification des spondylarthropathies. *Rev Rhum Mal Osteoartic* 1990;57:85-9.
 24. Dougados M, van der Linden S, Juhlin R, et al. The European Spondyloarthritis Study Group preliminary criteria for the classification of spondyloarthritis. *Arthritis Rheum* 1991;34:1218-27.
 25. Amor B. Usefulness of criteria for spondyloarthritis. *Joint Bone Spine* 2000;67:502-3.
 26. Collantes E, Veroz R, Escudero A, et al. Can some cases of "possible" spondyloarthritis be classified as "definite" or "undifferentiated" spondyloarthritis? Value of criteria for spondyloarthritis. Spanish Spondyloarthritis Study Group. *Joint Bone Spine* 2000;67(6):516-20.
 27. 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.
 28. Walsh NC, Manning CA, Boch JA et al. Wnt signaling antagonists are candidates for the suppression of osteoblast differentiation in inflammatory arthritis. *Arthritis Rheum* 20006;54:S511.
 29. Zwerina J, Tuerk B, Redlich K, et al. Imbalance of local bone metabolism in inflammatory arthritis and its reversal upon tumor necrosis factor blockade: direct analysis of bone turnover in murine arthritis. *Arthritis Res Ther* 2006;8:R22.
 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, Rudwaleit M, Baraliakos X, et al. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. *Ann Rheum Dis* 2009;68:ii1-ii44.
 32. 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.
 33. 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.
 34. Bellin MF, van der Molen AJ. Extracellular gadolinium-based contrast media: An overview. *Eur J Radiol* 2008;66:160-7.
 35. Maksymowych WP, Lambert RGW. Magnetic resonance imaging for spondyloarthritis--avoiding the minefield. *J Rheumatol* 2007;34:259-65.
 36. Calin A, Porta J, Fries JF, et al. Clinical history as a screening test for ankylosing spondylitis. *JAMA* 1977;237:2613-14.
 37. 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.
 38. Dougados M, d'Agostino MA, Benessiano J, et al. The DESIR cohort: A 10-year follow-up of early inflammatory back pain in France: Study design and baseline characteristics of the 708 recruited patients. *Jt Bone Spine* 2011;78:598-603.
 39. Braun A, Saracbas 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;70:1782-87.
 40. 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.
41. Madsen KB, Schiøttz-Christensen B, Jurik AG. Prognostic significance of magnetic resonance imaging changes of the sacroiliac joints in spondyloarthritis--a followup study. *J Rheumatol* 2010;37:1718-27.
42. Wanders AJB, Landewé RBM, Spoorenberg A, et al. What is the most appropriate radiologic scoring method for ankylosing spondylitis? A comparison of the available methods based on the outcome measures in rheumatology clinical trials filter. *Arthritis Rheum* 2004;50:2622-32.
43. Creemers MCW, Franssen MJAM, van 't Hof MA, et al. Assessment of outcome in ankylosing spondylitis: an extended radiographic scoring system. *Ann Rheum Dis* 2005;64:127-9.

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

Rosaline van den Berg, Manouk de Hooge, Floris van Gaalen, Monique Reijnerse,
Tom Huizinga, Désirée van der Heijde

Rheumatology (Oxford) 2013;52(8):1492-9

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.

Conclusions

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 IBD¹. 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 AS², however, they are of limited use in early disease or other subtypes of SpA³. 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 features⁶. 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 axSpA⁸. 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 feature⁹. Hence some have stated that not just chronic back pain, but specific inflammatory back pain (IBP) is typical of axSpA¹⁰. 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.

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 definition¹⁶), (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 (R.v.d.B. or M.d.H.) 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) index¹⁹. 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 handbook²⁰. 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 calculated²¹.

Patient-reported questionnaires

Patients completed the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)²² and Bath Ankylosing Spondylitis Functional Index (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 criteria⁶ was recorded. Treatment with NSAIDs, DMARDs and biologic therapies was recorded. NSAID intake is recorded according to the ASAS recommendations²⁴. A good response of back pain to a full dose of NSAID was defined as not present anymore or much better⁶. Furthermore, the Ankylosing Spondylitis Disease Activity Score (ASDAS) was calculated²⁵. 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 definition²⁶, according to the Spondyloarthritis Research Consortium of Canada (SPARCC) score²⁷ and on the presence of capsulitis/enthesitis. All radiographs of the SI joints (X-SIs) were scored according to the modified mNY criteria². 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 as 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 (s.d. 12.6) years, the mean symptom duration was 13.5 (s.d. 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) (Fig. 1).

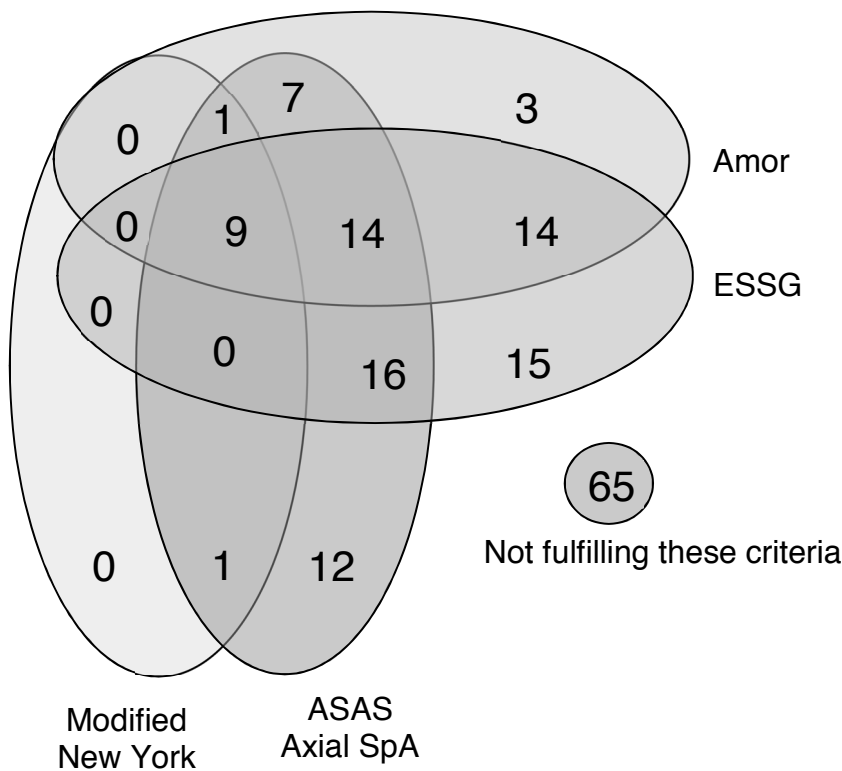


Figure 1: Venn-diagram representing the overlap between the various classification criteria for axial SpA

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-	Confidence about the diagnosis axSpA, mean (SD)
ASAS axSpA	55 (84.6)	87 (94.6)	15.6	0.16	6.4 ± 1.8
mNY	11 (16.9)	92 (100)	15.6	0.99	7.8 ± 1.1
ESSG	42 (64.6)	66 (71.7)	2.3	0.49	6.2 ± 1.7
Amor	31 (47.7)	75 (81.5)	2.6	0.64	6.5 ± 1.7
Modified ESSG (with MRI)	49 (75.4)	66 (71.7)	2.7	0.34	6.4 ± 1.7
Modified Amor (with MRI)	44 (67.7)	75 (81.5)	3.7	0.40	6.7 ± 1.6

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 no-axSpA 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 vs 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 $\pm$ SD	29.5 $\pm$ 8.7	32.3 $\pm$ 14.4	0.17
Male, n (%)	29 (48.3)	23 (23.7)	0.001
Duration of back pain (months), mean $\pm$ SD	13.4 $\pm$ 7.7	13.6 $\pm$ 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 $\pm$ SD	8.4 $\pm$ 11.9	5.8 $\pm$ 6.9	0.12
ESR (mm/h), mean $\pm$ SD	14.4 $\pm$ 16.7	10.1 $\pm$ 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 balanitis, 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- vs clinical-arm
	mNY+ n=11	mNY- n=19	Total n=30		
Age (years) at inclusion, mean $\pm$ SD	28.6 $\pm$ 9.6	32.9 $\pm$ 8.7	31.2 $\pm$ 9.0	28.2 $\pm$ 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 $\pm$ SD	15.6 $\pm$ 8.5	16.0 $\pm$ 6.9	15.5 $\pm$ 7.6	11.4 $\pm$ 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 $\pm$ SD	6.9 $\pm$ 7.2	7.6 $\pm$ 8.6	7.3 $\pm$ 8.0	9.4 $\pm$ 14.9	0.58
ESR (mm/h), mean $\pm$ SD	11.4 $\pm$ 13.9	14.2 $\pm$ 14.8	13.2 $\pm$ 14.3	15.6 $\pm$ 18.9	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 $\pm$ 1.8	4.0 $\pm$ 2.5	3.9 $\pm$ 2.3	3.9 $\pm$ 1.9	0.97
ASDAS	2.4 $\pm$ 0.7	2.5 $\pm$ 0.9	2.4 $\pm$ 0.8	2.4 $\pm$ 0.9	0.94
BASFI	3.3 $\pm$ 1.9	2.4 $\pm$ 2.2	2.7 $\pm$ 2.1	2.3 $\pm$ 2.2	0.50
BASMI	1.9 $\pm$ 0.7	1.6 $\pm$ 0.5	1.7 $\pm$ 0.6	1.6 $\pm$ 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 about the diagnosis axSpA, mean $\pm$ SD	7.8 $\pm$ 1.1	7.5 $\pm$ 0.6	7.7 $\pm$ 0.8	4.9 $\pm$ 1.5	<0.001

[†] Statistical 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 balanitis, 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 pain²⁹. 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 $\sim 60\%$ of the patients in the SPACE cohort fulfil 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 multicentre 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 symptom¹². 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 rheumatologists¹¹.

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 patients¹³. 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 study¹¹. 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 fulfil 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 sacroiliitis³⁰. 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 fulfil 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 fulfil the Amor criteria more easily.

To fulfil 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 fulfil 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 tool³². 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 results³³. 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.

Online supplementary data

Supplementary data are available at Rheumatology Online.

References

1. Khan MA. Update on spondyloarthropathies. *Annals of Internal Medicine* 2002;136:896-907.
2. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. *Arthritis Rheum* 1984;27:361-8.
3. Rudwaleit M, Landewe R, van der Heijde D, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part I): classification of paper patients by expert opinion including uncertainty appraisal. *Ann Rheum Dis* 2009;68:770-6.
4. Dougados M, Vanderlinden 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. Amor B, Dougados M, Mijiyawa M. Classification Criteria of Spondyloarthropathies. *Revue du Rhumatisme* 1990;57:85-9.
6. Rudwaleit M, van der Heijde D, Landewe R, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann Rheum Dis* 2009;68:777-83.
7. Rudwaleit M, van der Heijde D, Landewe 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 MA, 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, Saracbası 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 (Oxford)* 2009;48:404-9.
14. Rudwaleit M, van der Heijde D, Khan MA, 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, Landewe 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, Landewe 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 MA, Benessiano J, et al. The DESIR cohort: a 10-year follow-up of early inflammatory back pain in France: study design and baseline characteristics of the 708 recruited patients. *Joint Bone Spine* 2011;78:598-603.
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:ii1-44.
21. van der Heijde D, Landewe R, Feldtkeller E. Proposal of a linear definition of the Bath Ankylosing Spondylitis Metrology Index (BASMI) and comparison with the 2-step and 10-step definitions. *Ann Rheum Dis* 2008;67:489-93.
22. Garrett S, Jenkinson T, Kennedy LG, et al. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. *J Rheumatol* 1994;21:2286-91.
23. 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, Landewe 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 AG, Hermann KG, et al. Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. *Ann Rheum Dis* 2009;68:1520-7.
27. Maksymowych WP, Inman RD, 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 HS, Schouten JS. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC(3)-study. *Pain* 2003;102:167-78.
29. Deyo RA, Rainville J, Kent DL. What can the history and physical examination tell us about low back pain? *JAMA* 1992;268:760-5.
30. Feldtkeller E, Khan MA, 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 PP, 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 Hooze 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 Suppl 3:332.

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

Rosaline van den Berg*, Manouk de Hooge*, Martin Rudwaleit, Joachim Sieper, Floris
van Gaalen, Monique Reijnerse, Robert Landewé, Tom Huizinga, Désirée van der
Heijde

*Both authors contributed equally

Ann Rheum Dis 2013;72(10):1646-53

Abstract

Objective

To compare the original Berlin algorithm for diagnosing axial SpA (axSpA) with two modifications in the SPondyloArthritis Caught Early (SPACE) cohort and the 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: the original (inflammatory back pain (IBP) mandatory), modification 1 (IBP defined by $\geq 3/5$ IBP-features instead of $\geq 4/5$) and modification 2 (IBP deleted as obligatory entry criterion, added as SpA-feature). Diagnosis by rheumatologist, ASAS axSpA criteria and LR-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 (from 83.7% to 78.3%) than when using modification 2 (from 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%).

Conclusion

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-SIJ) and non-radiographic axSpA. The heterogeneity of SpA makes early detection challenging¹. 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 visiting the rheumatology outpatient clinic in the Leiden University Medical Centre, 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-SIJ and X-SIJ) 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-SIJ and X-SIJ were independently scored by two trained readers (MdH and RvdB) according to the ASAS/OMERACT definition (MRI-SIJ)⁶, and the modified New York (mNY) criteria (X-SIJ)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 before⁸. This included assessment of (past or current) SpA-features², 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-SIJ (mNY criteria)⁹, and the presence or absence of typical signs of active inflammation on MRI-SIJ⁸.

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-SIJ) were present.

In the original algorithm, IBP was defined according to the Calin criteria¹⁰. In the SPACE cohort and ASAS cohort, however, IBP was defined according to the ‘ASAS expert criteria’, which are slightly more specific⁵.

Subsequently, two modifications of the algorithm were constructed. In modification 1, fulfilment of the ASAS IBP criteria¹¹, 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 up⁵. 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 $\geq 4/5$ criteria, a higher sensitivity (95.1%) was reached at the cost of specificity (47.5%) if $\geq 3/5$ 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 fulfilment of the ASAS axSpA criteria⁷, the disease probability based on the likelihood ratio (LR)-product¹³, 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.

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.

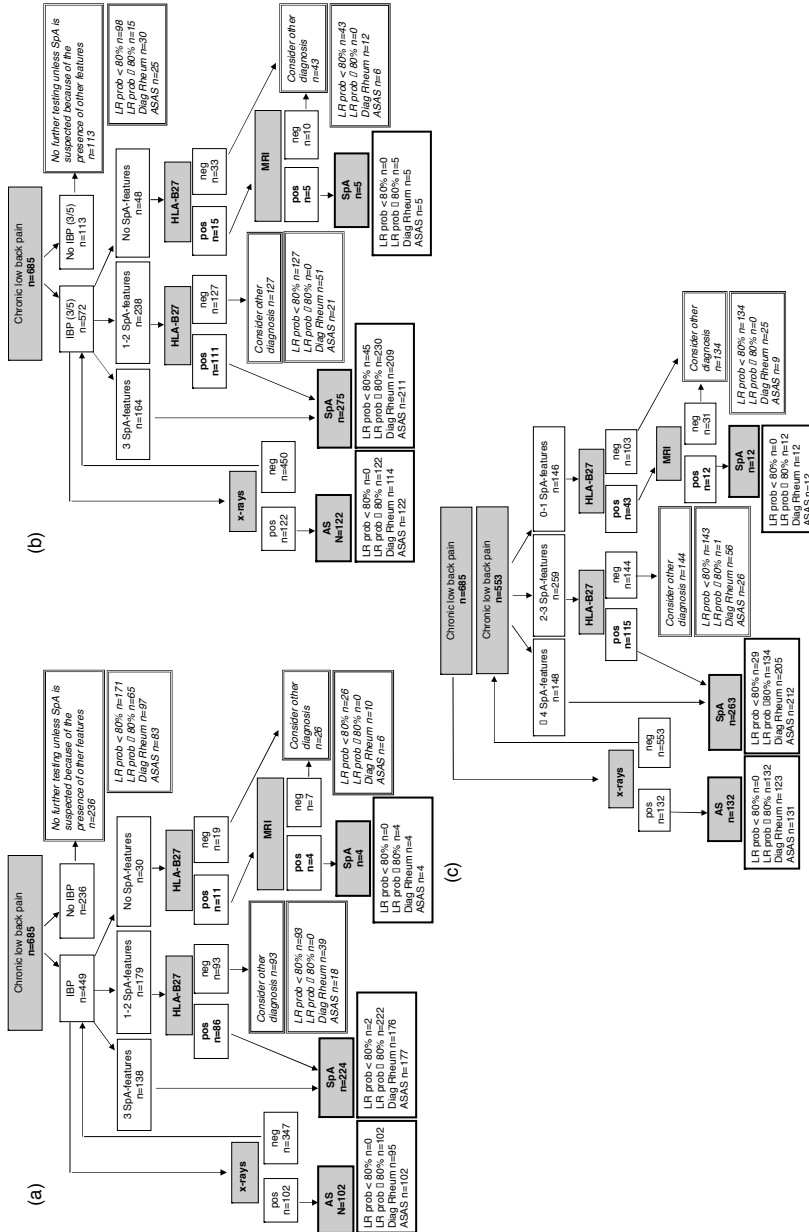


Figure 2: (a) Original Berlin algorithm (Assessment of SpondyloArthritis international Society (ASAS) cohort). (b) Modification 1 of the Berlin algorithm; IBP defined when 3 out of 5 criteria are fulfilled instead of 4 out of 5 criteria (ASAS cohort). (c) Modification 2 of the Berlin algorithm; ASAS modification of the Berlin algorithm; IBP is deleted as entry criterion and implemented as an additional SpA-feature (ASAS cohort). IBP, inflammatory back pain; SpA, spondyloarthritis; AS, ankylosing spondylitis; HLA-B27, Human Leukocyte Antigen; LR prob <80%, patients with a likelihood ratio probability below 80%; LR prob ≥80%, patients with a likelihood ratio probability of 80% or more; Diag Rheum, diagnosis of having axSpA according to the rheumatologist; ASAS, having axSpA according to the ASAS classification criteria.

Table 1: Baseline characteristics of patients in the SPACE cohort and the ASAS cohort; SpA vs no-SpA based on the diagnosis of the rheumatologist axSpA according to the rheumatologist; ASAS, having axSpA according to the ASAS classification criteria.

	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 $\pm$ SD	31.5 $\pm$ 16.6	31.1 $\pm$ 8.8	0.86	31.0 $\pm$ 10.8	35.8 $\pm$ 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 $\pm$ SD	13.4 $\pm$ 7.4 (months)	13.7 $\pm$ 7.1 (months)	0.79	6.3 $\pm$ 7.8 (years)	9.3 $\pm$ 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 $\pm$ SD	8.3 $\pm$ 11.6	5.7 $\pm$ 6.9	0.11	7.1 $\pm$ 14.9	2.4 $\pm$ 4.4	<0.001
ESR (mm/h), mean $\pm$ SD	13.6 $\pm$ 16.3	10.4 $\pm$ 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

p Values <0.05 are defined as statistically significant.

*Diagnosis according to rheumatologist.

†Not estimated in ASAS cohort.

ASAS, Assessment of SpondyloArthritis international Society; AxSpA, axial Spondyloarthritis; CRP, C reactive protein; ESR, erythrocyte sedimentation rate; HLA-B27, Human Leukocyte Antigen; IBD, Inflammatory Bowel Disease; IBP, Inflammatory Back Pain; preceding infection can be balanitis, urethritis, cervicitis and/or acute diarrhoea; SPACE, SpondyloArthritis Caught Early.

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-SIJ who was diagnosed as having SpA (figure 1C).

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-SIJ), 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-SIJ) ($n=7$), the algorithm suggests another diagnosis than axSpA.

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-SIJ but without SpA-features. In 150 HLA-B27 negative patients and in 10 HLA-B27 positive patients with a negative MRI-SIJ, 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-SIJ 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).

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 $\geq 80\%$ 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 $\geq 80\%$				
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 $\geq 80\%$				
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

ASAS, Assessment of SpondyloArthritis international Society; axSpA, axial Spondyloarthritis; LR-product, Likelihood Ratio-product, SPACE, SPondyloArthritis Caught Early.

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).

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-SIJ or MRI-SIJ) 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-SIJ), 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 axSpA was missed by the algorithm.

As shown in table 4, the majority of the patients (n=53) erroneously diagnosed as axSpA, have a negative MRI-SIJ. 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. 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-SIJ. Again, this trend is also seen using the two modifications.

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

Inclinometer 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=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	-

Imaging+: sacroiliitis present on MRI and/or x-rays

Imaging-: no sacroiliitis present on MRI and/or x-rays

A list of SpA-features is given in online supplementary table S1. *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. axSpA, axial Spondyloarthritis; HLA-B27, human leukocyte antigen; SPACE, SPondyloArthritis Caught Early.

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

Imaging+: sacroiliitis present on MRI and/or x-rays.

Imaging-: no sacroiliitis present on MRI and/or x-rays.

A list of SpA-features is given in online supplementary table S1. *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. ASAS, Assessment of SpondyloArthritis international Society; axSpA, axial Spondyloarthritis; HLA-B27, human leukocyte antigen.

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 ($\geq 3/5$ instead of $\geq 4/5$ 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-SIJ 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-SIJ in HLA-B27 negative (especially male) patients with 2–3 other SpA-features¹⁷. There are suggestions that an MRI-SIJ should be classified as positive on the basis of inflammatory lesions and structural changes to increase sensitivity of the MRI-SIJ. 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 patients¹². 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 a (atypical) presentation of ≥ 3 SpA-features,

but who are HLA-B27 negative and do not have sacroiliitis (X-SIJ and/or MRI-SIJ). 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-SIJ 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 $\geq 80\%$, 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.

Online supplementary data

Additional online supplementary data is available at the journal online:

<http://dx.doi.org/10.1136/annrheumdis-2012-201884>

References

1. Khan MA. Update on spondyloarthropathies. *Ann Intern Med* 2002;136:896-907.
2. Rudwaleit M, van der Heijde D, Khan MA, 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, Landewe R, van der Heijde D, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part I): classification of paper patients by expert opinion including uncertainty appraisal. *Ann Rheum Dis* 2009;68:770-6.
5. Sieper J, van der Heijde D, Landewe 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 AG, Hermann KG, et al. Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. *Ann Rheum Dis* 2009;68:1520-7.
7. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. *Arthritis Rheum* 1984;27:361-8.
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*;69:1264-8.
10. Calin A, Porta J, Fries JF, et al. Clinical history as a screening test for ankylosing spondylitis. *JAMA* 1977;237:2613-4.
11. Rudwaleit M, Khan MA, Sieper J. The challenge of diagnosis and classification in early ankylosing spondylitis: do we need new criteria? *Arthritis Rheum* 2005;52:1000-8.
12. Rudwaleit M, van der Heijde D, Landewe R, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann Rheum Dis* 2009;68:777-83.
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, Saracbası 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 HC, Spiller I, Song IH, 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 AG, 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

Magnetic resonance imaging of the sacroiliac joints in the early detection of spondyloarthritis: no added value of gadolinium compared to short tau inversion recovery sequence

Manouk de Hooge, Rosaline van den Berg, Victoria Navarro-Compán, Floris van
Galen, Désirée van der Heijde, Tom Huizinga, Monique Reijnerse

Rheumatology (Oxford) 2013;52(7):1220-4

Abstract

Objective

To investigate the additional value of T1 fatsat after Gd (T1/Gd) compared to T1 and STIR sequence in detecting active lesions of the sacroiliac joints (SIJ) typical for axSpA, in a prospective cohort study, the SPACE cohort. And to assess its influence on final MRI diagnosis of the SIJ (MRI-SIJ) based on the ASAS definition of active sacroiliitis.

Methods

Patients in the SpondyloArthritis Caught Early (SPACE) cohort received baseline and 3-month follow-up MRI-SIJ with coronal oblique T1, STIR and T1/Gd sequences. Bone marrow edema (BME), capsulitis/enthesitis and synovitis, and active sacroiliitis according to the ASAS definition were evaluated by 3 blinded readers.

Results

A total of 127 patients received an MRI-SIJ at baseline and 67 patients also received an MRI-SIJ at three months follow-up since the Gd protocol was added some months after the start of the SPACE-project. Twenty-five out of the 127 patients (19.7%) with a baseline MRI-SIJ and 14/67 patients (20.6%) with a follow-up MRI-SIJ presented BME on the STIR sequence sufficient enough to fulfill the ASAS definition for a positive MRI-SIJ. In eight patients additional synovitis and/or capsulitis/enthesitis was observed, however no additional BME was visualized on T1/Gd. One patient, without clinical diagnosis of axSpA, showed synovitis as an isolated finding.

Conclusions

Synovitis and capsulitis/enthesitis are detectable with the administration of gadolinium. However, they are always observed in the presence of BME. Therefore T1 and STIR sequence alone are sufficient in the MRI assessment which, among others, is used for diagnosing patients with early axSpA.

Introduction

Spondyloarthritis (SpA) is a spectrum of chronic inflammatory rheumatic diseases that share physiopathologic and clinical characteristics usually affecting the spine and the sacroiliac joints (SIJ). Sacroiliitis is defined by the presence of inflammation in one or both SIJ, and can be detected by plain radiographs or MRI¹. Radiographs only show structural changes but on MRI both inflammatory and structural changes are visible. Especially in early stages of the disease, MRI of the SIJ (MRI-SIJ) is of additional value in the diagnosis of axial SpA (axSpA), since plain radiographs can be normal in the presence of inflammation on MRI-SIJ. Therefore, the presence of bone marrow edema (BME) on MRI-SIJ has recently been included in the Assessment of Spondyloarthritis international Society (ASAS) axSpA classification criteria². Early detection of axSpA is essential to be able to start therapy early, as effective medication is available. In addition to early diagnosis of axSpA, MRI-SIJ is also used in the evaluation of treatment¹⁻³.

Generally, recommended sequences to detect active lesions are short tau inversion recovery (STIR) and T1 fat-saturated after gadolinium (T1/Gd), used in conjunction with T1-weighted sequence. While the STIR sequence shows active inflammation as increased signal intensity due to the presence of increased amounts of free water, T1/Gd is considered more sensitive since it depicts areas of increased vascularisation due to increased diffusion of Gd into the interstitial space⁴. There have been studies in which the STIR sequence is recommended over the T1 after gadolinium sequence, but these were studies concerning MRI of the spine^{5,6}. In other studies it is suggested that T1/Gd may improve the diagnostic capacity and the detection of active lesions, compared to the STIR sequence⁷⁻⁹. However, these were retrospective studies and since patient discomfort and costs are a drawback in using T1/Gd sequence its value ought to be properly determined.

The purpose of this study is to determine the additional value of T1/Gd sequence compared to the T1 and STIR sequences in detecting active lesions in the SIJ typical for axSpA in a prospective cohort study, the SPACE cohort, and to assess its possible influence on final definition of a positive MRI-SIJ based on the ASAS definition of active sacroiliitis.

Methods

All 127 patients in this study are part of the SpondyloArthritis Caught Early (SPACE) cohort. This cohort started in January 2009 and is an ongoing cohort including patients aged ≥ 16 years with short term chronic back pain (≥ 3 months, ≤ 2 years and

an onset <45 years) of unknown origin. At baseline all patients underwent a diagnostic work-up consisting of tests for the presence of acute anterior uveitis, inflammatory back pain (IBP), heel enthesitis, peripheral arthritis, psoriasis and IBD, positive family history for SpA, good response to NSAIDs, elevated ESR or CRP, HLA-B27 positivity and sacroiliitis on MRI and radiographs. Patients with a probable or certain diagnosis of axSpA were included in the follow-up. The SPACE cohort was approved by the Leiden University Medical Center Medical Ethics Committee and all patients provided written informed consent.

MRI-SIJ were performed on 1.5T (Philips Medical systems, Best, Netherlands) at baseline and after three months. The acquired sequences were coronal oblique T1-weighted TSE (TR 550/TE 10), STIR (TR 2500/TE 60) and T1-weighted TSE with fat suppression after intravenous gadolinium (TR 550/TE 10) with a slice thickness of 4mm. Gadolinium was administered with a dosage of 0.1 mmol/kg body weight. Written informed consent was obtained from all the participating patients. The SPACE cohort as well as this present study on the additional value of gadolinium were approved by the medical ethical committee of the Leiden University Medical Center.

MRI-SIJ were scored by two trained readers (MdH and RvdB). When both baseline and follow-up MRI-SIJ were available, these were scored simultaneously. Parameters scored for detecting active lesions of the SIJs typical for axSpA were: 1) Presence of BME. A positive MRI-SIJ was defined as the presence of one subchondral lesion of high signal intensity (BME lesion) on at least two consecutive slices or more than one lesion of high signal intensity on a single slice on the STIR and/or T1/Gd images, according to the ASAS/OMERACT definition¹; 2) The presence or absence of capsulitis/enthesitis, at the posterior cranial side of the SIJ, defined as high signal intensity on the T1/Gd or the STIR; 3) Presence or absence of synovitis detected on T1/Gd, defined as high signal intensity in the SI joint. The STIR and T1/Gd sequences were all scored in conjunction with T1 TSE images, showing low signal intensity of the BME lesions. In case of disagreement between the two readers, a third reader (VNC) scored the MRI-SIJ in the same way as described for the primary readers without knowledge of the scores of the first readers. To keep the third reader blinded for the outcomes of the scores, also cases without discrepancy were added. If two out of three readers rated the MRI-SIJ as positive (according to the ASAS definition), the MRI-SIJ was considered as such. All readers were blinded for clinical and laboratory data, the chronological order of the images as well as the report of the radiology department and the scores of the other readers.

Statistical analysis

All the scored abnormalities were categorical with a dichotomous measurement level. Agreement between T1-Gd and STIR sequences was measured by percentage of agreement¹⁰.

Results

At time of scoring the SPACE cohort included 157 patients. As the administration of Gd was added to the protocol some months after the start of the SPACE cohort, a baseline MRI-SIJ with Gd administration was obtained in 127 patients, in addition to the STIR and T1 sequences. In 90 patients also a follow-up MRI-SIJ after three months was obtained, including 67 patients with an MRI-SIJ with a T1/Gd sequence. At baseline, the average age of the patients was 30.8 (SD 8.4) years, mean duration of back pain was 13.4 (SD 7.1) months and the average age at onset of the back pain was 29.7 (SD 8.4) years. Forty-three patients (33.9%) were HLA-B27 positive and 87 patients (68.5%) had IBP at baseline. Of the patients studied at 3-month follow-up 25 out of the 67 patients (37.3%) were HLA-B27 positive and 49 patients (73.1%) had IBP.

In 102 patients there were no inflammatory abnormalities found on either the STIR or the T1/gd sequence of the MRI-SIJ. A positive MRI-SIJ according to the ASAS definition was seen in 19.7% of the patients (25/127) at baseline and in 20.9% of the patients (14/67) at follow-up, based on the STIR sequence. In the same 25 patients (baseline) and 14 patients (follow-up), the MRI-SIJ was also positive when judged by the T1/Gd sequence. The location of the BME lesions seen on the STIR corresponded with the location of lesions on the T1/Gd sequence in all patients at both time points. So in none of the patients additional BME was found on T1/Gd, resulting in 100% agreement between STIR and T1/Gd sequences.

Eleven of the 25 patients with a positive MRI-SIJ at baseline also had a positive MRI-SIJ at follow-up. Of the remaining 14 patients, four patients had a negative MRI-SIJ at 3-months follow-up and in 10 patients the follow-up MRI-SIJ was not taken yet. At follow-up, 20.9% (14/67) patients had a positive MRI-SIJ. In only three of these 14 patients the MRI-SIJ was negative at baseline, again when judged by STIR or by T1/Gd (table 1). So, over a period of three months three patients developed BME in the SIJ sufficient enough to meet the ASAS definition of a positive MRI-SIJ. Capsulitis/enthesitis was found in two patients (2.6%) at baseline. In one of these patients (1.2%) capsulitis/enthesitis was also found at follow-up, the other patient did not receive the follow-up MRI-SIJ. Capsulitis/enthesitis was only seen on the T1/Gd.

Table 1: Inflammation seen on the MRI-SIJ at baseline and follow-up after 3 months in the SPACE cohort.

	Baseline		Follow-up	
	STIR (n)	T1/Gd (n)	STIR (n)	T1/Gd (n)
<u>All patients (n=127)</u>				
Positive MRI (according to ASAS), n	25	25	14	14
Capsulitis/enthesitis, n	0	2	0	1
Synovitis, n	#	9	#	4
<u>Only baseline MRI available (n=60)</u>				
Positive MRI (according to ASAS), n	10	10	nd	nd
Capsulitis/enthesitis, n	0	1	nd	nd
Synovitis, n	#	5	nd	nd
<u>Baseline MRI and follow-up MRI available (n=67)</u>				
Positive MRI (according to ASAS), n	15	15	14	14
Capsulitis/enthesitis, n	0	1*	0	1*
Synovitis, n	#	4*	#	4*

nd: MRI at follow-up was not yet conducted. #Synovitis was not scored on the STIR sequence.

*The same patients at baseline and follow-up

With the STIR sequence this abnormality was not detected resulting in a 98.4% overall agreement between STIR and T1/Gd sequences at baseline and 98.9% overall agreement at follow-up. However, both at baseline and follow-up, capsulitis/enthesitis was always detected with the additional presence of BME on the STIR sequence and synovitis on the T1/Gd sequence. Synovitis was only assessed on the T1/Gd sequence and was detected in nine patients (6.3%) at baseline; five of those patients did not have a follow-up MRI-SIJ yet and in the remaining four patients (4.6%) synovitis was also shown at follow-up. Out of the eight patients showing synovitis at baseline, seven also showed BME on the MRI-SIJ. All patients with synovitis at follow-up showed BME on the follow-up MRI-SIJ. Figure 1 shows an example of a patient with capsulitis/enthesitis and synovitis as well as BME. The location of the BME lesions (indicated in the figure with circles) seen on the right image (T1/Gd sequence) is corresponding with the lesions on the left image (STIR sequence). The T1/Gd shows additional capsulitis/enthesitis (arrow) and synovitis (arrow head).

Out of the 25 patients at baseline, 22 patients did have axSpA according to the ASAS axSpA classification criteria. The remaining three patients show no other (clinical) SpA-features. Based on the absence of additional SpA-features these three patients could not be classified as axSpA by the ASAS axSpA classification criteria. The seven patients who show synovitis on the T1/Gd sequence and a positive MRI-SIJ according the ASAS definition also fulfilled the ASAS axSpA classification criteria. The one patient showing synovitis as a single symptom, so without BME, had normal radiographs, was HLA-B27 negative and did not have any other clinical SpA feature.

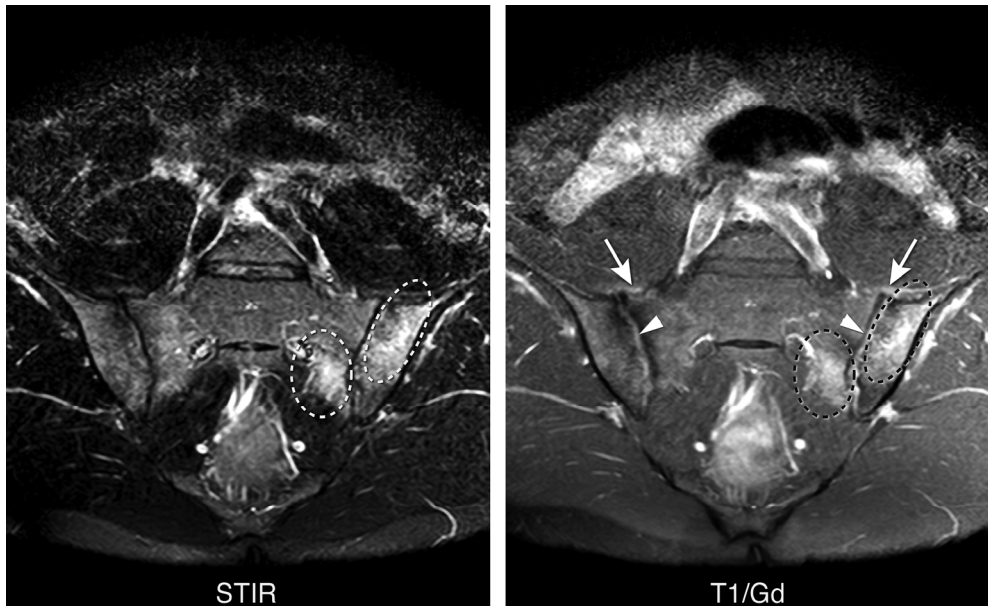


Figure 1: STIR sequence (left), T1/Gd sequence (right), BME (dotted circles), capsulitis/enthesitis (arrows) and synovitis (arrowheads).

Discussion

In this study the additional value of the intravenous administration of gadolinium (T1/Gd) in detecting BME, capsulitis/enthesitis and synovitis compared to the STIR sequence was investigated. In addition we also assessed the influence of T1/Gd on the final evaluation of the MRI based on the ASAS definition of a positive MRI-SIJ.

A 100% agreement was found between the STIR and T1/Gd sequence in detecting BME lesions. For the detection of BME lesions, the administration of gadolinium does not have additional value. This confirms the findings of a previous study by Madsen et al.¹¹, comparing STIR sequence with a gadolinium contrast-enhanced sequence to detect active abnormalities at the SIJ. However, that study included patients with chronic fatty lesions rather than patients with early BME lesions. The major focus of the current study is on the early stage of SpA, making this study not only a confirmation but also an important complementation to the findings of Madsen et al.

Our results also show that the presence of capsulitis/enthesitis is better detected with the administration of Gd. Although the T1/Gd sequence enabled the readers to score capsulitis/enthesitis while this was not seen on the STIR, capsulitis/enthesitis was always present in addition to BME lesions on the STIR sequence and synovitis on the

T1/Gd sequence. Synovitis, not detectable with the STIR sequence, was found in 8 patients at baseline and 4 at follow-up on the T1/Gd sequence. So in an early cohort as the SPACE cohort, with patients who have ≤ 2 years back pain complains, a small percentage of patients ($< 7\%$) show synovitis. All but 1 patient with synovitis also showed BME on the STIR sequence sufficient enough to meet the ASAS definition of a positive MRI and moreover had a diagnosis of axSpA. The patient with only synovitis did not have any sign related to axSpA. So synovitis as the only sign on MRI-SIJ can be considered as a false positive finding in this patient. This is in line with the recommendations by ASAS that the sole presence of synovitis is not enough for a positive MRI-SIJ¹². So, being unable to detect synovitis on the STIR sequence is not an issue. It confirms the recommendations by ASAS that only the presence of BME is essential in the early diagnosis of axSpA and that the STIR sequence is sufficient to detect BME. Considering this all, the administration of gadolinium for patients of the SPACE cohort was stopped in the beginning of April 2012, leaving the STIR and T1 sequences as only assessed MRI sequences in this cohort.

In summary, we found that the presence of synovitis and capsulitis/enthesitis is detectable with the administration of gadolinium. However, these parameters are present only in a low percentage of patients with suspected axSpA. Furthermore, synovitis and capsulitis/enthesitis is observed in the presence of BME. We conclude that T1 and STIR sequences alone are sufficient in the MRI assessment to diagnose patients with early axSpA.

References

1. Rudwaleit M, Jurik AG, Hermann KG, et al. Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. *Ann Rheum Dis* 2009;68:1520-7.
2. Rudwaleit M, van der Heijde D, Landewe R, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann Rheum Dis* 2009;68:777-83.
3. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. *Arthritis Rheum* 1984;27:361-8.
4. Maksymowych WP, Weber U. Diagnostic utility of MRI in early spondyloarthritis. *Curr Rheumatol Rep* 2011;13:402-8.
5. Baraliakos X, Hermann KG, Landewe R, et al. Assessment of acute spinal inflammation in patients with ankylosing spondylitis by magnetic resonance imaging: a comparison between contrast enhanced T1 and short tau inversion recovery (STIR) sequences. *Ann Rheum Dis* 2005;64:1141-4.
6. Hermann KG, Landewe RB, Braun J, et al. Magnetic resonance imaging of inflammatory lesions in the spine in ankylosing spondylitis clinical trials: is paramagnetic contrast medium necessary? *J Rheumatol* 2005;32:2056-60.
7. Adam G, Dammer M, Bohndorf K, et al. Rheumatoid arthritis of the knee: value of gadopentetate dimeglumine-enhanced MR imaging. *AJR Am J Roentgenol* 1991;156:125-9.
8. Algin O, Gokalp G, Baran B, et al. Evaluation of sacroiliitis: contrast-enhanced MRI with subtraction technique. *Skeletal Radiol* 2009;38:983-8.
9. Althoff CE, Feist E, Burova E, et al. Magnetic resonance imaging of active sacroiliitis: do we really need gadolinium? *Eur J Radiol* 2009;71:232-6.
10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
11. Madsen KB, Egund N, Jurik AG. 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.
12. 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:ii1-44.

5

Patients with chronic back pain of short duration from the SPACE cohort: which MRI structural lesions in the sacroiliac joints and inflammatory and structural lesions in the spine are most specific for axial spondyloarthritis?

Manouk de Hooge, Rosaline van den Berg, Victoria Navarro-Compán, Monique Reijnierse, Floris van Gaalen, Karen Fagerli, Robert Landewé, Maikel van Oosterhout, Roberta Ramonda, Tom Huizinga, Désirée van der Heijde

Ann Rheum Dis 2016;75(7):1308-14

Abstract

Objectives

To investigate the extent and performance of MRI lesions in the sacroiliac joint (MRI-SI) and spine (MRI-spine) in patients with suspected axial spondyloarthritis (axSpA).

Methods

MRI-SI/spine of patients with chronic back pain (onset <45 years) in the SPondyloArthritis Caught Early (SPACE) cohort were scored by two well-trained readers for inflammation, fatty lesions, erosions, sclerosis/ankylosis and syndesmophytes. MRI performances were tested against the Assessment of Spondyloarthritis international Society (ASAS) axSpA criteria (positive: imaging-arm+ or clinical-arm+; negative: possible axSpA (few spondyloarthritis (SpA) features present) or no SpA). Arbitrary cut-off levels for MRI lesions were set to assure at least 95% specificity (tested in the no SpA group).

Results

In total 126 patients were ASAS criteria positive (73 imaging-arm+ (22 by modified New York criteria (mNY)+; 51 by MRI+mNY-); 53 clinical-arm+) and 161 were ASAS criteria negative (89 possible axSpA and 72 no SpA). On MRI-SI (n=287), at least three fatty lesions (or at least three erosions) were seen in 45.5 (63.6)% of mNY+ patients, 15.7 (47.1)% of MRI+mNY- patients and 15.1 (13.2)% of clinical-arm+ patients versus 3.4 (6.7)% of possible axSpA patients and 2.8 (4.2)% of no SpA patients. A combined rule (at least five fatty lesions and/or erosions) performed equally well. Sclerosis and ankylosis were too rare to analyse. On MRI-spine (n=284), at least five inflammatory lesions (or at least five fatty lesions) were seen in 27.3 (18.2)% of mNY+ patients, 13.7 (21.6)% of MRI+mNY- patients and 3.8 (1.9)% of clinical-arm+ patients versus 4.5 (6.7)% of possible SpA patients and 2.9 (4.3)% of no SpA patients.

Conclusions

The presence of (1) at least five fatty lesions and/or erosions on MRI-SI, (2) at least five inflammatory lesions or (3) at least five fatty lesions on MRI-spine allows an acceptable discrimination of axSpA and no SpA, while assuring >95% specificity.

Introduction

Conventional radiographs of sacroiliac joints (SIJ) are frequently used for detecting sacroiliitis and diagnosing axial spondyloarthritis (axSpA). Limitations to this method are the major inter-reader variability and, as it takes time to develop structural damage visible on radiographs, the delay in diagnosis of several years^{1–5}. A good addition to using radiographs is MRI, since MRI can visualise both inflammatory lesions and structural damage while radiographs only depict structural damage⁴. Hence the increasing interest in MRI in the field of axSpA. MRI has a major role as entry criteria imbedded in the imaging arm of the Assessment of Spondyloarthritis international Society (ASAS) classification criteria for axSpA⁶. Patients fulfilling the imaging arm have sacroiliitis on conventional radiographs and/or MRI plus ≥ 1 additional spondyloarthritis (SpA) feature⁶. Radiographic sacroiliitis is defined as described in the modified New York (mNY) criteria.⁷ Sacroiliitis on MRI is defined according to the ASAS definition of a positive MRI of the SIJ (MRI-SI)⁸. Experts in the field created this definition in 2009, which is based on inflammatory lesions only. The question if structural lesions should be added to this definition is still outstanding. Trying to address this question, a study found that erosions are relatively specific for SpA and may increase the utility of MRI. They proposed an adjusted version of the ASAS definition of a positive MRI-SI adding the presence of erosions⁹. Another question that emerged was whether the inclusion of spinal lesions could contribute to detecting axSpA. In 2012, the same group of experts that created the definition of a positive MRI-SI, proposed a definition for a positive MRI of the spine (MRI-spine)¹⁰. The realisation of this definition was based on a systematic literature review; however, before its existence the Canada–Denmark MRI working group also described spinal lesions typical for SpA. Comparing these studies demonstrates the similarity of descriptions used by both working groups^{10–12}.

There were two objectives in this study. First, to investigate the quantity of structural MRI-SI lesions and inflammatory and structural MRI-spine lesions in patients with axSpA and patients with relatively short duration chronic back pain (CBP) due to other causes. Elaborating on this, the second objective was to investigate which lesions, and to what extent, assure sufficient specificity for axSpA.

Methods

Patients of ≥ 16 years with CBP (≥ 3 months, ≤ 2 years, onset < 45 years, cause unknown) were included in the SPondyloArthritis Caught Early (SPACE) cohort from five participating centres in the Netherlands (Leiden, Amsterdam and Gouda),

Norway (Oslo) and Italy (Padova). To classify patients according to the ASAS criteria, presence of SpA features was examined according to a fixed protocol including physical examination, laboratory tests, plain radiographs and MRI-SI/spine. The SpA features are radiographic sacroiliitis (mNY criteria: bilateral grade ≥ 2 or unilateral grade ≥ 3)⁷, sacroiliitis on MRI (presence of ≥ 2 inflammatory lesions highly suggestive of axSpA on a single slice or one lesion on ≥ 2 consecutive slices)⁸, HLA-B27 positivity, positive family history for SpA, inflammatory back pain (IBP), psoriasis, peripheral arthritis, dactylitis, heel enthesitis, uveitis, inflammatory bowel disease (IBD), good response to non-steroidal anti-inflammatory drugs (NSAIDs) and elevated C-reactive protein (CRP) or erythrocyte sedimentation rate (ESR)⁶.

Baseline data were used in the analyses. A detailed description of the recruitment and clinical assessment of the SPACE cohort has been published before¹³.

MRI assessment

Patients underwent an MRI-SI and MRI-spine performed on a 1.5T scanner. The acquired sequences for both modalities were Short Tau Inversion Recovery (TR2500/TE600) and T1-weighted Turbo Spin-Echo (TR550/TE10), with a slice thickness of 4 mm. The MRI-SI images were performed in coronal oblique plane and MRI-spine in sagittal plane¹⁴. Two well-trained readers (RvdB and MdH) independently scored all MRIs with both sequences viewed simultaneously, while blinded for all clinical data. MRI-SI was scored separately from MRI-spine. In some cases adjudication scores were obtained (see below).

MRI-SI lesions

Weber et al adapted the Spondyloarthritis Research Consortium of Canada (SPARCC) scoring system for inflammation in the SIJ to quantify fatty lesions, erosions, sclerosis and ankylosis in the SIJ⁹. In the current study this adapted SPARCC scoring system was used and structural lesions on ≥ 2 consecutive slices were scored, resulting in a maximum score of 40 per lesion (5 lesions per quadrant over all slices $\times$ 4 quadrants $\times$ 2 SIJ). Note that ankylosis always involves two quadrants (either upper iliac and upper sacral or lower iliac and lower sacral quadrants). In addition, readers gave a general dichotomous verdict (yes/no) on a 'positive MRI-SI' (ASAS definition)⁸.

An adjudicator (VNC) was introduced when readers disagreed on a 'positive MRI-SI' or the presence/absence of erosions according to a specific definition (≥ 2 erosions on the same slice or one erosion seen on ≥ 2 consecutive slices)⁹. We offered images that needed adjudication and images on which initial readers agreed to keep the adjudicator

blinded. In the analyses, mean scores of the two readers agreeing on a 'positive MRI-SI' were used for quantifying MRI-SI lesions.

MRI-spine lesions

Spinal MRI lesions were scored according to the Canada–Denmark score in which every vertebral unit (VU) is divided into quadrants. Per quadrant dichotomous scores (presence/absence) on corner inflammatory and structural lesions (fatty lesions, erosions, syndesmophytes) are given^{11,12}. Inflammatory and fatty lesions were scored when present on ≥ 2 consecutive slices. Erosions and syndesmophytes were scored when visible on ≥ 1 slice. Spinal lesions were only scored when considered typical for axSpA and not due to other causes like degenerative changes. Both readers also gave a general dichotomous verdict (yes/no) on a 'positive MRI-spine' (ASAS definition: ≥ 3 corner inflammatory lesions)¹⁰. An adjudicator (VNC) was introduced when inter-reader disagreement on a 'positive MRI-spine' (ASAS definition) or presence/absence of ≥ 3 fatty lesions was found.

The lesions of all quadrants throughout the 23 VUs were added up and the mean scores of the two readers agreeing on 'positive MRI-spine' were used to quantify spinal lesions for analyses.

Classification

All patients were classified according to the ASAS criteria for axSpA⁶. Patients with axSpA were divided into subgroups; the imaging arm contained patients fulfilling the mNY criteria (mNY+) and patients not fulfilling the mNY criteria but with a 'positive MRI-SI' (MRI+mNY-) and ≥ 1 SpA feature. HLA-B27-positive patients with ≥ 2 additional SpA features and mNY- and MRI-SI- were grouped (clinical arm). There were also patients not fulfilling the ASAS criteria but with a suspicion of axSpA because of the presence of SpA features with a high likelihood (eg, dactylitis) or three less specific (IBP, positive family history for SpA and good response to NSAIDs) SpA features (possible axSpA)¹⁵. The final group consists of patients in whom the possibility of having SpA at baseline was very low since patients did not have sacroiliitis on imaging, were HLA-B27 negative or had ≤ 2 SpA features with a low likelihood for SpA (no SpA). For a complete and detailed description on how patients were classified see figure 1.

To detect the influence of age we compared the extent of MRI lesions in patients with an age at onset of CBP <35 years and between 35 and 45 years.

	mNY+	MRI+mNY-	Clinical	Possible SpA			No-SpA		
Radiographic sacroiliitis	Pos	Neg	Neg	Neg	Neg	Pos	Neg	Neg	Neg
Sacroiliitis on MRI	Pos. or neg.	Pos	Neg	Neg	Neg	Pos	Neg	Neg	Neg
HLA-B27 positive			Pos	Pos	Neg	Neg	Pos	Neg	Neg
Peripheral arthritis					Neg	Neg	Neg	Neg	Neg
Dactylitis						Neg	Neg	Neg	Neg
Heel enthesitis						Neg	Neg	Neg	Neg
Psoriasis						Neg	Neg	Neg	Neg
IBD						Neg	Neg	Neg	Neg
Uveitis						Neg	Neg	Neg	Neg
Elevated ESR/CRP						Neg	Neg	Neg	Neg
Positive Family History						Neg	Neg	Neg	Neg
IBP						Neg	Neg	Neg	Neg
Good response to NSAIDs						Neg	Neg	Neg	Neg

Figure 1: Grouping of the patients. The first row depicts the different groups and the first column depicts all SpA features (starting with the major elements of the ASAS criteria (imaging and HLA-B27)). The other columns represent the various classification groups. Each patient can belong only to one of these subgroups. Per column, all conditions mentioned in the cells must be met. Some individual SpA features are mandatory: the red cells indicate all individual features that must be present and the grey cells indicate all individual features that must be absent. In addition, in some subgroups there are individual SpA features that have no impact on the classification (indicated as yellow). Finally, there are groups of SpA features for which exactly one SpA feature needs to be positive (green), ≥ 1 SpA feature needs to be positive (purple) or ≥ 2 SpA features need to be positive (blue). ASAS, Assessment of Spondyloarthritis international Society; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; feat., feature; HLA, human leucocyte antigen; IBD, inflammatory bowel disease; mNY, modified New York criteria; Neg, negative (grey); NSAIDs, non-steroidal anti-inflammatory drugs; Pos, positive (red); SpA, spondyloarthritis.

Data analysis

The proportion of patients with lesions was presented using several cut-offs. Lesions in the no SpA group were considered ‘false positive’. Arbitrarily, we defined the lowest cut-offs with $\leq 5\%$ false positives, reported in ***bold and italic***, as the optimal cut-off to ensure high specificity. In addition, the lowest cut-offs with $\leq 10\%$ false positives were reported in *italic*. Besides single MRI lesions, also combinations of lesions were tested in order to find the number and type of lesions most specific for axSpA. All analyses were performed in SPSS V.22.0.

Results

Baseline MRI-SI and clinical data were available in 287 patients, of which 103 (35.9%) are men. There were 126 patients (43.9%) fulfilling the ASAS criteria, 89 possible axSpA patients (31.0%) and 72 no SpA patients (25.1%). Of the no SpA patients, 21 had IBP only, 12 had both IBP and a good response to NSAIDs, 7 had a good response to NSAIDs, 5 had a positive family history and 2 patients were HLA-B27 positive, all in the absence of other SpA features. There were 18 patients without SpA features. Five possible axSpA patients had sacroiliitis on X-ray ($n=1$) or MRI ($n=4$) in the absence of other SpA features and therefore did not fulfil the ASAS criteria (table 1).

MRI-SI lesions

In 72 cases adjudication was needed due to disagreement on a positive MRI-SI (kappa primary readers=0.75) or the presence of erosions (kappa primary readers=0.57).

Table 2 shows the proportion of patients with single MRI-SI lesions. Fatty lesions and erosions were seen in all subgroups of patients, but more frequently in patients fulfilling the imaging arm (mNY+ and MRI+mNY-), and only somewhat more frequent in patients in the clinical arm compared with the possible axSpA and no SpA groups. Cut-off ‘ ≥ 3 fatty lesions’ and cut-off ‘ ≥ 3 erosions’ corresponded to $\leq 5\%$ false positives. Applying these cut-offs differentiated well between patients fulfilling the imaging arm, especially mNY+ patients, and no SpA patients. The proportion of patients with ‘ ≥ 3 fatty lesions’ and ‘ ≥ 3 erosions’ in the clinical arm and possible axSpA group is relatively low, ranging from 3.4% to 15.1%. Sclerosis and ankylosis were rarely seen (table 2).

Several combinations between structural lesions were investigated to see if this would increase specificity. Since sclerosis and ankylosis as single lesions occurred rarely, and were not specific, combinations containing these lesions were not shown, as these did

not change the results. When looking at ‘fatty lesions and/or erosions’ cut-off ≥ 5 yielded $\leq 5\%$ false positives, while the proportion of patients in the MRI+mNY- (37.3%) and mNY+ (63.6%) groups remain reasonably high (see table 2).

Table 1: Characteristics of patients with axial spondyloarthritis (axSpA) and chronic back pain due to unknown causes

ASAS classification 2/3 readers	No SpA	Possible axSpA	Clinical arm only	MRI+mNY-	mNY+
No. of patients	72	89	53	51	22
Age (years) at onset back pain, mean $\pm$ SD	30.6 (± 0.98)	30.4 (± 0.93)	27.5 (± 1.07)	30.5 (± 1.08)	25.7 (± 1.53)
Duration of back pain (months), mean $\pm$ SD	14.6 (± 0.86)	12.4 (± 0.72)	11.7 (± 0.92)	13.1 (± 1.09)	13.1 (± 1.85)
Male, n (%)	20 (27.8)	24 (27.0)	19 (35.8)	27 (52.9)	13 (59.1)
HLA-B27 positive, n (%)	2 (2.8)	9 (10.2)	53 (100)	31 (60.8)	15 (65.2)
Positive Family History, n (%)	14 (19.4)	37 (42.0)	33 (62.3)	18 (35.3)	9 (39.1)
IBP, n (%)	40 (55.6)	56 (63.6)	47 (88.7)	34 (66.7)	17 (73.9)
Psoriasis, n (%)	0	15 (17.0)	8 (15.1)	4 (7.8)	3 (13.4)
Peripheral Arthritis, n (%)	0	20 (22.7)	12 (22.6)	10 (19.6)	2 (8.7)
Dactylitis, n (%)	0	9 (10.2)	5 (9.4)	4 (7.8)	1 (4.3)
Heel enthesitis, n (%)	0	27 (30.7)	13 (24.5)	8 (15.7)	4 (17.4)
Uveitis, n (%)	0	6 (6.8)	10 (18.9)	5 (9.8)	3 (13.4)
IBD, n (%)	0	14 (15.9)	2 (3.8)	6 (11.8)	3 (13.4)
Good response to NSAIDs, n (%)	21 (29.2)	31 (35.2)	26 (49.1)	18 (35.3)	14 (60.9)
Elevated ESR/CRP, n (%)	0	26 (29.5)	13 (24.5)	11 (21.6)	7 (30.4)
CRP, mean $\pm$ SD	3.9 (± 0.40)	6.7 (± 1.21)	7.6 (± 1.68)	6.7 (± 1.01)	7.1 (± 1.33)
Radiographic sacroiliitis, n (%)	0	1 (1.1)	0	0	22 (100)
Sacroiliitis on MRI, n (%)	0	4 (4.5)	0	51 (100)	12 (52.2)

ASAS, Assessment of Spondyloarthritis international Society; axSpA, axial spondyloarthritis; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; HLA, human leucocyte antigen; IBD, inflammatory bowel disease; IBP, inflammatory back pain; mNY, modified New York criteria; NSAIDs, non-steroidal anti-inflammatory drugs; SpA, spondyloarthritis.

MRI-spine lesions

In 284 of 287 patients MRI-spine was available. Disagreement between readers resulted in adjudication in 43 patients. The reliability for the primary readers was good to moderate (kappa 0.73 for inflammatory lesions and 0.60 for fatty lesions).

More than half of the patients in the imaging arm had spinal inflammatory lesions (cut-off ≥ 1), but this was also the case in 36.3% of the no SpA patients. With ≥ 5 inflammatory lesions the proportion of false positives decreased $\leq 5\%$ with still a relatively high proportion of patients with inflammatory lesions in MRI+mNY- (13.7%) and especially in the mNY+ group (27.3%) (table 3). Fatty lesions (cut-off ≥ 1) were the most often seen spinal structural lesions (range 24.5%–50.0%). Cut-off ≥ 5 corresponds to $\leq 5\%$ false positives, while the proportion of patients in the mNY+ and MRI+mNY- groups were 18.2% and 21.6%, respectively.

Table 2. Proportions of patients with single or combined lesions seen on MRI of the SIJ

ASAS classification 2/3 readers	No SpA	Possible axSpA	Clinical arm only	MRI+mNY-	mNY+
No. of patients	72	89	53	51	22
Single lesions					
Fatty lesions ≥ 1	10	11	14	18	11
%	13.8	12.4	26.4	35.3	50.0
Fatty lesions ≥ 2	5	7	8	12	11
%	6.9	7.9	15.1	23.5	50.0
Fatty lesions ≥ 3	2	3	8	8	10
%	2.8	3.4	15.1	15.7	45.5
Erosions ≥ 1	15	29	21	40	17
%	20.8	32.6	39.6	78.4	77.3
Erosions ≥ 2	6	15	8	32	15
%	8.3	16.9	15.1	62.7	68.2
Erosions ≥ 3	3	6	7	24	14
%	4.2	6.7	13.2	47.1	63.6
Sclerosis ≥ 1	4	8	3	10	4
%	5.6	9.0	5.7	19.6	18.2
Sclerosis ≥ 2	1	3	1	4	1
%	1.4	3.4	1.9	7.8	4.5
Ankylosis ≥ 1	2	3	2	2	2
%	2.8	3.4	3.8	3.9	9.1
Combination of lesions					
Fatty lesions and/or erosion ≥ 1	20	32	26	41	18
%	27.8	36.0	49.1	80.4	81.8
Fatty lesions and/or erosion ≥ 2	11	17	14	32	17
%	15.3	19.1	26.4	62.7	77.3
Fatty lesions and/or erosion ≥ 3	6	11	13	26	17
%	8.3	12.4	24.5	51.0	77.3
Fatty lesions and/or erosion ≥ 4	4	10	7	22	14
%	5.6	11.2	13.2	43.1	63.6
Fatty lesions and/or erosion ≥ 5	1	6	6	19	14
%	1.4	6.7	11.3	37.3	63.6

False positives <10% in italic; false positives <5% in italic and bold

Between 22.7% and 33.3% of patients across all groups had ≥ 1 erosion. Increasing the cut-off decreased the proportion of false positives from 14.5% (cut-off ≥ 2) to 7.2% (cut-off ≥ 3) to 2.9% (cut-off ≥ 4). However, the proportion of patients in the imaging arm was nearly as low, decreasing from 18.2% to 13.6% to 4.5% (mNY+) and 15.7% to 7.8% to 3.9% (MRI+mNY-) with cut-off ≥ 2 , ≥ 3 and ≥ 4 , respectively, showing that spinal erosions are not suitable to discriminate between subgroups. This applies also to syndesmophytes, as similar proportions of patients with syndesmophytes were shown in all subgroups irrespective of the cut-off (table 3). Specificity did not increase when combining inflammatory and structural lesions. Even when combining the two most specific spinal lesions (inflammatory and/or

fatty lesions'), specificity did not increase and the proportion of false positives only drop below $\leq 5\%$ with cut-off ≥ 8 (data not shown).

Table 3. Proportions of patients with lesions seen on MRI of the spine

ASAS classification 2/3 readers	No SpA	Possible axSpA	Clinical arm only	MRI+ mNY-	mNY+
No. of patients	69	89	53	51	22
Inflammatory lesions ≥ 1	25	29	19	27	12
%	36.3	32.6	35.8	52.9	54.5
Inflammatory lesions ≥ 2	15	19	11	18	11
%	21.7	21.3	20.8	35.3	50.0
Inflammatory lesions ≥ 3	11	10	6	12	9
%	15.9	11.2	11.3	23.5	40.9
<i>Inflammatory lesions ≥ 4</i>	7	7	4	9	9
%	10.1	7.9	7.5	17.6	40.9
<i>Inflammatory lesions ≥ 5</i>	2	4	2	7	6
%	2.9	4.5	3.8	13.7	27.3
Fatty lesions ≥ 1	18	24	13	24	11
%	26.1	27.0	24.5	47.1	50.0
Fatty lesions ≥ 2	11	16	8	17	9
%	15.9	18.0	15.1	33.3	40.9
Fatty lesions ≥ 3	9	9	3	13	6
%	13.0	10.1	5.7	25.5	27.3
<i>Fatty lesions ≥ 4</i>	6	7	2	11	4
%	8.7	7.9	3.8	21.6	18.2
<i>Fatty lesions ≥ 5</i>	3	6	1	11	4
%	4.3	6.7	1.9	21.6	18.2
Erosions ≥ 1	23	24	16	15	5
%	33.3	27.0	30.2	29.4	22.7
Erosions ≥ 2	10	10	5	8	4
%	14.5	11.2	9.4	15.7	18.2
<i>Erosions ≥ 3</i>	5	4	4	4	3
%	7.2	4.5	7.5	7.8	13.6
<i>Erosions ≥ 4</i>	2	4	1	2	1
%	2.9	4.5	1.9	3.9	4.5
Syndesmophytes ≥ 1	17	21	8	19	2
%	24.6	23.6	15.1	37.3	9.1
<i>Syndesmophytes ≥ 2</i>	6	11	5	6	0
%	8.7	12.4	9.4	11.8	-
<i>Syndesmophytes ≥ 3</i>	5	7	3	3	0
%	7.2	7.9	5.7	5.9	-
<i>Syndesmophytes ≥ 4</i>	2	3	2	1	0
%	2.9	3.4	3.8	2.0	-

False positives <10% in italic; false positives <5% in italic and bold

Effect of age

In order to further investigate the specificity of lesions we stratified patients by age at onset of CBP: onset of CBP <35 years (younger age group) or onset of CPB between 35 and 45 years (older age group). On MRI-SI there was no effect of age for any single lesion, nor for any combinations (data not shown). However, on MRI-spine, the number of inflammatory lesions corresponding to $\leq 5\%$ false positives was far lower in the younger (≥ 3) compared with the older (≥ 6) age group (table 4). Furthermore, with this lower cut-off in the younger age group the proportion of patients in the imaging subgroups was still high, while with the higher cut-off in the older age group the proportion of patients was comparable throughout all subgroups. Similarly, fatty lesions (cut-off ≥ 3) were only specific for axSpA in the younger age group. For erosions and syndesmophytes the specificity was poor in both age groups (table 4).

Table 4. Proportions of patients with lesions seen on MRI of the spine, stratified by age

Patients with age at onset chronic back pain <35 years					
ASAS classification 2/3 readers	No SpA	Possible axSpA	Clinical arm only	MRI+ mNY-	mNY+
No. of patients	41	55	41	36	20
Inflammatory lesions ≥ 1	12	11	13	18	11
%	29.3	20.0	31.7	50.0	55.0
<i>Inflammatory lesions ≥ 2</i>	<i>3</i>	<i>7</i>	<i>7</i>	<i>12</i>	<i>10</i>
%	7.3	12.7	2.1	33.3	50.0
<i>Inflammatory lesions ≥ 3</i>	<i>2</i>	<i>1</i>	<i>2</i>	<i>10</i>	<i>8</i>
%	<i>4.8</i>	<i>1.8</i>	<i>4.8</i>	<i>27.8</i>	<i>40.0</i>
Fatty lesions ≥ 1	8	8	6	15	9
%	19.5	14.5	14.6	41.7	45.0
<i>Fatty lesions ≥ 2</i>	<i>4</i>	<i>5</i>	<i>5</i>	<i>9</i>	<i>7</i>
%	9.8	9.1	12.2	25.0	35.0
<i>Fatty lesions ≥ 3</i>	<i>3</i>	<i>3</i>	<i>3</i>	<i>6</i>	<i>5</i>
%	7.3	5.5	7.3	16.7	25.0
<i>Fatty lesions ≥ 3</i>	<i>2</i>	<i>2</i>	<i>2</i>	<i>5</i>	<i>4</i>
%	<i>4.8</i>	<i>3.6</i>	<i>4.8</i>	<i>13.9</i>	<i>20.0</i>
Erosions ≥ 1	9	12	14	9	4
%	22.0	21.8	34.1	25.0	20.0
<i>Erosions ≥ 2</i>	<i>3</i>	<i>3</i>	<i>4</i>	<i>4</i>	<i>3</i>
%	7.3	5.5	9.8	11.1	15.0
<i>Erosions ≥ 3</i>	<i>3</i>	<i>1</i>	<i>4</i>	<i>1</i>	<i>2</i>
%	7.3	1.8	9.8	2.8	10.0
<i>Erosions ≥ 4</i>	<i>1</i>	<i>1</i>	<i>1</i>	<i>1</i>	<i>1</i>
%	<i>2.4</i>	<i>1.8</i>	<i>2.4</i>	<i>2.8</i>	<i>5.0</i>
Syndesmophytes ≥ 1	8	12	6	12	1
%	19.5	21.8	14.6	33.3	5.0

<i>Syndesmophytes ≥2</i>	3	5	4	4	0
%	7.3	9.1	9.8	11.1	-
<i>Syndesmophytes ≥3</i>	2	3	3	1	0
%	4.8	5.5	7.3	2.8	-
Patients with age at onset chronic back pain between 35 and 45 years					
ASAS classification 2/3 readers	No SpA	Possible axSpA	Clinical arm only	MRI+ mNY-	mNY+
No. of patients	28	34	12	15	2
Inflammatory lesions ≥1	13	18	6	9	1
%	46.4	52.9	50.0	60.0	50.0
Inflammatory lesions ≥2	12	12	4	6	1
%	42.9	35.3	33.3	40.0	50.0
Inflammatory lesions ≥3	9	9	4	2	1
%	32.1	26.5	33.3	13.3	50.0
Inflammatory lesions ≥4	5	7	3	2	1
%	17.9	20.6	25.0	13.3	50.0
<i>Inflammatory lesions ≥5</i>	2	4	1	2	1
%	7.1	11.8	8.3	13.3	50.0
<i>Inflammatory lesions ≥6</i>	1	2	1	2	0
%	3.6	5.9	8.3	13.3	-
Fatty lesions ≥1	10	16	7	9	2
%	35.7	47.1	58.3	60.0	100.0
Fatty lesions ≥2	7	11	3	8	2
%	25.0	32.4	25.0	53.3	100.0
Fatty lesions ≥3	6	6	0	7	1
%	21.4	17.6	-	46.7	50.0
Fatty lesions ≥4	4	5	0	6	0
%	14.3	14.7	-	40.0	-
Fatty lesions ≥5	3	4	0	6	0
%	10.7	11.8	-	40.0	-
<i>Fatty lesions ≥6</i>	2	4	0	5	0
%	7.1	11.8	-	33.3	-
<i>Fatty lesions ≥7</i>	1	4	0	2	0
%	3.6	11.8	-	13.3	-
Erosions ≥1	14	12	2	6	1
%	50.0	35.3	16.7	40.0	50.0
Erosions ≥2	7	7	1	4	1
%	25.0	20.6	8.3	26.7	50.0
<i>Erosions ≥3</i>	2	3	0	3	1
%	7.1	8.8	-	20.0	50.0
<i>Erosions ≥4</i>	1	3	0	1	0
%	3.6	8.8	-	6.7	-
Syndesmophytes ≥1	9	9	2	7	1
%	32.1	26.5	16.7	46.7	50.0
Syndesmophytes ≥2	3	6	1	2	0
%	10.7	17.6	8.3	13.3	-

Syndesmophytes ≥ 3	3	4	0	2	0
%	10.7	11.8	-	13.3	-
Syndesmophytes ≥ 4	0	2	0	1	0
%	-	5.9	-	6.7	-

False positives <10% in italic; false positives <5% in italic and bold

Discussion

In patients with CBP included in the SPACE cohort we investigated the extent of structural lesions on MRI-SI, as well as inflammatory and structural lesions on MRI-spine. We found the rules ‘ ≥ 5 fatty lesions and/or erosions’ on MRI-SI, as well as ‘ ≥ 5 spinal inflammatory lesions’ and ‘ ≥ 5 spinal fatty lesions’ to be highly specific for axSpA, while still assuring an acceptable and useful level of discrimination between axSpA patients and no SpA patients. Similar results were found for single MRI-SI lesions: optimal cut-offs were ‘ ≥ 3 fatty lesions’ and ‘ ≥ 3 erosions’.

Erosions on MRI-SI are considered specific for axSpA. Studies have shown that erosions, either in the presence or in the absence of inflammation, are seen in both patients with early axSpA and ankylosing spondylitis (AS). These studies have purported that structural lesions, especially erosions, greatly contribute to the diagnostic value of MRI-SI^{16,17}. But there are limited data about the extent of MRI-SI lesions. Althoff et al and Krohn et al reported 2.4–7.6 fatty lesions and 3.8–8.2 erosions in patients with axSpA (nr-axSpA and AS)^{18,19}. The extent of these lesions seem to correspond to the cut-offs we find to be specific for axSpA in our study. Unfortunately, due to the absence of a control group in these studies it is impossible to confirm the specificity using our cut-offs in their studies. But Weber et al included patients with axSpA as well as patients with non-specific back pain and healthy controls when defining optimal cut-offs for MRI-SI. They reported data from two cohorts and found that ≥ 1 erosion (both cohorts) and ‘ $\geq 12/\geq 9$ fatty lesions’ (cohort A/B) yielded a specificity of $\geq 90.0\%$. When looking at lesion combinations they found ‘ $\geq 12/\geq 11$ fatty lesions and/or erosions’ (cohort A/B) to be most specific for axSpA²⁰. There is a difference in study design compared with our study. Weber et al classified patients based on the clinical opinion of a rheumatologist with access to the patient's clinical data and pelvic radiographs. We have classified patients according to the ASAS criteria. By doing so, we deliberately not interpreted the MRI features under study. Previously, we have shown that there is high agreement between the rheumatologist's diagnosis and fulfilment of the ASAS criteria in the SPACE cohort¹³. It might be a reasonable approach to consider fatty lesions in conjunction with erosions rather than each separately: one study has shown that fatty lesions alone only

contribute marginally to recognising patients with axSpA, while the majority of the patients with axSpA have fatty lesions alongside inflammation or erosions²¹. Summarising the literature, evidence is found that erosions have more diagnostic value than fatty lesions (alone). In keeping with our findings, a useful next step in determining the diagnostic utility of MRI-SI could be to test if adding '≥3 fatty lesions' as a criterion to the ASAS definition of a positive MRI-SI would increase the diagnostic performance. Similar analyses could be conducted with criteria such as '≥3 erosions' and the combination of '≥5 fatty lesions and/or erosions'.

The validation of our findings in independent cohorts could provide insight into the dilemma if and which structural lesions should be added to the ASAS definition of a positive MRI. Another relevant question is if MRI somehow could replace the role of plain pelvic X-rays in the determination of a positive imaging result.

In our study we have shown that reasonable and applicable cut-offs can be defined for inflammatory and fatty lesions on MRI-spine when specificity (>95%) for axSpA is the main aim. However, there are some conflicting data in the literature. A recent study states that adding spinal inflammatory and fatty lesions, in the classification of patients with axSpA, does not help since a better classification of patients with nr-axSpA is offset by increased misclassification of controls²². However, with a suitable cut-off value the number of false positives could be reduced. In our analyses, only '≥5 inflammatory lesions' and '≥5 fatty lesions' are having >95% specificity. This cut-off of '≥5 inflammatory lesions' is higher than the proposed ASAS definition of a positive MRI-spine (≥3), but Weber et al report an even higher cut-off (≥6 inflammatory lesions) corresponding to moderate to substantial diagnostic utility in patients with axSpA, while their cut-off of ≥6 fatty lesions only resulted in moderate specificity (82%/81% for two cohorts). In addition, like us, they found that combining spinal inflammatory and fatty lesions does not increase specificity²³. Also, Bennett et al found that >5 fatty lesions had diagnostic utility for axSpA²⁴. Comparing our data with the literature, even with different readers, different cohorts and different choices of external standard, we have arrived at approximately similar cut-offs for inflammatory and fatty spinal lesions (≥5 for both) with proven preservation of high specificity.

When the ASAS criteria were developed, MRI-spine was not included because the specificity of spinal lesions was argued⁶. The aforementioned study by Weber et al²² reported that combining MRI-spine and MRI-SI added little diagnostic value compared with MRI-SI alone. Testing our proposed cut-offs in independent cohorts could provide further insight in this matter.

The major limitation in studies testing the diagnostic value of imaging in axSpA is the lack of a 'gold standard'. In this study, patients were classified according to the ASAS criteria. Since we have investigated lesions not (yet) included in the ASAS criteria, circular reasoning is likely irrelevant here. An alternative external standard could be the diagnosis by the rheumatologist. However, physicians use similar clinical data to base their opinion on and may have had access to imaging data that could influence their opinion about the presence/absence of axSpA. A definite strength of this study is the availability of a no SpA control group that allows the determination of true specificity. Having multiple readers involved in the image scoring and this being a multicentre cohort study markedly added to the credibility of the final judgement about the presence or absence of structural lesions and the generalisability to other independent cohorts.

In conclusion, the presence of (1) at least five fatty lesions and/or erosions on MRI-SI, (2) at least five inflammatory lesions or (3) at least five fatty lesions on MRI-spine, allows an acceptable discrimination of axSpA and no SpA, while assuring >95% specificity.

References

1. Van den Berg R, Lenczner G, Feydy A, et al. Agreement between clinical practice and trained central reading in reading of sacroiliac joints on plain pelvic radiographs: Results from the DESIR cohort. *Arthritis Rheumatol* 2014;66:2403-11.
2. Poddubnyy D, Gaydukova I, Hermann KG, 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.
3. 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.
4. Mau W, Zeidler H, Mau R, et al. Clinical features and prognosis of patients with possible ankylosing spondylitis. Results of a 10-year followup. *J Rheumatol* 1988;15:1109-14.
5. Van Tubergen A, Heuft-Dorenbosch L, Schulpens G, et al. Radiographic assessment of sacroiliitis by radiologists and rheumatologists: does training improve quality? *Ann Rheum Dis* 2003;62:519-25.
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. Van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. *Arthritis Rheum* 1984;27:361-8.
8. Rudwaleit M, Jurik AG, Hermann KG, et al. Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. *Ann Rheum Dis* 2009;68:1520-7.
9. Weber U, Lambert RGW, Østergaard M, et al. The diagnostic utility of magnetic resonance imaging in spondylarthritis: An international multicenter evaluation of one hundred eighty-seven subjects. *Arthritis Rheum* 2010;62:3048-3058.
10. Hermann KG, Baraliakos X, van der Heijde D, et al. Descriptions of spinal MRI lesions and definition of a positive MRI of the spine in axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI study group. *Ann Rheum Dis* 2012;71:1278-88.
11. Lambert RGW, Pedersen SJ, Maksymowych WP, et al. Active inflammatory lesions detected by magnetic resonance imaging in the spine of patients with spondyloarthritis - Definitions, assessment system, and reference image set. *J Rheumatol Suppl* 2009;36:3-17.
12. Østergaard M, Maksymowych WP, Pedersen SJ, et al. Structural lesions detected by magnetic resonance imaging in the spine of patients with spondyloarthritis - Definitions, assessment system, and reference image set. *J Rheumatol Suppl* 2009;36:18-34.
13. Van den Berg R, de Hooze 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. *Rheumatol (Oxford)* 2013;52:1492-9.
14. De Hooze M, van den Berg R, Navarro-Compán V, et al. Magnetic resonance imaging of the sacroiliac joints in the early detection of spondyloarthritis: No added value of gadolinium compared with short tau inversion recovery sequence. *Rheumatol (Oxford)* 2013;52:1220-4.
15. Rudwaleit M, van der Heijde D, Khan MA, et al. How to diagnose axial spondyloarthritis early. *Ann Rheum Dis* 2004;63:535-43.
16. Wick MC, Weiss RJ, Jaschke W, et al. Erosions are the most relevant magnetic resonance imaging features in quantification of sacroiliac joints in ankylosing spondylitis. *J Rheumatol* 2010;37:622-7.
17. Weber U, Lambert RGW, Pedersen SJ, et al. Assessment of structural lesions in

- sacroiliac joints enhances diagnostic utility of magnetic resonance imaging in early spondylarthritis. *Arthritis Care Res (Hoboken)* 2010;62:1763-71.
18. Althoff CE, Sieper J, Song IH, et al. Active inflammation and structural change in early active axial spondyloarthritis as detected by whole-body MRI. *Ann Rheum Dis* 2013; 72:967-73.
19. Krohn M, Braum LS, Sieper J, et al. Erosions and fatty lesions of sacroiliac joints in patients with axial spondyloarthritis: Evaluation of different mri techniques and two scoring methods. *J Rheumatol* 2014;41:473-80.
20. Weber U, Zubler V, Pedersen SJ, et al. Development and validation of a magnetic resonance imaging reference criterion for defining a positive sacroiliac joint magnetic resonance imaging finding in spondyloarthritis. *Arthritis Care Res* 2013;65:977-85.
21. Weber U, Pedersen SJ, Zubler V, et al. Fat infiltration on magnetic resonance imaging of the sacroiliac joints has limited diagnostic utility in nonradiographic axial spondyloarthritis. *J Rheumatol* 2014;41:75-83.
22. 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* 2015;74:985-92.
23. Weber U, Zhao Z, Rufibach K, et al. Diagnostic utility of candidate definitions for a positive MRI of the spine in patients with axial spondyloarthritis. *Arthritis Rheum* 2015;67:924-33.
24. Bennett AN, Rehman A, Hensor EM, et al. The fatty Romanus lesion: a non-inflammatory spinal MRI lesion specific for axial spondyloarthropathy. *Ann Rheum Dis* 2010;69:891-4.

6

Reliability of mSASSS scoring in everyday practice in DESIR cohort study centres: cross-sectional study of agreement with trained readers

Pascal Claudepierre*, Manouk de Hooge*, Antoine Feydy, Monique Reijnierse, Alain
Saraux, Maxime Dougados, Désirée van der Heijde

*Both authors contributed equally

Ann Rheum Dis (accepted for publication)

The assessment of axial SpondyloArthritis (axSpA) lesions is important; not only to evaluate a major contributor to disease severity, but also to determine whether drugs are effective in inhibiting structural progression. In trials, assessment of axSpA lesions can be done locally in a study center, or in a centralised manner by specifically trained readers. The potential impact of choosing one method rather than the other is unknown. In earlier work, we found the same -moderate- level of agreement between local reading (LocR) and central reading (CentR) as between the 2 central readers regarding radiographic sacroiliitis assessment¹. However, differences between LocR and CentR resulted in misclassification of some patients (AS or axSpA meeting ASAS criteria)². The modified Stoke Ankylosing Spondylitis Spine Score (mSASSS) is the recommended tool to assess erosions, sclerosis, squaring and (bridging) syndesmophytes at each anterior vertebral corner of the cervical and lumbar spine^{3,4}. The objective of this study was to assess agreement between baseline mSASSS values of patients with recent-onset inflammatory back pain (DESIR cohort) determined by local readers and by two specifically trained readers.

Local readers were “experienced” because of their medical practice but did not participate in any calibration nor training session about mSASSS scoring. The central readers participated in a calibration session led by four AS experts (DvdH, MD, MR and AF), in which definitions, examples and pitfalls were discussed. Subsequently, the two readers independently scored a training set of 20 radiographs ($\kappa=0.61$). After a consensus meeting was held with the same four AS experts, the two readers started reading for this study. An adjudicator (AF) assessed the mSASSS when initial readers disagreed on the presence of syndesmophytes. Consequently, the mSASSS in the analysis was either the mean score of the two initial readers, either the mean of the adjudicator score and the score of the initial reader closest to the adjudicator.

When dividing the 635 patients into 2 groups ($mSASSS < 1$ or $mSASSS \geq 1$), agreement was good between central readers and only fair between LocR and CentR. Disagreement between two central readers was balanced, with neither reader overestimating mSASSS relative to the other reader, in contrast to LocR who systematically overestimated the radiographic spinal lesions when compared to CentR (figure 1a, 1b and 2). There were 121 false-positive cases when CentR was considered as external standard; in 74% of these cases, mSASSS was overestimated by LocR of 2 or more points and in 22% even at least 5 points. Of the 19 false-negative cases, in 26.3% underestimation was by a single point and in 10.5% underestimation was by ≥ 5 points. When separating patients with $mSASSS < 2$ from those with $mSASSS \geq 2$, agreement improved between two central readers but not between LocR and CentR

(figure 1c and 1d). This indicates that most of the disagreement was related to local readers, not to central readers.

Syndesmophytes were scored differently between LocR and CentR for 79 patients: 31 cases had one or more syndesmophytes according to CentR and no syndesmophytes according to LocR, opposite results were seen in 48 cases. Analysis of the local results showed false-positives (≥ 1 syndesmophyte not identified by CentR) in 36 cases and false-negatives (no syndesmophyte while CentR scored ≥ 1 syndesmophyte) in 14 cases.

When radiographic severity in axSpA patients is assessed based on mSASSS or number of patients with syndesmophytes, overestimation of radiographic spinal lesions by LocR would translate into overestimation of radiographic severity. This overestimation may indicate challenges in distinguishing degenerative lesions from axSpA lesions, a point that was addressed in the calibration sessions attended by the central readers. These data support the use of radiographic spinal scores determined by specifically trained central readers rather than local readers at various centres in clinical trials, and highlights the need for specific training when assessing axSpA spinal lesions in clinical practise.

a		Central reader 2			b		Central reading		
		mSASSS ≥ 1	mSASSS < 1				mSASSS ≥ 1	mSASSS < 1	
Central reader 1	mSASSS ≥ 1	41	19	60	Local reading	mSASSS ≥ 1	39	121	160
	mSASSS < 1	16	559	575		mSASSS < 1	19	456	475
		57	578	635			58	577	635
$\kappa=0.67$ / Percentage positive agreement = 70.1%					$\kappa=0.26$ / Percentage positive agreement = 35.8%				
c		Central reader 2			d		Central reading		
		mSASSS ≥ 2	mSASSS < 2				mSASSS ≥ 2	mSASSS < 2	
Central reader 1	mSASSS ≥ 2	39	11	50	Local reading	mSASSS ≥ 2	29	97	126
	mSASSS < 2	8	577	585		mSASSS < 2	16	493	509
		47	588	635			45	590	635
$\kappa=0.79$ / Percentage positive agreement = 80.4%					$\kappa=0.26$ / Percentage positive agreement = 33.9%				

Figure 1: Agreement about the mSASSS between two central readers and between LocR and CentR

a: Agreement between two central readers comparing mSASSS < 1 vs. ≥ 1

b: Agreement between LocR and CentR comparing mSASSS < 1 vs. ≥ 1

c: Agreement between two central readers comparing mSASSS < 2 (indicating no lesions or a single grade-1 lesion) vs. ≥ 2

d: Agreement between LocR and CentR comparing mSASSS < 2 (indicating no lesions or a single grade 1 lesion) vs. ≥ 2

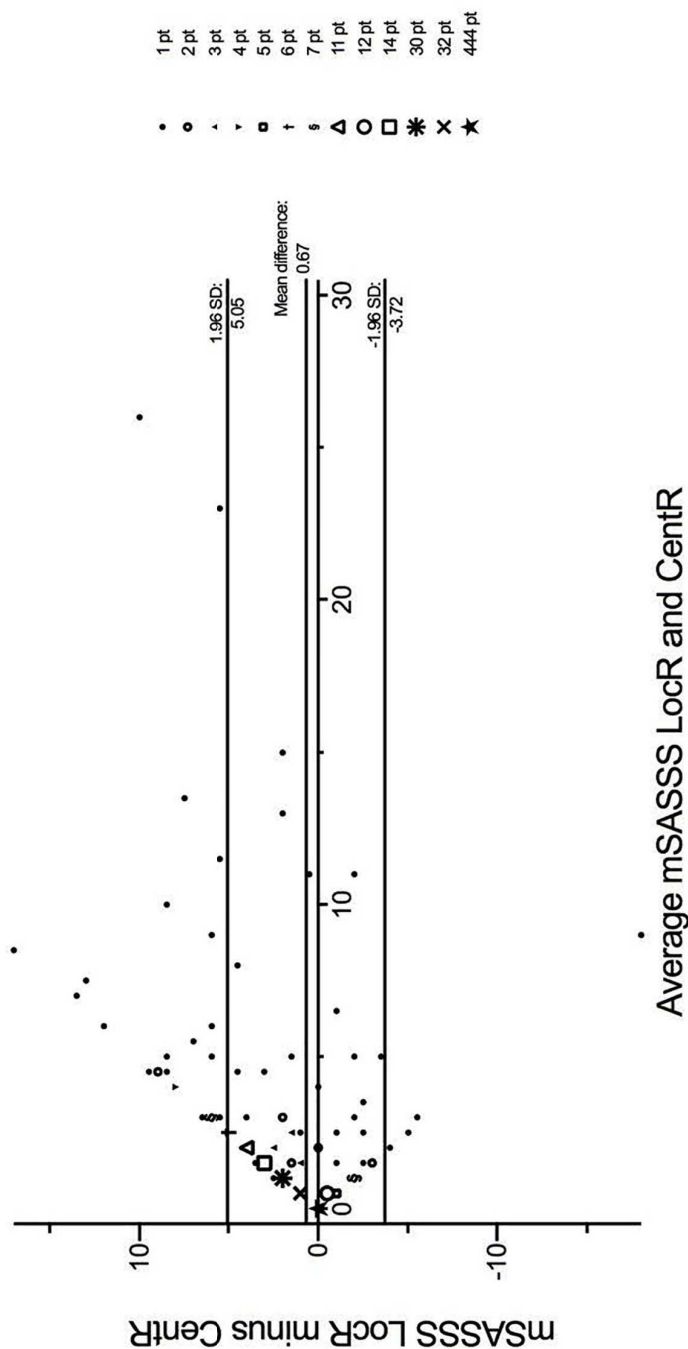


Figure 2: Bland-Altman plot
The differences between mSASSS values assigned by local and trained readers were plotted against the mean of the two values. Each type of symbol represents a specific number of patients (pt).
X axis: Mean of LocR and CentR for mSASSS
Y axis: Scores of LocR minus CentR for mSASSS

References

1. van den Berg R, Lenczner G, Feydy A, et al. Agreement between clinical practice and trained central reading in reading of sacroiliac joints on plain pelvic radiographs: Results from the DESIR cohort. *Arthritis Rheumatol* 2014;66:2403-11.
2. van den Berg R, Lenczner G, Thévenin F, et al. 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* 2015;74:2016-21.
3. van Der Heijde D, Landewé R. Selection of a method for scoring radiographs for ankylosing spondylitis clinical trials, by the Assessment in Ankylosing Spondylitis Working Group and OMERACT. *J Rheumatol* 2005;10:2048-9.
4. Creemers MCW, Franssen MJAM, van 't Hof MA, et al. Assessment of outcome in ankylosing spondylitis: an extended radiographic scoring system. *Ann Rheum Dis* 2005;64:127-9.

7

Assessment of typical SpA lesions on MRI of the spine: do local readers and central readers agree in the DESIR cohort at baseline?

Manouk de Hooge, Jean-Baptiste Pialat, Monique Reijnierse, Désirée van der Heijde, Pascal Claudepierre, Alain Saraux, Maxime Dougados, Antoine Feydy

Submitted

Abstract

Objective

Comparing local reading (LocR) with central reading (CentR) of typical spondyloarthritis lesions including bone marrow edema (BME) and structural lesions on Magnetic Resonance Imaging of the spine (MRI-spine), in patients with inflammatory back pain (IBP; ≥ 3 months, < 3 years).

Methods

Baseline data of 667 patients, age 18-50 years, from the Devenir des Spondylarthropathies Indifférenciées Récentes DESIR cohort were used. Two trained central readers scored anterior and posterior corner BME, fatty lesions, erosions, and syndesmophytes on MRI-spine. Presences of lesions, based on average scores, were used for CentR. A local radiologist and/or rheumatologist scored MRI-spine on presence/doubt/absence of 'inflammation' and 'structural lesions'. Agreement between central readers and readings was calculated (Cohen's Kappa's; κ).

Results

Agreement between central readers was moderate (BME $\kappa=0.55$, fatty lesions $\kappa=0.50$) to slight (erosions $\kappa=0.12$, syndesmophytes $\kappa=0.19$). Agreement between LocR and CentR was $\kappa=0.32$ (BME) and $\kappa=0.13$ (structural lesions). In 78/160 patients (48.8%) LocR were in doubt while CentR scored BME lesions, for structural lesions this was 17.8% (28/157 patients).

Conclusion

Agreement between 2 central readers for scoring spondyloarthritis-like lesions on MRI-spine was moderate but better compared to LocR and CentR agreement. LocR often doubt about the presence of MRI-spine lesions while central trained readers score lesions.

Introduction

In the field of axial spondyloarthritis (axSpA) the role of magnetic resonance imaging (MRI) in daily practice and studies increased over the past years. Sacroiliitis on MRI (MRI-SI) is part of the Assessment of SpondyloArthritis international Society (ASAS) classification criteria for axSpA¹. Although spinal MRI-lesions are not part of the criteria, the interest in these lesions is growing. Bone marrow edema (BME) on MRI of the spine (MRI-spine) is increasingly used to monitor the effect of anti-inflammatory therapy in axSpA patients²⁻⁵. Activity on either MRI-SI or MRI-spine is frequently used to define if patients have a 'positive MRI' used as prognostic factor to treatment response in patients with non-radiographic axSpA⁶. Besides this, spinal inflammation in early disease and structural damage on MRI-spine in later disease seem to be related to functional impairment in AS patients⁷.

In clinical trials and cohort studies there are always trained readers providing the scores. This is different for daily practice where a local radiologist (in consensus with a rheumatologist) performs the MRI readings aware of the clinical and biological data. In daily practice and clinical studies, readers are assumed to assess the compatibility of MRI-lesions with axSpA. However it is unknown whether there are discrepancies in assessment between daily practice and central readers, and if readings of specialists in daily practice are adequate without specific training concerning the recognition of typical MRI-lesions associated with axSpA. One study compared local and central readings concerning radiographic sacroiliitis and they found a moderate agreement between central readers ($\kappa=0.54$) and between central vs. local readers ($\kappa=0.55$)⁸. Disagreement between central readers was balanced in two directions while local readers report a large proportion of false positives and a small proportion of false negatives, when central reading was considered as external standard. To our knowledge there are no studies addressing this for MRI-spine readings. Therefore the objective of this study was to compare results of local reading (LocR) to central reading (CentR) as external standard for BME and structural lesions seen on MRI-spine, in patients with inflammatory back pain (IBP) of relative short duration.

Methods

Study population, data collection and classification

Baseline data from patients included in the Devenir des Spondylarthropathies Indifférenciées Récentes (DESIR) cohort was used. Inclusion period was from December 2007 until April 2010 in 25 regional centres in France, where patients aged 18-50 years with IBP according to the Calin or Berlin criteria^{9,10}, persisting ≥ 3 months

but <3 years, were included. A detailed description of exclusion criteria can be found elsewhere¹¹. The DESIR cohort was conducted in accordance with the Declaration of Helsinki and approved by the local ethical committee and health authorities. Written informed consent was obtained from participating patients before inclusion. Study population and data collection assessment of the DESIR cohort have previously been published¹². The study was registered on clinicaltrials.gov (ID: #NCT01648907) and the clinical database used for the current study was locked on October 30th 2012.

In this study, patients were classified according to the ASAS classification criteria for axSpA based on the central imaging readings (MRI and radiographs) of the sacroiliac (SI) joints¹³.

Imaging

Patients enrolled in the DESIR cohort underwent an MRI-spine performed on a 1T-1.5T scanner with acquired sequences T1-weighted Turbo Spin-Echo (TR500-700/TE10-55) and Short Tau Inversion Recovery (TR4000/TE50-70/TI130-160 (1T);140-170msec (1.5T)), with 4mm slice thickness. MRI-spine was performed in sagittal plane and both sequences were viewed simultaneously. Upper (C2-T10) and lower part of MRI-spine (T8-S1) were conducted separately with an overlap of at least 2 vertebrae. In total 708 patients were enrolled in the DESIR cohort but since one or both sequences were missing in 41 patients, MRI-spine data of 667 patients were available.

Central reading

Two central readers (MdH and JBP), both familiar with scoring MRI-spine, participated in a calibration session before starting the reading. The calibration session was a systematic conducted exercise, executed by two senior radiologists (MR and AF) and two senior rheumatologists (DvdH and MD), who already did such calibration sessions before. During the calibration process, definitions of lesions, examples and pitfalls were discussed. Subsequently, the two readers independently read a training set of 20 MRI-spine, followed by the calculation of agreement based on the presence of ≥ 2 BME lesions ($\kappa=0.60$) and ≥ 3 fatty lesions ($\kappa=0.47$) and a consensus meeting where the two readers discussed discrepancies, difficult images and possible arrangements concerning the procedures of scoring. After this meeting, the two readers started to read the baseline images of the DESIR cohort.

The central readers independently scored all MRI-spine images, while blinded for clinical and other imaging data. Anterior and posterior corner BME, fatty lesions, erosions, and syndesmophytes were scored per vertebral unit (VU). In total 23 VUs

were scored with C2-C3 as VU1 and L5-S1 as VU23. Central readers scored all lesions only when considered typical for axSpA. When the readers considered the lesions to be due to other causes, like degenerative changes or Scheuermann disease, lesions were not scored. BME and fatty lesions suggestive of spondylitis were scored when visible on ≥ 2 consecutive slices. For erosions and syndesmophytes suggestive of spondylitis presence on ≥ 1 slice was sufficient.

When central readers disagreed on the presence of ≥ 2 BME, an adjudicator (AF) scored all 23 VUs on inflammation. When central readers disagreed on the presence of ≥ 3 fatty lesions, the adjudicator provided scores in all 23 VUs on the presence of structural lesions for fatty lesions, erosions and syndesmophytes separately. When there was disagreement on both BME and fatty lesions, the adjudicator provided the scores for all MRI-spine lesions. The cut-off values for BME (≥ 2) and fatty lesions (≥ 3) were chosen arbitrarily since, at the time, there was no official definition of a positive MRI-spine.

The central reading (CentR), presence/absence of MRI-lesions, was obtained from the average of both central readers scores or, in case of adjudication, the average of the adjudicator score and central reader closest to this score. Where presence represented an average score of ≥ 1 and absence an average score of < 1 lesion. This was done for BME and structural lesions separately.

In addition, the central readers gave an overall verdict whether MRI-spine fulfilled the ASAS consensus definition of a positive MRI. In this definition a spinal MRI is considered positive when there are ≥ 3 corner BME, each seen on ≥ 2 consecutive sagittal slices¹⁴.

Local reading

In each participating centre, a local radiologist or rheumatologist decided on the presence of 'BME' and 'structural lesions' on MRI-spine. The local readers were not trained apart from the usual medical education. The local scoring was different and more global compared to the central scoring. Local readers separately scored 'BME' and 'structural lesions', without further distinction for different types of structural lesions. The verdict options the local readers could choose were 'presence', 'doubt' or 'absence' of lesions separately for cervical, thoracic and lumbar spine. These scores will be referred to as 'local reading' (LocR).

Statistical analysis

Agreement between central readers as well as between LocR and CentR was expressed in Cohen's Kappa's (κ) and percentages positive agreement (PPA). PPA is the number of positive readings scored by both readers/readings divided by all of the positive readings of either readers/readings¹⁵. The strength of agreement related to κ value was interpreted as suggested by Landis and Koch: 0.00-0.20 corresponds to slight agreement, 0.21-0.40 to fair, 0.41-0.60 to moderate, 0.61-0.80 to substantial and 0.81-1.00 to a (almost) perfect agreement¹⁶. All statistical analyses were performed with SPSS version 22.0 (SPSS Inc., Chicago, IL, USA).

Results

LocR and CentR of MRI-spine were available in 667 (BME) and 666 (structural lesions) patients. Less than half of the patients ($n=306$; 45.9%) were male. The mean age at onset back pain was 31.9 years ($SD \pm 8.7$). Human leukocyte antigen (HLA)-B27 was positive in 58.9% of the patients. There were 596 patients with an onset of IBP <45 years, in whom the ASAS classification criteria for axSpA could be applied. Based on the central imaging readings (MRI and radiographs) of the sacroiliac joints 430 patients fulfilled the ASAS criteria; 242 patients based on sacroiliitis on imaging (imaging arm) and 188 patients based on HLA-B27 positivity (clinical arm).

Central reading agreement

In 86/667 patients adjudication on BME and in 63/667 patients adjudication on structural lesions was needed. Table 1 shows the agreement between central readers on the presence/absence (≥ 1) for all spinal MRI-lesions. The agreement for BME ($\kappa=0.55$) and fatty lesions ($\kappa=0.50$) was moderate and the agreement for erosions ($\kappa=0.12$) and syndesmophytes ($\kappa=0.19$) was slight only. The disagreement on BME lesions was balanced between the readers: in 54 and 64 patients. However, the distribution in disagreement was unequal concerning structural lesions. Central reader 1 scored many more lesions than central reader 2, especially fatty lesions.

Table 1 also shows the agreement based on a 'positive MRI-spine' according to the ASAS definition ($\kappa=0.58$) and this was similar to the agreement of central readers on any BME lesion present.

Table 1: Agreement between central reader pair on BME and structural spinal MRI lesions.

Central reader 2	Central reader 1	
	BME lesion present	BME lesion absent
	142	54
	BME lesion absent	407
	Kappa = 0.58 / Percentage positive agreement = 54.6%	
	Central reader 1	
	Fatty lesion present	Fatty lesion absent
	82	4
	Fatty lesion absent	470
	Kappa = 0.50 / Percentage positive agreement = 41.6%	
Central reader 2	Central reader 1	
	Erosion present	Erosion absent
	7	8
	Erosion absent	580
	Kappa = 0.12 / Percentage positive agreement = 8.0%	
	Central reader 1	
	Syndesmophyte present	Syndesmophyte absent
	11	18
	Syndesmophyte absent	587
	Kappa = 0.19 / Percentage positive agreement = 13.8%	
Central reader 2	Central reader 1	
	MRI-spine +	MRI-spine -
	77	58
	MRI-spine -	509
	Kappa = 0.58 / Percentage positive agreement = 48.7%	
	Central reader 1	
Central reader 2	Structural lesion present	Structural lesion absent
	86	9
	Structural lesion absent	426
	Kappa = 0.41 / Percentage positive agreement = 35.7%	

Agreement between LocR and CentR

In 91 patients (13.6%) spinal BME lesions were present according to LocR. In 118 patients (17.7%) there was doubt about the presence of BME lesions. LocR scored structural lesions less frequently: in 50 patients (7.5%) structural lesions were scored 'present' and in 48 patients (7.2%) there was doubt. CentR scored BME lesions in 160 patients (24.0%) and structural lesions in 157 patients (23.6%).

Table 2 shows the agreement between LocR and CentR. Agreement on BME lesions ($\kappa=0.27$) was slightly higher than on structural lesions ($\kappa=0.13$), but both were lower compared to the agreement between the two central readers. Also when looking at the individual scores of the central readers the agreement with LocR on BME lesions (reader 1 $\kappa=0.32$; reader 2 $\kappa=0.25$) and structural lesions (reader 1 $\kappa=0.09$; reader 2 $\kappa=0.09$) was slight (table 2). Table 2 also showed the agreement between LocR and CentR when LocR score 'doubt' would be considered as 'lesion present' or as 'lesion

absent'. When doubtful lesions were considered present the agreement increased, indicating that a lot of the lesions considered doubtful by LocR are scored present by CentR. Of the 160 patients in whom CentR scored BME lesions, 48.8% (n=78) got a doubt score by LocR. LocR doubted about the presence of structural lesions in 17.8% (n=28) of the 157 patients in whom CentR scored structural lesions.

In 492 patients, a radiologist performed the LocR. In 206/205 (BME/structural lesions) patients rheumatologists provided the LocR. There is an overlap of 31 patients in which the LocR (BME and structural lesions) was given by a radiologist and rheumatologist. When LocR was given by a radiologist agreement was higher compared to LocR given by a rheumatologist. For BME lesions there was a big difference in agreement of LocR and CentR between specialists; $\kappa=0.36$ for radiologists and $\kappa=0.006$ for rheumatologists. This difference was less for structural lesions; $\kappa=0.15$ for radiologists and $\kappa=0.12$ for rheumatologists (table 3).

Agreement per spinal segment

Cervical BME lesions were present in 16 patients according to LocR and 17 patients according to CentR. There were 33 (LocR) and 57 (CentR) patients with thoracic BME lesions and 55 (LocR) and 55 (CentR) patients with lumbar BME lesions. However, LocR and CentR agreed on the presence of BME lesions in 2 (cervical), 13 (thoracic) and 15 (lumbar) patients. LocR scored the presence of structural lesions in 9 (cervical), 27 (thoracic) and 23 patients (lumbar). CentR scored structural lesions in 19 (cervical), 91 (thoracic) and 67 patients (lumbar). LocR and CentR agreed in only 1 (cervical), 12 (thoracic) and 5 (lumbar) patients on the presence of structural lesions (table 4).

In general the agreement on BME as well as on structural lesions was the lowest in the cervical spine. Agreement between the central readers was moderate (cervical $\kappa=0.45$; lumbar $\kappa=0.55$) to substantial (thoracic $\kappa=0.64$) for BME and fair (cervical $\kappa=0.22$) to moderate (thoracic $\kappa=0.51$; lumbar $\kappa=0.45$) for structural lesions. The agreement between LocR and CentR was lower than between the central readers. κ between LocR and CentR was the lowest in cervical spine, then lumbar spine and though κ was the highest in thoracic spine it was still rather low (table 4).

Only for cervical structural lesions, LocR by rheumatologists had a higher agreement with CentR ($\kappa=0.16$) than when LocR was performed by radiologists ($\kappa=0.00$). In the rest of the segments agreement between LocR by radiologists and CentR was higher compared to LocR by rheumatologists.

Table 2: Agreement between local (LocR) and central (CentR) reading on the presence/absence of BME and structural spinal MRI lesions

LocR	CentR		
		BME lesion present	BME lesion absent
	BME lesion present	33	58
	BME lesion absent	49	409
	Doubt	78	40
	Kappa (leaving out 'doubt' group) = 0.27; PPA = 23.6%		
	Kappa ('doubt' group considered as 'lesion present') = 0.45; PPA = 43.0%		
	Kappa ('doubt' group considered as 'lesion absent') = 0.11; PPA = 15.1%		
	Central reader 1		
		BME lesion present	BME lesion absent
LocR	BME lesion present	47	44
	BME lesion absent	73	385
	Doubt	86	32
	Kappa (leaving out 'doubt' group) = 0.32; PPA = 28.7%		
	Kappa ('doubt' group considered as 'lesion present') = 0.48; PPA = 47.2%		
	Kappa ('doubt' group considered as 'lesion absent') = 0.16; PPA = 18.8%		
	Central reader 2		
		BME lesion present	BME lesion absent
	BME lesion present	40	51
	BME lesion absent	74	384
LocR	Doubt	82	36
	Kappa (leaving out 'doubt' group) = 0.25; PPA = 24.2%		
	Kappa ('doubt' group considered as 'lesion present') = 0.43; PPA = 43.1%		
	Kappa ('doubt' group considered as 'lesion absent') = 0.11; PPA = 16.2%		
	CentR		
		Structural lesion present	Structural lesion absent
	Structural lesion present	21	29
	Structural lesion absent	108	460
	Doubt	28	20
	Kappa (leaving out 'doubt' group) = 0.13; PPA = 13.3%		
LocR	Kappa ('doubt' group considered as 'lesion present') = 0.55; PPA = 23.8%		
	Kappa ('doubt' group considered as 'lesion absent') = 0.10; PPA = 11.3%		
	Central reader 1		
		Structural lesion present	Structural lesion absent
	Structural lesion present	26	24
	Structural lesion absent	172	396
	Doubt	34	14
	Kappa (leaving out 'doubt' group) = 0.09; PPA = 11.7%		
	Kappa ('doubt' group considered as 'lesion present') = 0.20; PPA = 22.2%		
	Kappa ('doubt' group considered as 'lesion absent') = 0.07; PPA = 10.2%		
LocR	Central reader 2		
		Structural lesion present	Structural lesion absent
	Structural lesion present	11	39
	Structural lesion absent	61	507
	Doubt	23	25
	Kappa (leaving out 'doubt' group) = 0.09; PPA = 9.9%		
	Kappa ('doubt' group considered as 'lesion present') = 0.24; PPA = 21.4%		
	Kappa ('doubt' group considered as 'lesion absent') = 0.06; PPA = 8.2%		

Table 3. Agreement between local (LocR) and central (CentR) reading on presence/absence of spinal BME and structural MRI lesions. LocR reported separately when performed by radiologists or rheumatologists.

LocR	BME	CentR			
				Present	Absent
		Radiologist	Present	30	41
			Absent	29	304
			Doubt	58	30
		Kappa (leaving out 'doubt' group) = 0.36; PPA = 30.0%			
		Kappa ('doubt' group considered as 'lesion present') = 0.50; PPA = 46.8%			
		Kappa ('doubt' group considered as 'lesion absent') = 0.17; PPA = 19.0%			
		CentR			
				Present	Absent
Structural	Rheumatologist	Present	4	24	
		Absent	20	126	
		Doubt	21	11	
	Kappa (leaving out 'doubt' group) = 0.006; PPA = 8.3%				
	Kappa ('doubt' group considered as 'lesion present') = 0.30; PPA = 31.3%				
	Kappa ('doubt' group considered as 'lesion absent') = 0.00; PPA = 5.8%				
	Radiologist	Present	16	15	
		Absent	87	338	
		Doubt	21	15	
		Kappa (leaving out 'doubt' group) = 0.15; PPA = 13.6%			
		Kappa ('doubt' group considered as 'lesion present') = 0.26; PPA = 24.0%			
		Kappa ('doubt' group considered as 'lesion absent') = 0.12; PPA = 11.5%			
	Rheumatologist	Present	6	14	
		Absent	26	146	
		Doubt	7	6	
	Kappa (leaving out 'doubt' group) = 0.12; PPA = 13.0%				
	Kappa ('doubt' group considered as 'lesion present') = 0.23; PPA = 22.0%				
	Kappa ('doubt' group considered as 'lesion absent') = 0.09; PPA = 11.3%				

Discussion

In this study we compared BME and structural lesions scored on MRI-spine by local and central readers in patients included in the DESIR cohort. We found that agreement between central readers is moderate at most but agreement between LocR and CentR was even lower.

Until now the DESIR cohort is unique in having reported data of both local and central assessments of imaging readings. Van der Berg et al. showed a moderate agreement at best ($\kappa=0.55$), between LocR and CentR of radiographic sacroiliitis⁸.

In another study, van der Berg et al. found a substantial agreement ($\kappa=0.70$) between LocR and CentR concerning the presence of sacroiliitis on MRI¹⁷. In both studies

Table 4. Agreement per spinal segment between LocR and CentR for the presence of inflammatory and structural spinal MRI lesions.

		Inflammatory lesions			Structural lesions		
		Cervical	Thoracic	Lumbar	Cervical	Thoracic	Lumbar
Central reader 1	Kappa	0.45	0.64	0.55	0.22	0.51	0.45
vs central reader 2	PPA in %	19.0	26.6	24.0	10.9	22.5	20.8
Central reader 1	Kappa	0.05	0.19	0.26	0.03	0.10	0.02
vs LocR*	PPA in %	3.3	11.3	14.2	2.3	7.6	3.9
Central reader 2	Kappa	0.14	0.27	0.17	0	0.07	0.03
vs LocR*	PPA in %	7.5	14.2	11.6	0	5.7	3.8
	Kappa	0.10	0.23	0.20	0.05	0.15	0.06
CentR vs LocR*	PPA in %	5.7	12.6	12.0	3.4	9.2	5.3
CentR vs LocR*	Kappa	0.12	0.31	0.25	0	0.14	0.07
(radiologist)	PPA in %	6.9	15.4	13.8	0	9.0	5.6
CentR vs LocR*	Kappa	0	0.08	0.04	0.16	0.18	0.03
(rheumatologist)	PPA in %	0	6.7	5.7	8.3	10.8	4.0

*LocR score 'doubt' was not taken into account for calculating kappa

PPA = percentage positive agreement

the agreement between LocR and CentR was similar to the agreement between central readers ($\kappa=0.54$ for radiographic sacroiliitis and $\kappa=0.73$ for sacroiliitis on MRI).

In this study, the agreement on inflammatory spinal lesions between central readers was similar to the agreement van den Berg et al reported for radiographic sacroiliitis. However, this is as far as similarities go, because we found a much lower agreement between LocR and CentR for spinal inflammatory as well as structural lesions compared to the agreement between LocR and CentR for sacroiliitis on radiographs or MRI^{8,17}. More important; in our study agreement between LocR and CentR was much lower compared to agreement between central readers; fair versus moderate for BME and slight versus moderate for structural lesions. Unfortunately LocR did not have the option to differentiate between the different types of structural lesions. Neither did they have the option to specify the quantity of any of the lesions. This limited the possibilities to compare the agreement between LocR and CentR.

In BME lesions, we see that disagreement between CentR and LocR was balanced in two directions, when not taking patients with 'doubt' score into account. But for structural lesions the disagreement was unequally distributed; when CentR would be considered as external standard, LocR showed a large proportion of false negatives and a small proportion of false positives. From our data it seems that LocR underrate MRI spinal lesions. CentR more often score BME and structural spinal lesions and in a majority of these positive cases LocR did not take a decision and used the possibility to express their doubt about the presence of MRI spinal lesions due to axSpA. CentR

did not have the possibility to score 'doubt'. This discrepancy in scoring options for LocR and CentR is an issue. LocR was a general and concise score and CentR a detailed and quantified score. The agreement between LocR and CentR improved when 'doubt' scores were allocated to the 'lesions present' group and decreased when they were considered as 'lesions absent'. Also the PPA dramatically increased when 'doubt' scores were considered 'lesions present'. This indicates that LocR do not score as many spinal lesions as CentR. When looking at the LocR of radiologists and rheumatologists separately we see that there is less agreement with CentR when rheumatologists perform LocR compared to radiologists. Especially when looking at BME, rheumatologists more often doubt about the presence of lesions.

Although we can only guess, it seems that LocR would only score MRI-spine lesions when they were very certain. This phenomenon is facilitated by the option for the LocR to neither score present nor absent. Also, local readers might not have enough confidence on the typical appearance of spinal lesions due to axSpA compared to lesions due to other causes like degenerative disc disease or Scheuermann. If this is the case, it seems that local readers are not skilled enough to judge whether spinal MRI-lesions are typical for axSpA and therefore often revert to the option 'doubt'. This scenario is plausible since it is considered difficult to distinguish between SpA and degenerative lesions¹⁸. Besides, it might be possible to see typical axSpA and degenerative lesions within one patient. Our data could be in line with and supportive to this explanation, since the lesions that were seen by the LocR were rarely confirmed by the CentR. Finally, it is likely that the local readers had access to clinical and other imaging data like MRI and radiographs of the SI joints. This extra information, which the central readers did not have, could influence the judgement of LocR. Taking this in consideration, it is interesting for future investigations to explore whether spinal lesions due to axSpA are seen on the same location as degenerative lesions.

It has been shown that training does not improve the reliability in judgment of sacroiliitis on radiographic or structural lesions seen on MRI-SI^{19,20}. We may speculate that training can improve agreement, since the agreement between (trained) central readers was better as compared to the agreement between central and local readers. However, the agreement between central readers was moderate at best and there is still a lot of room for improvement.

A limitation of this study was the difference in scoring procedure of LocR and CentR. Having the option to score 'doubt' in LocR may caused the low agreement between both readings since LocR often used this scoring option. As it is in several studies concerning MRI readings in axSpA, another limitation was the lack of a gold standard.

In conclusion, we found a moderate agreement between two central readers for scoring inflammatory and structural lesions seen on MRI-spine. This agreement was better compared to the agreement between LocR and CentR, which was only considered fair at most.

So, LocR do not score many lesions compared to CentR and doubt in the majority of cases about the presence of spinal MRI lesions.

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. Rudwaleit M, Listing J, Brandt J, et al. Prediction of a major clinical response (BASDAI 50) to tumour necrosis factor alpha blockers in ankylosing spondylitis. *Ann Rheum Dis* 2004;63:665-70.
3. Baraliakos X, Davis J, Tsuji W, et al. Magnetic resonance imaging examinations of the spine in patients with ankylosing spondylitis before and after therapy with the tumor necrosis factor alpha receptor fusion protein etanercept. *Arthritis Rheum* 2005;52:1216-23.
4. Braun J, Baraliakos X, Golder W, et al. Magnetic resonance imaging examinations of the spine in patients with ankylosing spondylitis, before and after successful therapy with infliximab: Evaluation of a new scoring system. *Arthritis Rheum* 2003;48:1126-36.
5. Braun J, Landewe R, Hermann K-GA, 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.
6. 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.
7. Machado P, Landewé R, Braun J, et al. Both structural damage and inflammation of the spine contribute to impairment of spinal mobility in patients with ankylosing spondylitis. *Ann Rheum Dis* 2010;69:1465-70.
8. van den Berg R, Lenczner G, Feydy A, et al. Agreement between clinical practice and trained central reading in reading of sacroiliac joints on plain pelvic radiographs: Results from the DESIR cohort. *Arthritis Rheumatol* 2014;66:2403-11.
9. Calin A, Porta J, Fries JF, et al. Clinical history as a screening test for ankylosing spondylitis. *JAMA* 1977;237:2613-4.
10. 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.
11. Dougados M, d'Agostino MA, Benessiano J, et al. The DESIR cohort: A 10-year follow-up of early inflammatory back pain in France: Study design and baseline characteristics of the 708 recruited patients. *Jt Bone Spine* 2011;78:598-603.
12. Dougados M, Etcheto A, Molto A, et al. Clinical presentation of patients suffering from recent onset chronic inflammatory back pain suggestive of spondyloarthritis: The DESIR cohort. *Jt Bone Spine* 2015;82:345-51.
13. van den Berg R, Lenczner G, Thévenin F, et al. 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* 2015;74:2016-21.
14. Hermann K, Baraliakos X, van der Heijde D, et al. Descriptions of spinal MRI lesions and definition of a positive MRI of the spine in axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI study group. *Ann Rheum Dis* 2012;71:1278-88.
15. Ahrens W, Pigeot I. *Handbook of Epidemiology*. New York, Springer-Verlag 2006:520-2.
16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
17. Van den Berg R, Thevenin F, Feydy A, Claudepierre P, Reijniers M, Saraux A, 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

- [abstract]. *Arthritis Rheum* 2013;65:S475.
18. Baraliakos X, Hermann KG, Braun J. Imaging in axial spondyloarthritis: Diagnostic problems and pitfalls. *Rheum Dis Clin North Am* 2012;38:513-22.
19. Poddubnyy D, Gaydukova I, Hermann KG, 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.
20. A van Tubergen, L Heuft-Dorenbosch, G Schulpen, et al. Radiographic assessment of sacroiliitis by radiologists and rheumatologists: does training improve quality? *Ann Rheum Dis* 2003;62:519-25.

8

Is site of back pain related to location of MRI lesion in patients with chronic back pain included in the SPACE cohort?

Manouk de Hooge*, Freek de Bruin*, Lieke de Beer, Pauline Bakker, Rosaline van den Berg, Sofia Ramiro, Floris van Gaalen, Karen Fagerli, Robert Landewé, Maikel van Oosterhout, Roberta Ramonda, Tom Huizinga, Hans Bloem, Monique Reijnierse, Désirée van der Heijde

*Both authors contributed equally

Arthritis Care Res (accepted for publication)

Abstract

Objective

To determine associations between MRI-lesions originating from either axial spondyloarthritis (axSpA) or from degeneration and pain in patients with chronic back pain of <2 years duration.

Methods

Patients from the SPondyloArthritis Caught Early cohort localized the sites of pain (thoracic, lumbar, buttock). Average MRI-scores from two readers for axSpA lesions and from two different readers for degenerative lesions were used. Associations between sacroiliac joint (SIJ) lesions and buttock pain were investigated by logistic regression analysis and associations between axSpA or degenerative lesions and pain in the spine (thoracic and lumbar) were investigated by generalized estimating equations. Interactions with gender, age, HLA-B27 and fulfilment of ASAS criteria were tested.

Results

In 348 patients (126 males, 127 fulfilling ASAS-criteria, mean age 29.4 years) spinal MRI (n=342 also SIJ images) were available. Pain was localized in the thoracic spine (35.9%), the lumbar spine (82.5%) or in the buttock(s) (57.8%). Inflammatory lesions of the SIJs (OR 1.06, $p=0.04$) -and erosions of the SIJs in patients <25 years (OR 1.16, $p=0.04$)- were associated with buttock pain. AxSpA spinal lesions were not associated with pain.

Modic type 1 lesions in patients >35 years (OR 5.19, $p=0.001$), high intensity zone lesions in females not fulfilling ASAS-criteria (OR 5.09, $p=0.001$) and herniation in various subgroups (OR ranged 2.07–4.66) were associated with pain.

Conclusions

Specific degenerative lesions -but not typical axSpA lesions- of the spine are associated with pain at the same location in given subgroups. Inflammatory lesions in the SIJ are associated with buttock pain.

Introduction

Axial spondyloarthritis (axSpA) is a rheumatic disease characterized by inflammation of joints and entheses with predominantly spinal involvement¹. Sacroiliitis on radiographs or magnetic resonance imaging (MRI) plays a role in the diagnostic process and as one of the entry criteria of the Assessment of SpondyloArthritis international Society (ASAS) classification criteria for axSpA². Besides inflammation, MRI can also visualize structural lesions considered typical for axSpA like fatty lesions, erosions and excessive bone formation³. Although the MRI of the spine (MRI-spine) does not contribute to the ASAS criteria, it may visualize lesions rather typical for axSpA. However, MRI-spine can also depict degenerative changes, either occurring in isolation or in combination with lesions considered typical for axSpA. These degenerative changes are common in both symptomatic and asymptomatic populations⁴.

Frequently patients present to the rheumatologist with back pain complaints of unknown origin. An important aim of the clinician is to find an explanation for the symptoms and to make a diagnosis. It is often hypothesized that MRI lesions due to axSpA will cause pain. In addition, MRI-spine is often included in the (diagnostic) imaging protocol to exclude other pathologies like degeneration, which may also explain pain reported by patients.

For clinicians, it is relevant to know whether there is a true relationship between MRI lesions, either due to axSpA or to degenerative changes, and pain localized at the same site. Two previous studies investigated the possible association between the location of MRI lesions and site of pain^{5,6}. However none of these studies investigated this association in an early chronic back pain (CBP) cohort nor combined the data of axSpA associated lesions and degenerative lesions.

Our objective was twofold: a) to evaluate if the presence of axSpA associated lesions seen on MRI of the sacroiliac joints (MRI-SI) is related to presence of buttock pain and b) if the presence of MRI-spine lesions typical for axSpA or for degenerative disease is related to CBP localized at the same site. This was executed in the SPondyloArthritis Caught Early (SPACE) cohort, comprising young patients with chronic back pain of relatively short duration of unknown origin, part of whom fulfil the ASAS classification criteria for axSpA.

Methods

Patients with CBP (≥ 3 months, ≤ 2 years, onset < 45 years) of unknown origin were included in the SPACE cohort. Five participating centers in the Netherlands (Leiden,

Amsterdam and Gouda), Norway (Oslo) and Italy (Padova) recruited patients of ≥ 16 years. At intake, all patients were asked to localize the sites of pain. Answering options were: thoracic and/or lumbar spinal segment, and/or buttock region. Patients were classified according to the ASAS axSpA classification criteria ('ASAS+' or 'ASAS-')². The large majority of patients in the ASAS- group were labelled as having 'nonspecific back pain'. Recruitment details and the clinical assessment of the study population have been published previously⁷.

MRI assessment

MRI-SI and MRI-spine was performed on a 1.5T scanner with the acquired sequences T1-weighted and short tau inversion recovery (STIR) in coronal oblique (MRI-SI) and sagittal view (MRI-spine). Two trained readers scored axSpA lesions (MdH and PB) in the SI joints and spine and two other readers scored degenerative changes (FdB and LdB) in the spine, with both sequences viewed simultaneously. Readers were blinded for patient characteristics, clinical outcome and the scores of the other readers. When specific conditions arose adjudication scores were obtained (explained below).

AxSpA lesions

On MRI-SI, inflammatory lesions were scored according to the SpondyloArthritis Research Consortium of Canada (SPARCC) method. In this method, the presence of inflammation is scored as present/absent in each quadrant of the SI joint in 6 consecutive slices, which represent the larger part of the synovial compartment of the SI joints. In addition, scores for lesion intensity (brightness equal to spinal fluid or veins) and depth (homogeneous and unequivocal signal increase extending >1 cm for articular surface) were provided per quadrant⁸. With an adapted version of the SPARCC method structural MRI-SI lesions (fatty lesions, erosions, sclerosis and ankylosis) were scored according to the same procedure as inflammatory lesions⁹. Corner inflammatory and structural MRI-spine lesions (fatty lesions, erosions (vertebral corner erosions) and syndesmophytes) were scored in 23 vertebral units (VU) according to the Canada-Denmark score^{10,11}. Inflammatory and fatty lesions were scored when present on ≥ 2 consecutive slices. Erosions and syndesmophytes were scored when visible on ≥ 1 slice.

In addition, readers gave a general dichotomous verdict about the MRI-SI (positive/negative), based on the following ASAS definition: ≥ 2 inflammatory lesions highly suggestive of axSpA on one slice or one lesion seen on ≥ 2 consecutive slices¹². A dichotomous score of positive MRI-spine was based on the following ASAS definition: ≥ 3 corner inflammatory lesions each seen on ≥ 2 consecutive slices¹³. An

adjudicator (RvdB) was introduced when readers disagreed on a positive MRI-SI or positive MRI-spine. If the primary readers agreed on a positive (or negative) MRI-SI, the mean SPARCC scores were calculated based on the scores of these primary readers. In case of disagreement, the mean scores were based on the consensus scores of the adjudicator and one primary reader. A similar process was followed for calculating the mean Canada-Denmark scores in the MRI-spine.

Degenerative lesions

We scored each VU for: disc degeneration using the Pfirrmann classification. This is a morphologic grading system, based on various stages of pathologic changes in the disc. Grade 1 is considered normal and grade 2 is 'any visual deviation of normal without signs of actual degeneration'. Therefore, a cut-off of >2 was used (as is customary in many other publications)^{14,15}. High intensity zone (HIZ): protrusion of the nucleus pulposus in the posterior element of the discus, resulting in a higher signal on STIR in the posterior part of the discus on MR imaging, hence the name^{16,17}. Herniation: projection of disc material outside the vertebral contour¹⁷. Schmorls' nodes: these are projection of disc material into the vertebral endplate, resulting in an indentation of the vertebral endplate^{18–20}. Modic changes: Type I are reactive end plate changes with inflammation (seen as high signal on STIR images), type II are reactive endplate changes with fat deposition (resulting in high signal on T1 images)^{21,22}. In addition, we defined degenerative disc disease (DDD) as a degenerative complex with presence of disc degeneration and an HIZ and herniation together in an intervertebral disc^{17,18}. An adjudicator (MR) was introduced in case of disagreement between the readers for presence/absence of Modic changes, herniation, HIZ and Schmorls' nodes. For Pfirrmann classification, adjudication was performed if the difference between the 2 readers was 2 classes or more (if there was only a 1 class difference, the lowest class was used).

Statistical analysis

Separate analyses were performed for lesions found on MRI-SI and MRI-spine. The association between MRI-SI lesions and buttock pain was studied with a logistic regression model. First a univariable analysis was performed, followed by multivariable to adjust for potential confounders. Scores of lesions in the SI joints varied widely between patients and therefore continuous scores were used. For MRI-spine, a multilevel model with generalized estimating equations (GEE) with an exchangeable correlation structure was used to investigate the association between the presence of each MRI-spine lesion separately and the occurrence of pain at the same

site, taking both levels of the spine into account²³. This analysis adjusts for spurious associations within patients presenting with lesions at different sites and provides the opportunity to interpret the effect of a lesion on pain as an independent effect of other investigated lesions. Since the majority of patients had a low score for spine lesions and only a few patients had a high score, dichotomous scores (presence/absence) were used. Interactions were tested to explore whether fulfilment of the ASAS axSpA-classification criteria, Human-Leukocyte-Antigen-B27 (HLA-B27) positivity, gender or age were possible effect modifiers. An interaction term with $p < 0.10$ was *a priori* considered relevant. In case of a relevant interaction data were presented per subgroup (eg. for age, HLA-B27, gender). For age, strata based on tertiles were explored.

All analyses were performed using Stata 13 and p-values < 0.05 were considered statistically significant.

Results

There were 348 patients with available baseline MRI-spine and clinical data, and 342/348 patients also had available MRI-SI. Patients' mean age was 29.4 (SD 8.2) years, 36.2% (n=126) were men and 36.5% (n=127) fulfilled the ASAS classification criteria for axSpA. Of the 348 patients, 37.9% (n=132) were HLA-B27 positive; of the axSpA patients 87.4% (n=111) were HLA-B27 positive. All patients had back pain since this was an inclusion criterion for the study. Patients could localize their pain at multiple levels and 35.9% (n=125) of the patients indicated pain at the thoracic site, 82.5% (n=287) at the lumbar site and 57.8% (n=201) at the buttock site.

Interaction analyses showed that in the SI joints, an interaction with $p < 0.10$ was found with age between buttock pain and sclerosis ($\beta = -0.06$; $p = 0.059$) and erosions ($\beta = -0.02$; $p = 0.014$). In the spine an interaction was found for age between pain and Modic type 1 lesions ($\beta = 0.14$; $p = 0.025$), and DDD ($\beta = 0.07$; $p = 0.029$), for sex between pain and HIZ ($\beta = 0.90$; $p = 0.05$), for HLA-B27 status: HIZ ($\beta = -0.87$; $p = 0.053$), herniation ($\beta = -0.84$; $p = 0.016$) and disc degeneration ($\beta = 0.25$; $p = 0.096$); for ASAS status: HIZ ($\beta = -0.87$; $p = 0.055$) and herniation ($\beta = -0.65$; $p = 0.056$).

AxSpA lesions

Table 1 showed that inflammatory lesions (20.8%) and erosions (29.2%) are found on MRI-SI in $> 20\%$ of the patients, and fatty lesions on MRI-SI in 14.6% of the patients. Sclerosis and ankylosis on MRI-SI and axSpA lesions on MRI-spine were not very

prevalent, occurring in 5-23 patients (1.4%-6.6%). Lesions were equally distributed among the thoracic and lumbar spine. In patients with lesions, inflammatory lesions in the SI joints had the widest range per patient; 0.5–53 (table 2). In general, a wider range of lesions was seen on MRI-SI than on MRI-spine. In patients with axSpA spinal lesions, the number of lesions and range was higher in the thoracic than in the lumbar spine.

For MRI-SI lesions, a significant association with buttock pain was found for inflammatory lesions (OR 1.06, $p=0.038$). Because of a significant interaction with age the data on erosions were provided for subgroups of age (<25 years, between 25 and 34 years, ≥ 35 years). Only in patients <25 a significant association was present (OR 1.16, $p=0.04$) (table 3). While the interaction analysis reported age as a relevant effect modifier in the relation between sclerosis and pain, the subgroup associations were not statistically significant. There was no association between ankylosis and pain. Table 4 showed that none of the spinal axSpA lesions were significantly associated with pain.

Table 1. Number of patients with specific MRI lesions, specified by location

	Thoracic n = 348 (%)	Lumbar n = 348 (%)	SI joints n = 342 (%)
<u>AxSpA lesions</u>			
Inflammatory lesions	43 (12.4)	30 (8.6)	68 (19.9)
Sclerosis	NA	NA	20 (5.8)
Fatty lesions	30 (8.6)	27 (7.8)	51 (14.9)
Erosions	20 (5.7)	18 (5.2)	101 (29.5)
Syndesmophytes (spine)/ankylosis (SI joints)	16 (4.6)	2 (0.6)	7 (2.0)
	Thoracic n = 348 (%)	Lumbar n = 348 (%)	SI joints n = 342 (%)
<u>Degenerative lesions</u>			
Modic type 1 (BME)	2 (0.6)	37 (10.6)	NA
Degenerative disc disease	4 (1.1)	40 (11.4)	NA
High intensity zone	0	81 (23.3)	NA
Schmorls' nodes	125 (35.9)	82 (23.6)	NA
Modic type 2 (fat)	1 (0.3)	11 (3.2)	NA
Herniation	6 (1.7)	118 (33.9)	NA
Disc degeneration (Pfirrmann >2)	75 (2.2)	156 (44.8)	NA

NA; not applicable

Degenerative lesions

Degenerative lesions were not equally distributed among the thoracic and lumbar spine. Table 1 showed that Modic changes, DDD, HIZ, herniation and disc degeneration were more often seen in the lumbar than in the thoracic spine. Table 2 showed that in patients with at least one degenerative lesion, only a few patients had >1 lesions (median was 1) in both the thoracic and the lumbar spine. As expected, Schmorls' nodes (median=2, range 1-10) and disc degeneration (median=2, range 1-10) were more frequent in the thoracic spine and DDD (median=2, range 1-6) was more frequent in the lumbar spine (table 2). HIZ was absent in the thoracic spine.

For degenerative lesions, significant associations with pain were found for various lesions in specific subgroups. Splitting the population in age-tertiles yielded strongest associations for DDD and Modic type 1 lesions with pain in the upper tertile (patients aged ≥ 35 years), as shown in table 4. In stratified analyses, an association of HIZ with spinal pain was found in HLA-B27 negative (OR=2.88, $p=0.001$), ASAS- (OR=3.43, $p<0.001$) and female patients (OR=3.92, $p<0.001$). In addition, we further analysed these subgroups by stratifying on both ASAS status and gender and found the strongest association in ASAS- female patients (OR=5.09, $p=0.001$). Herniation was associated with pain in HLA-B27 negative and ASAS+ and ASAS- patients, but in the latter with a stronger relationship than in ASAS+ patients (OR range 1.45-4.66).

To analyse disc degeneration and pain, patients were divided in an HLA-B27 positive and negative group. However, none of the subgroups showed significant associations between MRI-lesion and pain (table 4). When analysing the complete group ($n=348$) the association between DDD and pain was statistically significant (OR=1.46; CI 1.07-1.96 ($p=0.015$)).

We analysed the overlap between the degenerative reading (the adjudicated scores) and the axSpA readings of the individual readers to ensure that the same lesion is not interpreted as degenerative and as axSpA. We found that in 7 VU (6 patients) inflammatory axSpA lesions and Modic type 1 lesions co-exist, based on the adjudicated scores of degeneration and axSpA reader 1 (6 VU in 4 patients for reader 2). No overlap has been found between the degenerative reading and the axSpA readings for co-existence between Modic type 2 and fatty lesions.

Table 2. Number of specific MRI lesions in patients with at least 1 MRI lesion, specified by location (notation: [p0, p25] median [p75, p100])

	Thoracic	Lumbar	SI joint
AxSpA lesions			
Inflammatory lesions	[0.5, 0.5] 1.25 [2, 15]	[0.5, 0.5] 1 [1, 3.5]	[0.5, 1] 3 [11, 53]
Sclerosis	NA	NA	[0.5, 1] 2.75 [5.5, 15]
Fatty lesions	[0.5, 1] 3 [4.5, 12.5]	[0.5, 1] 1 [2.5, 6.5]	[0.5, 1] 4 [8.75, 27.5]
Erosions	[0.5, 0.5] 0.5 [0.5, 2]	[0.5, 0.5] 0.5 [0.5, 1.5]	[0.5, 1] 3.5 [6.5, 18.5]
Syndesmophytes (spine) /ankylosis (SI joints)	[0.5, 0.5] 0.5 [1, 1.5]	[0.5, 0.5] 0.5 [0.5, 0.5]	[1, 1.5] 3 [8, 11]
Degenerative lesions			
Modic type 1 (BME)	[1, 1] 1 [1, 2]	[1, 1] 1 [1, 1]	NA
Degenerative disc disease	[1, 1] 1 [1.5, 2]	[1, 1] 2 [2, 5]	NA
High intensity zone	0	[1, 1] 1 [1, 2]	NA
Schmorls' nodes	[1, 1.5] 2 [4, 10]	[1, 1] 1 [3, 5]	NA
Modic type 2 (fat)	[1, 1] 1 [1, 1]	[1, 1] 1 [1, 1]	NA
Herniation	[1, 1] 1 [1, 2]	[1, 1] 1 [2, 3]	NA
Disc degeneration (Pfirrmann >2)	[1, 1] 2 [3, 10]	[1, 1] 1 [2, 5]	NA
NA; not applicable			

Table 3. Relationship between MRI-SI lesions and buttock pain

AxSpA lesions (SI joints)	n	Stratification group $\pm$	OR	95% CI	P-value
Inflammatory SI lesions	342	No	1.06	1.00 – 1.13	0.04*
Sclerosis	117	<25 years $\S$	1.75	0.12 – 26.25	0.69
	117	25-34 years $\S$	0.96	0.77 – 1.20	0.73
	108	≥ 35 years $\S$	0.71	0.40 – 1.27	0.25
Fatty lesions	342	No	1.04	0.97 – 1.11	0.26
Erosions	117	<25 years $\S$	1.16	1.00 – 1.34	0.04*
	117	25-34 years $\S$	1.19	1.00 – 1.41	0.06
	108	≥ 35 years $\S$	0.94	0.78 – 1.12	0.47
Ankylosis	342	No	1.23	0.83 – 1.82	0.36

$\pm$ Where a statistically and clinically relevant interaction was found, results are shown stratified for the variable(s) indicated in this column. Where no stratification is shown, it means that there was no relevant interaction with age, sex, HLA-B27 status or fulfilment of the ASAS axSpA criteria

* P value < 0.05

$\S$ Analyses were stratified according to age tertiles

Table 4. Relationship between MRI spinal lesions and back pain - results from a multilevel analysis (levels: thoracic and lumbar spine)

	n	Stratification group [±]	OR	95% CI	P-value
AxSpA lesions					
Inflammatory spinal lesions	348	No	0.98	0.58 – 1.66	0.95
Fatty changes	348	No	0.80	0.47 – 1.36	0.40
Erosions	348	No	0.67	0.24 – 1.82	0.43
Syndesmophytes	348	No	1.36	0.24 – 7.83	0.73
Degenerative lesions					
Modic type 1 (BME)	117	<25 years [§]	0.92	0.09 – 9.05	0.94
	123	25-34 years [§]	1.31	0.39 – 4.39	0.66
	108	≥35 years [§]	5.19	1.90 – 14.19	0.001*
Degenerative disc disease	117	<25 years [§]	0.92	0.09 – 9.05	0.94
	123	25-34 years [§]	0.95	0.32 – 2.83	0.93
	108	≥35 years [§]	3.70	1.58 – 8.68	0.003*
High intensity zone	132	HLA-B27+	1.25	0.63 – 2.46	0.52
	216	HLA-B27-	2.88	1.58 – 5.23	0.001*
	127	ASAS+	1.15	0.48 – 2.75	0.75
	221	ASAS-	3.43	1.73 – 6.80	0.000*
	126	Male	1.39	0.65 – 2.99	0.39
	222	Female	3.92	1.86 – 8.24	0.000*
	62	ASAS+ & male	0.65	0.18 – 2.38	0.51
	64	ASAS- & male	1.69	0.62 – 4.64	0.31
	65	ASAS+ & female	1.84	0.53 – 6.34	0.34
	157	ASAS- & female	5.09	1.99 – 13.04	0.001*
Schmorls' nodes	348	No	1.03	0.77 – 1.39	0.81
Modic type 2 (fat)	348	No	2.41	0.66 – 8.77	0.18
Herniation	132	HLA-B27+	1.45	0.93 – 2.26	0.10
	216	HLA-B27-	3.14	1.91 – 5.17	0.000*
	127	ASAS+	2.07	1.03 – 4.15	0.041*
	221	ASAS-	4.66	2.24 – 8.97	0.000*
Disc degeneration (Pfirrmann >2)	132	HLA-B27+	1.25	0.97 – 1.62	0.08
	216	HLA-B27-	0.97	0.86 – 1.09	0.59

[±] Where a statistically and clinically relevant interaction was found, results are shown stratified for the variable(s) indicated in this column. Where no stratification is shown, it means that there was no relevant interaction with age, sex, HLA-B27 status or fulfilment of the ASAS axSpA criteria

* P value < 0.05

[§] Analyses were stratified according to age tertiles

Discussion

In this study the relationship between site of lesions on MRI-SI or MRI-spine and the localization of pain by patients was investigated in patients presenting with chronic back pain between 16 and 45 years of short duration, with 36.5% patients fulfilling the ASAS axSpA criteria. We have found in the SI joints an association between inflammatory lesions and buttock pain as well as an association between erosions and buttock pain. To calculate the likelihood that a patient indicates buttock pain the OR

of 1.06 should be multiplied for every inflammatory lesion present. As the median number of lesions per patient is 3 (with a range up to 53) this may result in a clinically (very) relevant association (e.g. for 3 inflammatory lesions an 19% increased risk). In patients younger than 25 years, the increased risk is 16% for a single erosion and 81% for 4 erosions (median number of erosions 3.5) present on MRI-SI.

In addition, an association between several types of degenerative lesions but not with axSpA spinal lesions and pain was found in the spine.

AxSpA lesions were not very prevalent in the spine of CBP patients with a maximum of two years symptom duration (<10% of patients), but were far more prevalent in the SI joints. Degenerative lesions were more prevalent in the lumbar spine than in the thoracic spine, with the exception of Schmorls' nodes.

To our knowledge there is one other study in which the relation between the localization of pain and axSpA lesions on MRI was investigated and the findings were partly matching the findings in our study. Like in the current study, Blachier et al. found an association between buttock pain and inflammatory and structural MRI-SI lesions, but with higher ORs of 2.86 and 1.89, respectively⁵. Unfortunately the readers in that study did not differentiate between structural lesions, so it is unclear what type(s) of structural lesion(s) were associated with buttock pain. Using logistic regression analysis, Blachier et al. also found the thoracic and lumbar pain site to be related to inflammatory lesions at the same site. Logistic regression analyses takes the site (lumbar vs thoracic) rather than the patient as the dependent variable, does not adjust for within-patient correlation and consequently may yield spurious correlations. This is why we have used GEE which adjusts for within-patient correlation. Different analysis methods may explain lower ORs in our study as compared to Blachier's.

Moreover, a larger proportion of patients in the current study did not fulfil the ASAS criteria compared to Blachier et al. and thus, consequently subgroups of axSpA/AS associated interactions (ASAS status, HLA-B27 status and gender) were smaller. In addition, patients were enrolled in the SPACE cohort in an early stage of the disease, resulting in only few patients with lesions in the spine and in low numbers of lesions per patient who had at least one lesion. Moreover, the reading of the erosions and syndesmophytes was done by a 1-slice approach. This resulted in a difference in reading between the two readers: frequently a single erosion was scored by one reader but not by the other reader. This can be seen when comparing the low percentage of patients with erosive disease in table 1 and the average of the reader scores of 0.5 in at least 75% of the patients. The same applies to the syndesmophytes. The absence of a relation between erosions or syndesmophytes and pain is most likely due to the very

low number of lesions in patients early in disease, but we cannot rule out that a discrepancy in reading between the readers played a role.

Modic type 1 lesions, herniation, HIZ and DDD showed the strongest association with pain in various subgroups. For DDD and Modic type 1 lesions, this association was strongest in the group of patients ≥ 35 (upper tertile). Obviously, these lesions are more common in the 'older' patient group, but it should be noted that the maximum age did not exceed 45 years. Expectedly, associations may even be stronger in older patients than those investigated in this cohort.

For herniation and HIZ, associations with pain were found in the subgroups that are least associated with axSpA/AS: females, HLA-B27 negative patients and patients not fulfilling ASAS criteria. In contrast, disc degeneration seemed to be associated with pain in the entire group.

Many other studies have confirmed the association between degenerative lesions and back pain. Kjaer et al. investigated a randomly selected sample of 413 40-year-old individuals from the general population⁶. The associations between the presence of various degenerative lesions in the lumbar spine and reported back pain during the previous 3 months, reported back pain during the last year, as well as reported back pain for which medical care was sought, were analysed. The 3 types of back pain, defined by Kjaer, were statistically significantly associated with Modic changes (ORs ranging from 1.9–4.2) and disc degeneration (ORs ranging from 1.5–2.6). HIZ were associated with back pain in the last year and back pain for which care was sought (ORs were 2.0 and 2.5 respectively), as was herniation with ORs 1.9 and 2.2. In a meta-analysis including data of 3097 patients from 14 studies, aged 50 at most, an association was found between back pain and disc degeneration (OR 2.24), herniation (OR 2.65) and Modic type 1 (OR 4.01) but not with HIZ²⁴.

We are not aware of studies that have analysed a complex of degenerative lesions comparable to the DDD as was used in the current study. As far as we know, no single study has investigated the association between degenerative changes and pain in a cohort of young patients referred and included with a certain suspicion of axSpA.

Degenerative changes on MRI are rather prevalent in otherwise asymptomatic patients^{4,25}, which makes it difficult to assume a causal relationship. On the other hand, the rather strong positive associations that we have found between MRI-signs of degenerative spine disease and back pain in patients with characteristics least associated with axSpA, as well as the regional co-localization of MRI-signs and symptoms, make us feel confident to hypothesize that -at least in the patients in our

cohort, who experienced back pain almost daily for at least 3 months and sought help for that- the degenerative changes may indeed cause back pain at the same sites.

Our study has a few limitations. Due to the cross-sectional nature of our study, we were not able to analyse causality. Therefore, future analyses could focus on whether the lesions seen on MRI-SI and MRI-spine are 'site-stable' or that the site of MRI lesions may change over time with longitudinal analyses. It would be interesting to see whether the localization of pain will change when the site of MRI-lesions changes. Such a detailed analysis requires multiple well-trained readers, as was performed in this study, and also adjudication, in order to assure reliability of findings and a high level of internal validity. Needless to say that internal validity may go at the expense of generalizability: it will still be difficult to attribute localized pain in an individual patient to a site-specific lesion on MRI reported by a radiologist in usual clinical practice. Another possible limitation is the exclusion of the posterior elements. However, in an earlier analysis of this cohort, we found that inflammation in the posterior element was present in only 4% of the patients and this was always in patients who also have inflammation in the vertebral bodies²⁶. This may be different in patients with more longstanding disease.

In conclusion, inflammatory lesions and erosions in the SI joints that are typical for axSpA are associated with buttock pain and degenerative lesions on MRI of the spine are associated with back pain in corresponding regions, especially in the subgroups of patients least associated with axSpA. No association was found between pain and axSpA lesions in the spine.

References

1. Dougados M, Baeten D. Spondyloarthritis. *Lancet* 2011;377:2127-37.
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, Rudwaleit M, Baraliakos X, et al. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. *Ann Rheum Dis* 2009;68:ii1-ii44.
4. Weishaupt D, Zanetti M, Hodler J, et al. MR imaging of the lumbar spine: prevalence of intervertebral disk extrusion and sequestration, nerve root compression, end plate abnormalities, and osteoarthritis of the facet joints in asymptomatic volunteers. *Radiology* 1998;209:661-6.
5. 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.
6. Kjaer P, Leboeuf-Yde C, Korsholm L, et al. Magnetic resonance imaging and low back pain in adults: a diagnostic imaging study of 40-year-old men and women. *Spine* 2005;30:1173-80.
7. van den Berg R, de Hooze 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. *Rheumatol* 2013;52:1492-9.
8. Maksymowych WP, Inman RD, Salonen D, et al. Spondyloarthritis Research Consortium of Canada magnetic resonance imaging index for assessment of sacroiliac joint inflammation in ankylosing spondylitis. *Arthritis Care Res* 2005;53:703-9.
9. Weber U, Lambert RGW, Østergaard M, et al. The diagnostic utility of magnetic resonance imaging in spondylarthritis: An international multicenter evaluation of one hundred eighty-seven subjects. *Arthritis Rheum* 2010;62:3048-58.
10. Lambert RGW, Pedersen SJ, Maksymowych WP, et al. Active inflammatory lesions detected by magnetic resonance imaging in the spine of patients with spondyloarthritis - Definitions, assessment system, and reference image set. *J Rheumatol* 2009;suppl 84:3-17.
11. Østergaard M, Maksymowych WP, Pedersen SJ, et al. Structural lesions detected by magnetic resonance imaging in the spine of patients with spondyloarthritis - Definitions, assessment system, and reference image set. *J Rheumatol* 2009;suppl 84:18-34.
12. 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;68:1520-27.
13. Hermann K, Baraliakos X, van der Heijde D, et al. Descriptions of spinal MRI lesions and definition of a positive MRI of the spine in axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI study group. *Ann Rheum Dis* 2012;71:1278-88.
14. Pfirrmann CW, Metzendorf A, Zanetti M, et al. Magnetic resonance classification of lumbar intervertebral disc degeneration. *Spine* 2001;26:1873-8.
15. Pearce RH, Thompson JP, Beabout GM, et al. Magnetic resonance imaging reflects the chemical changes of aging degeneration in the human intervertebral disk. *J. Rheumatol Suppl* 1991;27:42-3.
16. Aprill C, Bogduk N. High-intensity zone: a diagnostic sign of painful lumbar disc on magnetic resonance imaging. *Br J Radiol* 1992;65:361-9.
17. Fardon DF, Milette PC. Nomenclature and classification of lumbar disc pathology. Recommendations of the Combined task Forces of the North American Spine Society, American Society of Spine Radiology, and American Society of Neuroradiology. *Spine* 2001;26:E93-E113.

18. Modic MT, Carter JR, Masaryk TJ, et al. Imaging of degenerative disk disease. *Radiology* 1988;168:177-86.
19. Modic MT, Steinberg PM, Ross JS, et al. Degenerative disk disease: assessment of changes in vertebral body marrow with MR imaging. *Radiology* 1988;166:193-9.
20. Emch TM, Modic MT. Imaging of lumbar degenerative disk disease: History and current state. *Skeletal Radiol* 2011;40:1175-89.
21. Pfirrmann CW, Resnick D. Schmorl nodes of the thoracic and lumbar spine: radiographic-pathologic study of prevalence, characterization, and correlation with degenerative changes of 1,650 spinal levels in 100 cadavers. *Radiology* 2001;219:368-74.
22. R. J. Moore. The vertebral endplate: Disc degeneration, disc regeneration. *Eur Spine J* 2006;15 Suppl 3:S333-7.
23. Twisk JWR. *Applied Multilevel Analysis*. 2006.
24. Brinjikji W, Diehn FE, Jarvik JG, et al. MRI findings of disc degeneration are more prevalent in adults with low back pain than in asymptomatic controls: A systematic review and meta-analysis. *Am J Neuroradiol* 2015;36:2394-9.
25. Boden SD, McCowin PR, Davis DO, et al. Abnormal magnetic-resonance scans of the cervical spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am* 1990;72:1178-84.
26. van den Berg R, de Hooge M, Reijnen M, et al. Inflammation in the posterior elements of the spine is infrequent in patients with recent onset axial spondyloarthritis, especially if no inflammation in the vertebral bodies is observed. *Ann Rheum Dis* 2013;72 Suppl:663-4.

9

Summary and conclusions

Summary and conclusions

The first part of this thesis covered studies that focus on the diagnosis and classification of (early) axial spondyloarthritis (axSpA) patients. More specifically, the performance of two modifications of the diagnostic Berlin algorithm and the performance of the Assessment of Spondyloarthritis international Society (ASAS) axSpA classification criteria were tested. In the second and major part of this thesis the focus shifted to imaging in axSpA. This part addressed several aspects of imaging by MRI. First, the use and additional value of the paramagnetic contrast agent gadolinium-DTPA in the assessment of sacroiliitis on MRI as part of the classification of axSpA patients. Thereafter, the discrimination between patients meeting and not meeting the ASAS axSpA criteria based on structural lesions seen on magnetic resonance imaging of the sacroiliac joints (MRI-SI) as well as lesions on MRI of the spine (MRI-spine) was tested. Also, the reliability of reading MRIs and radiographs of the spine for signs of axSpA was investigated. Finally, this thesis covered a study estimating the association between pain and MRI lesions due to axSpA or of degeneration.

The two cohorts in which all studies in this thesis were conducted are the DEvenir des Spondylarthropathies Indifférenciées Récentes (DESIR) and Spondyloarthritis Caught Early (SPACE) cohorts. Patients between 18 and 50 years, with inflammatory back pain (IBP) according to the Calin¹ or Berlin criteria² for ≥ 3 months but < 3 years were included in the DESIR cohort. In addition, all patients in the DESIR cohort had to have symptoms moderately to highly suggestive of SpA according to the local rheumatologist. Recruitment took place in 25 participating centres across France and the final patient was recruited on the 29th of April 2010.

The SPACE cohort is a multicentre but also multinational, inception cohort of patients with chronic back pain. The cohort was launched in January 2009 and recruitment of patients is ongoing. Patients from the Netherlands (Leiden, Gouda and Amsterdam), Norway (Oslo) and Italy (Padova) who are referred to a rheumatologist are included in the SPACE cohort when they are ≥ 16 years, have (almost) daily chronic back pain for ≥ 3 months but ≤ 2 years, with an onset of back pain < 45 years. In both cohorts clinical and imaging data was collected according to a fixed protocol of the corresponding cohort, which included several questionnaires, history taking, physical examination, and laboratory and imaging evaluations.

This chapter summarizes the most important findings per chapter and presents the possibilities for future research with respect to the studies included in this thesis.

Compatibility axSpA criteria sets

The most recently developed criteria to classify axSpA patients are the ASAS axSpA criteria, developed in 2009. The study in which these criteria were created showed a good sensitivity and specificity of the ASAS axSpA criteria, using the expert physician's diagnosis as gold standard³. The study was performed in a cohort containing patients with a high suspicion of SpA collected by experts in the field of SpA. Therefore the cohort had a high prevalence of axSpA patients. Whether the ASAS axSpA criteria would perform well in a clinical setting with possibly a lower pre-test probability was investigated in **chapter 2**. The performance of the ASAS axSpA criteria was compared to the performance of other classification criteria namely the Amor and European Spondyloarthropathy Study Group (ESSG) criteria in patients included in the SPACE cohort. The diagnosis of the treating rheumatologist was used as gold standard to test the performance of the ASAS axSpA classification, Amor and ESSG criteria. Almost 60% of the patients fulfilled the ASAS axSpA classification, Amor or ESSG criteria, and approximately 2 out of 5 patients (38.2%) fulfilled the ASAS axSpA classification criteria. Moreover, the ASAS axSpA criteria outperformed the other criteria with regard to sensitivity, specificity, positive and negative likelihood ratio, also after modifying the Amor and ESSG criteria by adding active sacroiliitis on MRI as a possible component.

Furthermore, we also investigated whether patients fulfilling the clinical arm had a similar disease as patients fulfilling the imaging arm of the ASAS axSpA criteria. Indeed, the patients in the SPACE cohort fulfilling the clinical arm were remarkably similar to those fulfilling the imaging arm. Patients showed the same level of disease activity and symptom duration as well as the same frequency of SpA features. The only SpA feature, which appeared more frequent in the clinical arm, was a positive family history for SpA or related diseases. This was not surprising, since a positive family history is associated with Human Leukocyte Antigen B27 (HLA-B27) positivity in a patient, and HLA-B27 is mandatory in the clinical arm⁴⁻⁶. However, the level of confidence of the rheumatologist about the diagnosis was lower in the clinical arm than in the imaging arm, reflecting a feeling of confidence induced by imaging findings and this dissimilarity indicates that the findings on MRI or radiographs strongly influence the judgement of the rheumatologist.

The presence of SpA features was compared between axSpA and no-SpA patients in **chapter 2**. It is interesting to see that though there was a significant difference in the presence of IBP in axSpA and no-SpA patients, approximately 17% of axSpA patients did not have IBP and nearly 57% of no-SpA patients had IBP. These findings confirm previous findings that even though IBP is a strong characteristic of axSpA, IBP

should not be considered pathognomonic ^{7,8}. Despite the fact that not all axSpA patients have IBP, IBP is used as entry criterion in the original Berlin Algorithm for diagnosing axSpA. This algorithm is a tool assisting clinicians in the diagnostic process of (early) axSpA by advising the clinician if and what tests should be performed next to reach an informed diagnosis. IBP as an entry criterion may be too stringent and consequently axSpA may not be considered in patients with chronic back pain without IBP. In **chapter 3** the performance of two modifications to the algorithm were tested in the SPACE and ASAS cohorts. The ASAS cohort was assembled for the validation of the ASAS classification criteria for axSpA and included patients with chronic back pain (CBP) for ≥ 3 months with an onset < 45 years and with a suspicion of SpA but without a definite diagnosis. Patients were evaluated according to a fixed protocol by rheumatologists with an expertise in the field of SpA. The diagnostic work-up contained assessment of clinical, laboratory and imaging data³. The two modifications of the Berlin algorithm that were tested in this chapter focussed both on IBP. In the Berlin algorithm the fulfilment of the the ASAS IBP criteria (≥ 4 out of 5 criteria) is the first mandatory step; One modification loosened these IBP criteria: patients who fulfilled ≥ 3 out of 5 criteria were considered having IBP. In the other modification IBP was deleted as obligatory entry criterion and added as SpA-feature. When investigating the performance of a diagnostic algorithm the lack of a genuine gold standard is a limitation. To diminish this limitation several external standards were used to test the performance of the modified algorithms. Besides the diagnosis of the rheumatologist and fulfilment of the ASAS criteria also the likelihood ratio product, which represents the disease probability, was used as an external standard. The results in this chapter showed that both modifications greatly increase the sensitivity of the algorithm, while specificity decreases only slightly. The best balance between sensitivity and specificity was reached with the latter modification: deleting IBP as entry criterion and including it as SpA-feature.

Imaging

MRI sequences recommended to detect inflammatory lesions are Short Tau Inversion Recovery (STIR) and T1-weighted Spin Echo fat-saturated after intravenous administration of gadolinium-DTPA (T1/Gd), used in concordance with T1-weighted Spin Echo sequence⁹. However, intravenous contrast administration has drawbacks: patients experience discomfort, there are potential side effects, it is more costly, and administration is time consuming. Therefore we have conducted a study to determine the additional value of intravenous administration of gadolinium (T1/Gd) in detecting inflammatory lesions compared to the STIR sequence. This was studied at baseline

and follow-up after 3 months in patients of the SPACE cohort. We also assessed the possible influence of using T1/Gd instead of STIR sequence on the final definition of a positive MRI-SI based on the ASAS definition of active sacroiliitis¹⁰. **Chapter 4** showed that there is 100% agreement between STIR and T1/Gd sequences on the presence of bone marrow edema (BME) at baseline and follow-up. This implicates that for the detection of BME lesions, the administration of gadolinium does not have additional value. This was in line with results found in a previous study in patients with axSpA showing nearly equal performance between both sequences in detecting BME in the periphery of chronic changes¹¹. In our study we found that T1/Gd was more sensitive in detecting capsulitis and enthesitis since this was not seen on the STIR sequence. However, capsulitis and enthesitis were always present alongside BME lesions and synovitis. Synovitis is only detectable with the T1/Gd sequence, and all but one of eight patients showed synovitis in the presence of BME on the T1/Gd. The single patient showing synovitis without BME did not have any characteristic associated with axSpA and therefore the isolated synovitis found in this patient was considered a false positive finding. This is in agreement with the recommendations by ASAS that the sole presence of synovitis is insufficient to label an MRI-SI as positive⁹. As a direct result of these findings, the administration of gadolinium for patients included in the SPACE cohort was stopped at the beginning of April 2012, leaving only the assessment of the STIR and T1-weighted MRI sequences in this cohort.

The ASAS group that developed the ASAS axSpA criteria refrained from including lesions other than BME MRI-SI lesions due to lack of evidence on the utility of structural MRI-SI and MRI-spine lesions in the classification of (early) axSpA³. Before it is possible to answer the question if the definition of ‘a positive MRI’ used in the ASAS criteria should be modified by adding structural MRI-SI lesions and/or spinal lesions seen on MRI, one needs to know if there are particular MRI lesions that are typically seen in axSpA and to what extent. There have been several studies on MRI-SI lesions in SpA patients, but due to the absence of a control group it was impossible to assess the specificity of these lesions^{12–15}. The study reported in **chapter 5** did include a control group of CBP patients due to other causes than axSpA from the SPACE cohort. Because of this control group we were able to quantify MRI lesions in patients with and without axSpA. Moreover, we investigated which lesions had a specificity of >95% for CBP patients meeting the ASAS axSpA criteria. On purpose we selected a high specificity to reduce misclassification of patients as the imaging arm is fulfilled in case of positive imaging plus only one additional SpA feature. The studies that investigated the quantity of structural MRI-SI lesions solely in axSpA

patients corresponded to the results we found in our study; fatty lesions and erosions on MRI-SI were considered specific for axSpA with optimal cut-offs of '≥3 fatty lesions' and '≥3 erosions'. In one study evidence was found that erosions are more relevant MRI-SI lesions than fatty lesions alone as the majority of axSpA patients have fatty lesions alongside inflammation or erosions¹⁶. Therefore we also investigated the specificity of combinations of MRI-SI lesions and we found the combination of '≥5 fatty lesions and/or erosions' to be the only one assuring >95% specificity.

Also the quantity and specificity of lesions on MRI-spine was investigated in this chapter and results showed that '≥5 inflammatory lesions' and '≥5 fatty lesions' had specificity of more than 95%. The cut-off of '≥5 inflammatory lesions' is higher than the proposed ASAS definition of 'a positive MRI-spine' (≥3 corner inflammatory lesions suggestive of SpA)¹⁷. However, the results of previous studies approximated our data more than the ASAS definition of 'a positive MRI-spine'. These studies found that ≥6 inflammatory lesions and ≥6 fatty lesions corresponded to a moderate to substantial diagnostic utility in axSpA patients^{18,19}. In addition, and in agreement with our study, Weber et al. found that combining inflammatory and fatty lesions of the spine does not increase specificity¹⁸.

Comparing the data from **chapter 5** to the literature, even with different readers, different cohorts and diverse choices of external standard, we arrived at about the same cut-off values for structural lesions on MRI-SI and lesions on MRI-spine. This resulted in the recommendation of using 'a rule of five' when looking at MRI lesions representative of axSpA: '≥five fatty lesions and/or erosions' on MRI-SI, as well as '≥five spinal inflammatory lesions' and '≥five spinal fatty lesions' are highly specific for axSpA, and thereby this practical rule provides a useful level of discrimination between axSpA and no-SpA patients.

In clinical trials and cohort studies like the ones described above, the imaging scores are always provided by trained readers. Often the trained readers have participated in one or several calibration sessions. Regularly more than 1 reader provides the imaging scores used in clinical studies. This is dissimilar to daily practice where a local radiologist alone, or in consensus with a rheumatologist, performs the imaging readings. The specialist reading the images in daily practice is aware of the clinical data while the readers performing scores for study purposes are usually blinded for clinical data. In both settings readers are assumed to assess the compatibility of imaging abnormalities with axSpA. In earlier work, moderate to substantial levels of agreement between local readers (LocR) and central readers (CentR) were found regarding SI joint assessment on radiographs and MRI, respectively^{20,21}. LocR classified a large

proportion of the patients as ankylosing spondylitis, in disagreement with CentR. While in just a small amount of patients the classification ankylosing spondylitis was assigned by CentR but not by the LocR. Agreement of SI joint assessment on MRI was considerably better and was evenly distributed between over- and underclassification. So, differences between LocR and CentR considering radiographs and to a lesser extent MRI, would lead to misclassification in some patients (ankylosing spondylitis meeting the modified New York (mNY) criteria or axSpA according to the ASAS axSpA criteria)^{20,21}. These studies were performed in the DESIR cohort. Thus far the DESIR cohort is unique in having reported data of both local and central assessments of radiograph and MRI readings of the SI joints as well as the spine. Like the agreement on abnormalities in the SI joints on radiographs, the agreement between LocR and CentR on lesions in the spine is low (**chapter 6** and **chapter 7**). In **chapter 6** the radiographic changes in the spine due to axSpA were quantified using the modified Stoke Ankylosing Spondylitis Spine Score (mSASSS) by evaluating the anterior vertebral corners of the cervical and lumbar part of the spine on the presence of erosions, squaring, sclerosis and bridging or non-bridging syndesmophytes. The agreement between the two central readers was moderate. Still this agreement was higher than the fair agreement we found between LocR and CentR. In addition, the disagreement between the two central readers was balanced, with neither reader systematically overestimating the mSASSS relative to the other reader. In the sensitivity analysis, using a higher cut-off to define radiological damage, agreement between the two central readers improved. However, the agreement between LocR and CentR remained only fair. Because of their specific training we used the central readers as external standard and therefore disagreements between LocR and CentR were attributed to the performance of the local readers. Since LocR more frequently gave an mSASSS of ≥ 1 when the CentR did not score any radiographic abnormalities, we concluded that disagreement between LocR and CentR consistently occurred as overestimation of radiographic spinal signs by the local readers. This phenomenon is similar to the results regarding the agreement on radiographic abnormalities of the SI joints²⁰.

The agreement between the central readers who scored lesions on MRI-spine was also found to be higher than the agreement between LocR and CentR. For inflammatory as well as structural lesions on MRI-spine the agreement between central readers was moderate. The agreement between LocR and CentR was found to be fair for inflammatory lesions and just slight for structural lesions seen on MRI-spine (**chapter 7**). Due to a discrepancy in the scoring methods between LocR and CentR, the agreement on MRI-spine lesions could only be examined on the level of a

dichotomous score -presence/absence- for inflammatory and structural abnormalities suggestive of axSpA. The CentR conducted a detailed and quantified score while the LocR provided a global assessment of 'presence', 'doubt' or 'absence' of lesions. These differences in scoring options was unfortunate and made interpretation of the data challenging. In **chapter 7** it seemed that LocR underrated MRI-spine lesions. CentR more often scored inflammatory lesions as well as structural lesions. In the vast majority of these positive cases LocR did not decide on the presence or absence of lesions and used the possibility to express their doubt about the presence of MRI-spine lesions due to axSpA. The overall agreement and percentage positive agreement between LocR and CentR improved when 'doubt' scores were allocated to the 'lesions present' group and decreased when considered as 'lesions absent'. LocR appeared to only score MRI-spine lesions when they were very confident of the presence of a lesion. Judging from the numerous times the option 'doubt' was chosen the local readers were often not confident enough to make a decision. Although we can only guess, the explanation for this perhaps lies in the reduced confidence of local readers on the typical appearance of lesions in the spine. They may be uncertain about the cause and find it challenging to distinguish lesions due to axSpA from lesions due to other causes like degenerative disc disease. This may be possible since it is considered difficult to discriminate between SpA and degenerative lesions²². Moreover, it is suggested that typical axSpA and degenerative lesions may possibly coexist in one patient. In another study we found that trained readers could reliably differentiate between lesions due to axSpA and lesions due to degeneration with little overlap between axSpA and degenerative lesions (F. Bruijn et al. submitted for publication). Two of the readers in that study were the same as the central readers in **chapter 7**. All readers in this unpublished study were specifically educated on the differences between degenerative and axSpA findings on MRI by experienced musculoskeletal radiologists. This suggests that specific training on the characteristics of degenerative and axSpA lesions improves assessment of spinal lesions. Disagreement between LocR and CentR on lesions in the spine, both on radiographs and MRI, may indicate challenges for local non-trained readers in distinguishing degenerative lesions from axSpA lesions in the spine.

Frequently patients present to the rheumatologist with back pain complaints of unknown origin. In the process of finding an explanation for the symptoms and to make a diagnosis, an MRI of the spine may be included in the (diagnostic) imaging protocol. For clinicians, it could be relevant to know whether there is a relation between MRI findings, either due to axSpA or to degenerative causes, and pain at the

same site. In **chapter 8** the associations between pain and MRI lesions at the same site were investigated in the SPACE cohort. The results showed associations between buttock pain and both inflammatory lesions and erosions on MRI-SI. The likelihood that a patient had buttock pain increased with 6% for every additional inflammatory lesion seen on MRI-SI. In patients who are younger than 25 years at the onset of their back pain complaints, the likelihood of having buttock pain even increased with 16% for every additional erosion on MRI-SI. In agreement with our findings Blachier et al found an association between buttock pain and inflammatory and structural MRI-SI lesions in the DESIR cohort²³. Unfortunately no differentiation between structural lesions was made in that study, so it was unclear what type(s) of structural lesion(s) were related to buttock pain.

Whereas associations with axSpA MRI-SI lesions were seen in young adults in our study, associations between pain and degenerative lesions were seen in relatively older patients although it has to be noted that due to the inclusion criteria of the SPACE cohort the maximum age at inclusion did not exceed 47 years. For degenerative disc disease and Modic type 1 lesions, the association with pain at the same location was the strongest in the group of patients with the onset of back pain after the age of 35. Pain was associated with herniation and high intensity zone -which was determined as protrusion of the nucleus pulposus in the posterior element of the discus- in the subgroups of patients least associated with axSpA: female patients, HLA-B27 negative patients and patients that did not fulfil the ASAS axSpA classification criteria. When stratifying for HLA-B27 none of the subgroups showed significant associations between disc degeneration and pain. However, disc degeneration seemed to be associated with pain in the entire group. These findings confirmed results found in prior studies^{24,25}. In contrast to degenerative spinal lesions and MRI-SI lesions due to SpA, no associations were found between MRI-spine lesions due to SpA and pain at the same site.

Future perspective for research

The modification of the Berlin algorithm suggested in **chapter 3** was validated in two independent cohorts and this modification showed improvement over the original algorithm by reducing the patients that falsely exit the algorithm. Based on these results, the ASAS group accepted this modified Berlin algorithm. Since the algorithm is presented as a tool to help rheumatologists in the early diagnosis of axSpA, the use of the algorithm by rheumatologists could be investigated to assess performance of the current algorithm with the goal of continuous improvement of the algorithm.

Besides the overall performance of the algorithm, the cost-effectiveness of imaging in

the algorithm could be investigated. Initially, the algorithm was developed with the consideration to keep the use of MRI-SI to a minimum. It would however be interesting to investigate if the benefits of further testing on the presence of abnormalities on MRI-SI in some patients would outweigh the costs of performing an MRI-SI. For example in HLA-B27 negative patients with 2-3 SpA features the algorithm does not recommend making an MRI and suggest another diagnosis than axSpA but we have shown that in the SPACE cohort where an MRI is made in all patients, a considerable percentage of these patients have axSpA according to the rheumatologist (**chapter 3**). In these patients it may be worthwhile to perform an MRI-SI. Furthermore, the cost-effectiveness of radiography of the SI joints in the Berlin algorithm could be investigated. Even though performing radiographs is less expensive compared to MRI, it is questionable whether radiographs of the SI joints should be performed in all patients visiting the rheumatologist with back pain of unknown origin like the algorithm suggests. In particular, because in axSpA with a short symptom duration only few patients (already) have abnormalities on radiography. Elaborating on this, it is interesting to review if the rheumatologists approve the order of the algorithm. The algorithm starts with the radiographs of the SI joints while a consultation with a rheumatologist starts with questions about the clinical history, then the physical examination and if necessary additional tests like laboratory and imaging research. When the algorithm is used retrospectively, when all test results are available, the current order is useful but when the rheumatologist would use the algorithm as a guiding tool during the consultation it might not be convenient in practice to perform a radiograph of the SI joints even before history taking. As physicians are supposed to use the algorithm as a tool in the daily practice, it would make sense that the algorithm follows the order of the typical clinical consult (history taking and physical examination followed by laboratory and imaging tests).

With the introduction of MRI-SI to the ASAS axSpA classification criteria, the discussion on the definition of a positive MRI -as defined in the classification criteria- intensified. Patients not fulfilling the mNY criteria and without active sacroiliitis but with structural damage on MRI-SI suggestive of axSpA are not classified as axSpA by the current form of the ASAS criteria when they are HLA-B27 negative. An interesting question is whether the physician would diagnose such patients as having axSpA. In keeping with our results in **chapter 5**, a useful next step is to test the influence of adding ‘ ≥ 3 fatty lesions’ as a criterion to the definition of a positive MRI-SI to the performance of the ASAS criteria, with the rheumatologist’s diagnosis as gold standard. Parallel to this investigation, similar analyses with the criteria ‘ ≥ 3

erosions' and '≥5 fatty lesions and/or erosions' should be performed. Also, spinal MRI lesions are currently not part of any classification criteria sets for axSpA. So, when considering altering the definition of a positive MRI, one should also reflect on the possible additional value of axSpA lesions seen on MRI-spine. It could be of interest to investigate if adding '≥5 inflammatory spinal MRI lesions' and '≥5 fatty spinal MRI lesions' increases the performance of the ASAS axSpA criteria. In this light, the definition of a positive MRI of the spine given by the ASAS/OMERACT group -≥3 inflammatory lesions suggestive of SpA- should also be tested.

To further investigate the utility of MRI-spine an intriguing additional question is if - and to what extent- MRI-spine data would influence the diagnosis of the rheumatologist. Ideally, the rheumatologist would first be blinded for the MRI-spine when giving the diagnosis. The MRI-spine could be provided after the rheumatologist has given the diagnosis based on the clinical data and imaging data of the SI joints only. Subsequently, the rheumatologist would give a diagnosis again with the MRI-spine data at his/her disposal. Should MRI of the spine be found useful for diagnosing axSpA, the added value should nevertheless be weighed against the relatively high costs of MRI spine.

Furthermore, one could wonder if axSpA should be considered in patients with radiographic lesions of the spine in absence of radiographic sacroiliitis. Conceivably, the presence of either radiographic sacroiliitis or spinal bone formation might be a better criterion to define radiographic axSpA compared to the sole presence of radiographic sacroiliitis. This could be an intriguing focus for future research.

When considering the incorporation of spinal lesions in the ASAS criteria it should be taken into account that the reliability between readers appears to be moderate at most. At the moment, disease classification is not influenced by disagreement on spinal lesions, since spinal lesions are not part of the current classification criteria for AS or axSpA. But if the time arrives where they will be part of criteria sets or when spinal lesions are part of the eligibility for a clinical study, then reliably determining the presence of inflammatory/structural spinal lesions becomes at least as relevant as determining whether the patient has (radiographic) sacroiliitis. Although, it has been shown that training does not improve the reliability in judgment of sacroiliitis on radiographic or structural lesions seen on MRI-SI ^{26,27}, we speculate that training can improve the assessment of spinal lesions. In **chapter 7** we observed that local readers are not confident in their decision on the presence/absence of MRI-spine lesions. Possibly there is room for improvement of the confidence on MRI-spine lesions,

because it appears to be fairly challenging to distinguish axSpA lesions from degenerative changes in the spine for readers without specific training²². Moreover, unpublished data shows that trained readers were able to reliably differentiate between lesions due to axSpA and lesions due to degeneration with little overlap between both sorts of lesions. Educating readers in the typical appearance of degenerative changes aside from axSpA lesions may improve the ability of the readers to differentiate one from the other. This training should focus on the characteristics of degenerative and axSpA findings on MRI and the training ought to be supervised by at least one experienced musculoskeletal radiologist. Whether training improves assessment of spinal lesions by a reader should be investigated. Should this improve it will have a positive effect on MRI-spine readings and the assessment of lesions including change over time.

To further understand the development of axSpA lesions, using only cross-sectional data will not be sufficient. Longitudinal data is required to answer questions concerning the occurrence and potential disappearance of lesions associated with axSpA. The longitudinal design of the DESIR and SPACE cohort offer a solid base for these kind of studies. The progression of inflammatory and structural lesions can be studied using follow-up data from these cohorts. In that perspective, when we assess inflammatory lesions we have to factor in that these lesions can be volatile. Therefore future research may also focus on whether axSpA lesions seen on MRI-SI and MRI-spine are ‘site-stable’ and if these lesions evolve over time. This may provide noteworthy insight in whether taking an MRI at one moment is rather a snapshot or gives a stable perspective of the abnormalities over time. If the physician bases his/her diagnosis (partly) on the results of the MRI, as findings in **chapter 2** suggest, it is of importance to know if repeating an MRI after a few months or a year would have influenced this diagnosis.

In order to identify axSpA patients as early as possible, it is important to know which patients with chronic back pain develop axSpA lesions on imaging or who develop other clinical SpA features over time. This means that the baseline characteristics of patients that eventually developed axSpA over time should be investigated and that development of SpA features over time should be monitored. If and when these patients and their characteristics can be identified, it will improve our understanding of how to identify (early) axSpA. Moreover, it will improve referral of chronic back pain patients by general practitioner and other physicians to the rheumatologist.

References

1. Calin A, Porta J, Fries JF, et al. Clinical history as a screening test for ankylosing spondylitis. *JAMA* 1977;237:2613-4.
2. 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.
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. van Onna M, Jurik AG, 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.
5. Heuft-Dorenbosch L, Landewé R, Weijers R, et al. Combining information obtained from magnetic resonance imaging and conventional radiographs to detect sacroiliitis in patients with recent onset inflammatory back pain. *Ann Rheum Dis* 2006;65:804-8.
6. Reveille JD. Genetics of spondyloarthritis—beyond the MHC. *Nat Rev Rheumatol* 2012;8:296-304.
7. 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.
8. 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.
9. 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:ii1-ii44.
10. 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;68:1520-27.
11. Madsen KB, Schiøttz-Christensen B, Jurik AG. Prognostic significance of magnetic resonance imaging changes of the sacroiliac joints in spondyloarthritis—a followup study. *J Rheumatol* 2010;37:1718-27.
12. Wick MC, Weiss RJ, Jaschke W, et al. Erosions are the most relevant magnetic resonance imaging features in quantification of sacroiliac joints in ankylosing spondylitis. *J Rheumatol* 2010;37:622-7.
13. Weber U, Lambert RGW, Pedersen SJ, et al. Assessment of structural lesions in sacroiliac joints enhances diagnostic utility of magnetic resonance imaging in early spondylarthritis. *Arthritis Care Res (Hoboken)* 2010;62(12):1763-71.
14. Althoff CE, Sieper J, Song I, et al. Active inflammation and structural change in early active axial spondyloarthritis as detected by whole-body MRI. *Ann Rheum Dis* 2013;72:967-3.
15. Krohn M, Braum LS, Sieper J, et al. Erosions and fatty lesions of sacroiliac joints in patients with axial spondyloarthritis: Evaluation of different mri techniques and two scoring methods. *J Rheumatol* 2014;41:473-80.
16. Weber U, Pedersen SJ, Zubler V, et al. Fat infiltration on magnetic resonance imaging of the sacroiliac joints has limited diagnostic utility in nonradiographic axial spondyloarthritis. *J Rheumatol* 2014;41:75-83.
17. Hermann K, Baraliakos X, van der Heijde D, et al. Descriptions of spinal MRI lesions and definition of a positive MRI of the spine in axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI study group. *Ann*

- Rheum Dis 2012;71:1278-88.
18. Weber U, Zhao Z, Rufibach K, et al. Diagnostic utility of candidate definitions for a positive MRI of the spine in patients with axial spondyloarthritis. *Arthritis Rheum* 2015;67:924-33.
19. Bennett A, Rehman A, Hensor EM, et al. The fatty Romanus lesion: a non-inflammatory spinal MRI lesion specific for axial spondyloarthropathy. *Ann Rheum Dis* 2010;69:891-4.
20. van den Berg R, Lenczner G, Feydy A, et al. Agreement between clinical practice and trained central reading in reading of sacroiliac joints on plain pelvic radiographs: Results from the DESIR cohort. *Arthritis Rheumatol* 2014;66:2403-11.
21. van den Berg R, Lenczner G, Thévenin F, et al. 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* 2015;74:2016-21.
22. Baraliakos X, Hermann KG a, Braun J. Imaging in Axial Spondyloarthritis: Diagnostic Problems and Pitfalls. *Rheum Dis Clin North Am* 2012;38:513-22.
23. 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.
24. Kjaer P, Leboeuf-Yde C, Korsholm L, et al. Magnetic resonance imaging and low back pain in adults: a diagnostic imaging study of 40-year-old men and women. *Spine* 2005;30:1173-80.
25. Brinjikji W, Diehn FE, Jarvik JG, et al. MRI Findings of Disc Degeneration are More Prevalent in Adults with Low Back Pain than in Asymptomatic Controls: A Systematic Review and Meta-Analysis. *Am J Neuroradiol* 2015;36:2394-9.
26. Poddubnyy D, Gaydukova I, Hermann KG, 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.
27. A van Tubergen, L Heuft-Dorenbosch, G Schulp, et al. Radiographic assessment of sacroiliitis by radiologists and rheumatologists: does training improve quality? *Ann Rheum Dis* 2003;62:519-25.

10

Nederlandse samenvatting

Samenvatting en conclusies

Spondyloartritis (SpA) is een reumatische ziekte die bij ongeveer 1% van de bevolking voorkomt. SpA is een ontstekingsziekte waarbij zowel gewrichten van de extremiteiten zoals knie of pols, als de wervelkolom en het bekken, voornamelijk de sacro-iliacale (SI) gewrichten, aangedaan kunnen zijn.

De SI-gewrichten zijn de gewrichten tussen het onderste gedeelte van de wervelkolom -het heiligbeen- en de darmbeenderen van het bekken. Wanneer de ontsteking voornamelijk gewrichten in de extremiteiten treft, spreekt men van perifere SpA. Bij axiale SpA (axSpA) betreft het de ontsteking van het axiale skelet, oftewel de wervelkolom en de SI-gewrichten. Ontsteking in de SI-gewrichten kan resulteren in klachten van stijfheid of pijn in de rug. Bij het merendeel van de patiënten ontstaan deze rugklachten tussen het 20^{ste} en 30^{ste} levensjaar en maar zelden ontstaat de ziekte na het 45^e levensjaar.

De symptomen van axSpA zijn erg uiteenlopend. Er is niet één symptoom dat bij alle axSpA patiënten voorkomt. Wel zijn er symptomen die veel vaker voorkomen bij patiënten met axSpA dan bij patiënten met andere ziekten van de rug. Deze symptomen worden 'SpA kenmerken' genoemd. Een veelvoorkomend SpA kenmerk is inflammatoire rugpijn. Artsen stellen de aanwezigheid van inflammatoire rugpijn vast aan de hand van meerdere symptomen: de aanwezigheid van rugpijn in rust, vermindering van rugpijn bij beweging, last van stijve rug in de ochtend die meer dan 30 minuten aanhoudt, nachtelijke rugpijn, pijn in het bilgebied waarbij de pijn wisselt van kant en het begin van de rugklachten voor het 40^{ste} levensjaar. Naast inflammatoire rugpijn kan een pijnlijke zwelling in de perifere gewrichten -perifere artritis- of ter hoogte van een peesaanhechting

-enthesis- of van een gehele vinger of teen -dactylitis- waargenomen worden. Deze zwelling is het gevolg van ontsteking. Ook ontstekingen buiten het skelet zoals een acute inwendige oogontsteking -acute uveïtis-, inflammatoire darmziektes zoals de ziekte van Crohn of colitis ulcerosa en psoriasis zijn symptomen die kunnen voorkomen bij patiënten met axSpA. Het substantieel afnemen van de pijn en stijfheid bij het gebruik van niet-steroïde ontstekingsremmende geneesmiddelen (NSAIDs) zoals diclofenac of naproxen -goede response op NSAIDs- en het voorkomen van axSpA of axSpA gerelateerde ziekten bij eerstegraads of tweedegraads bloedverwanten -positieve familieanamnese- worden ook gezien als kenmerken van axSpA. Patiënten met axSpA hebben vaak ook verhoogde ontstekingswaarden in het bloed zoals het C-reactief proteïne (CRP) en bezinkingssnelheid van erythrocyten (BSE). Daarnaast wordt

het gen HLA-B27 veel vaker gevonden in patiënten met axSpA. De hierboven genoemde ontsteking in (een van) de SI-gewrichten is kenmerkend voor axSpA. Deze ontsteking wordt sacroiliitis genoemd en kan zichtbaar gemaakt worden met o.a. Magnetic Resonance Imaging (MRI) of röntgenonderzoek.

Of en welke van de bovengenoemde SpA kenmerken een patiënt heeft, kan via lichamelijk onderzoek, de anamnese van de patiënt, laboratoriumonderzoek en beeldvormend onderzoek, zoals MRI of röntgenfoto's, worden uitgezocht. Met behulp van deze onderzoeken krijgt de arts een gedetailleerd klinisch beeld met als doel een passende diagnose en eventuele behandeling voor de patiënt te bieden. Om goed onderzoek te kunnen doen is het echter belangrijk om goed gedefinieerde groepen te creëren van alle patiënten die de diagnose axSpA van de arts gekregen hebben. Om dit soort groepen te creëren worden classificatie sets gebruikt waarbij aan de hand van overeenkomstige kenmerken patiënten worden ondergebracht in een categorie, in dit geval axSpA.

Een van de meest bekende classificatiecriteria in het veld van axSpA stammen uit 1984 en zijn de 'modified New York' (mNY) criteria. Deze criteria classificeren patiënten met Ankyloserende Spondylitis (AS), wat in de volksmond ook wel de ziekte van Bechterew wordt genoemd. AS is de meest uitgesproken vorm van axSpA en wordt gekenmerkt door de aanwezigheid van radiografische sacroiliitis. Radiografische sacroiliitis betreft de schade aan het bot, ten gevolge van ontsteking, die middels röntgenfoto's van de SI-gewrichten wordt vastgelegd. Deze structurele schade wordt in de mNY criteria uitgedrukt via een in ernst oplopende gradatie van 0 tot 4 per SI-gewricht. Radiografische sacroiliitis staat in de mNY criteria gelijk aan graad ≥ 2 aan beide SI-gewrichten of graad 3-4 aan een enkel SI-gewricht. Om een patiënt volgens de mNY criteria te classificeren dient de patiënt naast radiografische sacroiliitis ook minimaal een van de volgende klinische kenmerken te hebben:

- 1) lage rugpijn gedurende minstens 3 maanden die verbetert door beweging maar niet door rust
- 2) verminderde beweeglijkheid van de onderrug bij het zijwaarts buigen en bij naar voren en achteren buigen
- 3) verminderd vermogen om de borstkas uit te zetten bij inademing vergeleken met de normaalwaarde voor geslacht en leeftijd

Lange tijd waren röntgenfoto's van de SI-gewrichten de belangrijkste manier van beeldvorming in het axSpA-veld. Röntgenfoto's hebben echter als nadeel dat ze alleen structurele afwijkingen aan het bot tonen. Deze afwijkingen worden beschouwd als de gevolgen van ontsteking terwijl de daadwerkelijke ontsteking niet te detecteren is via

radiografie. Tussen het ontstaan van de rugklachten en de zichtbare sacroiliitis op de röntgenfoto's zit vaak 6 tot 8 jaar. Daarnaast zal maar een gedeelte van de patiënten met axSpA uiteindelijk AS ontwikkelen. Om ook de groep axSpA patiënten zonder radiologische afwijkingen te kunnen classificeren zijn begin jaren '90 van de vorige eeuw de Amor en European Spondyloarthritis Study Group (ESSG) criteria ontwikkeld waarin ook rekening gehouden wordt met patiënten zonder structurele afwijkingen op de röntgenfoto's. Beide criteria sets hechten veel belang aan de klinische eigenschappen en het hebben van radiologische afwijkingen is wel opgenomen in de criteria maar is niet verplicht gesteld. De Amor criteria hebben zelfs helemaal geen verplichte symptomen. Factoren die geassocieerd worden met axSpA hebben een score van 1 tot 3 gekregen. Om te voldoen aan de Amor criteria moeten er in totaal 6 of meer punten gescoord worden. De ESSG criteria hebben wel verplichte karakteristieken, namelijk de aanwezigheid van inflammatoire rugpijn of synovitis (gewrichtsontsteking) die asymmetrisch of overwegend in de onderste extremiteiten aanwezig is.

Met de komst van de ESSG en Amor criteria betekende de afwezigheid van radiografische sacroiliitis niet langer ook de afwezigheid van ziekte. Hierdoor raakte men ook meer geïnteresseerd in beeldvormende methoden die het mogelijk maken ontsteking af te beelden. Een techniek van beeldvorming die dit mogelijk maakt is MRI, waarmee naast scherpe anatomische details ook informatie over de samenstelling van het bot en weke delen kan worden verkregen. Derhalve is dit voor het vinden van afwijkingen die passen bij axSpA in een vroege fase een uitstekende techniek. Om het resultaat van een MRI-scan goed in beeld te brengen wordt het SI-gewricht door een computer als een aantal achtereenvolgende coupes gepresenteerd. Er wordt op deze manier veel meer informatie verzameld dan bij een röntgenfoto. Een MRI bestaat op zijn minst uit twee series die tezamen worden beoordeeld. Om ontsteking zichtbaar te maken is een T2-gewogen of Short Tau Inversion Recovery (STIR) opname de beste methode, omdat met deze opname water (die vrijkomt bij ontsteking zogenaamd ontstekingsoedeem) zichtbaar wordt. Bij structurele schade is het bot zo aangetast dat er meer definitieve veranderingen zichtbaar zijn in de begrenzing en de samenstelling van het bot vergeleken met een gezond SI-gewricht. Om structurele afwijkingen te visualiseren kan het best een T1-gewogen opname gebruikt worden.

De opkomende interesse voor MRI leidde ertoe dat deze beeldvormende methode werd geïncorporeerd in het Berlijn algoritme, een hulpmiddel dat ontwikkeld werd om reumatologen te begeleiden in het diagnostische proces van (vroege) axSpA. Middels

dit algoritme wordt de reumatoloog geadviseerd voor welke patiënten met rugpijn het raadzaam is om HLA-B27 typering te testen of een MRI te maken teneinde het vertrouwen van de reumatoloog in de diagnose te verhogen. Door het Berlijn algoritme te gebruiken wordt het aantal onnodige diagnostische testen beperkt, hetgeen kosten, tijd en moeite bespaart.

In 2009 is MRI officieel opgenomen in classificatiecriteria voor axSpA. Naast de MRI zijn de SpA kenmerken gebundeld in de classificatiecriteria voor axSpA die ontwikkeld zijn door de ‘Assessment of Spondyloarthritis international Society’ (ASAS) werkgroep. Deze criteria kunnen worden toegepast op patiënten die 3 maanden of langer rugklachten hebben, die voor het 45ste levensjaar begonnen zijn en waarbij de diagnose axSpA door de arts is gesteld. Een patiënt kan worden geclassificeerd met axSpA wanneer er sacroiliitis op de röntgenfoto of MRI wordt waargenomen naast minimaal een van de SpA kenmerken. Voor radiografische sacroiliitis bestaat het gradatiesysteem van de mNY criteria en dat wordt ook in de ASAS criteria aangehouden om radiografische sacroiliitis vast te stellen. Ook sacroiliitis op MRI heeft een vaste omschrijving in de ASAS criteria. De werkgroep heeft een definitie opgesteld en wanneer aan deze definitie wordt voldaan, spreekt men van een ‘positieve MRI’. De belangrijkste kenmerken van de definitie zijn:

- 1) de ontsteking heeft karakteristieken die typisch zijn voor axSpA
- 2) ontsteking van alleen het gewrichtskapsel of de peesaanhechting zijn niet genoeg. Op de MRI moet ontsteking in het beenmerg van het heiligbeen en/of de darmbeenderen aanwezig zijn
- 3) ontstekingsoedeem dat op meerdere plaatsen op een MRI coupe te zien zijn. Wanneer er slechts 1 plek van ontstekingsoedeem aanwezig is, moet dit op minimaal 2 achtereenvolgende MRI coupes te zien zijn.

Wanneer patiënten geen ontsteking van het SI-gewricht vertonen, kunnen ze toch geclassificeerd worden als axSpA wanneer het HLA-B27 gen aanwezig is in combinatie met minimaal 2 andere SpA kenmerken. Deze manier van classificeren binnen de ASAS criteria wordt ook wel de ‘klinische arm’ genoemd. Deze arm is toegevoegd omdat niet alle axSpA patiënten sacroiliitis op beeldvorming ontwikkelen of dat niet altijd (alle) beeldvorming aanwezig is voor alle patiënten.

Dit proefschrift

In dit proefschrift wordt er aandacht besteed aan 2 verschillende onderzoeksvelden. Het eerste deel van het proefschrift gaat over de diagnose en **classificatie** van axSpA patiënten. Het beschrijft niet alleen de overeenkomsten tussen verschillende criteria

sets maar ook hoe goed de criteria sets presteren ten opzichte van de diagnose van de reumatoloog. Daarnaast gaat dit gedeelte in op de prestaties van het Berlijn algoritme bij verscheidene wijzigingen van de toelatingscriteria en of deze eventuele wijzigingen van toegevoegde waarde zijn bij het herkennen van axSpA patiënten.

Het tweede gedeelte van dit proefschrift heeft betrekking op de **beeldvormend onderzoek** en de interpretatie van beeldvormend onderzoek. Hierin wordt de waarde van het gebruik van intraveneuze toediening van contrastvloeistof bij het maken van MRI's besproken. Ook het onderscheidend vermogen van structurele afwijkingen in de SI-gewrichten en afwijkingen in de rug op MRI tussen axSpA en niet-axSpA patiënten wordt behandeld. Tot slot worden bevindingen over de betrouwbaarheid van de beoordeling van afwijkingen in de rug die typisch zijn voor axSpA toegelicht en wordt er gekeken of er een relatie is tussen pijn in de SI-gewrichten of in de rug en MRI afwijkingen op dezelfde plek.

In deze onderzoeken wordt gebruik gemaakt van gegevens uit 2 grote patiëntonderzoeken die veel overeenkomsten vertonen maar ook een aantal belangrijke verschillen hebben. Het DEvenir des Spondylarthropathies Indifférenciées Récentes (DESIR) cohort bestaat uit patiënten van 18 tot en met 50 jaar afkomstig uit 25 verschillende ziekenhuizen in Frankrijk. Alle patiënten hadden inflammatoire rugpijn die 3 maanden of langer, maar niet langer dan 3 jaar, aanwezig was. Bovendien waren de reumatologen van mening dat de waarschijnlijkheid dat deze patiënt de diagnose axSpA heeft meer dan 50% is. Het Spondyloarthritis Caught Early (SPACE) cohort bevat patiënten in dezelfde leeftijdscategorie: alle patiënten waren 16 jaar of ouder maar waren bij inclusie niet ouder dan 47 jaar zijn omdat de rugklachten voor het 45^{ste} levensjaar ontstaan moesten zijn en minimaal 3 maanden maar niet langer dan 2 jaar aanwezig mochten zijn. In tegenstelling tot het DESIR cohort, hadden niet alle patiënten inflammatoire rugpijn maar de patiënten in het SPACE cohort hadden wel allemaal chronische rugpijn waarvan de oorzaak onbekend was.

Classificatiecriteria axSpA

Toen de ASAS classificatiecriteria voor axSpA werden ontwikkeld, is er gekeken naar de overeenkomst tussen de classificatie van patiënten volgens de ASAS criteria en de diagnose axSpA van expert reumatologen. De criteria bleken goed overeen te komen met de diagnose van de reumatoloog. Echter, het cohort dat is gebruikt voor de ontwikkeling van de ASAS criteria bevatte alleen patiënten met een hoge verdenking op axSpA. In **hoofdstuk 2** wordt beschreven of de ASAS criteria ook zo goed presenteren wanneer de kans dat een patiënt met rugklachten axSpA heeft kleiner is.

Niet alleen de prestaties van de ASAS criteria worden afgezet tegen de diagnose van de reumatoloog maar ook de prestaties van de eerder ontwikkelde classificatiecriteria (ESSG en Amor). Dit onderzoek hebben we uitgevoerd met gegevens van het SPACE cohort en er is gebleken dat bijna 60% van de patiënten met rugpijn aan minimaal een van de drie criteria sets (ASAS, ESSG of Amor) voldoet en dat 2 op de 5 patiënten in het SPACE cohort voldoen aan de ASAS criteria. Afgezet tegen de diagnose van reumatoloog, presteerden de ASAS criteria het best. Omdat de ESSG en Amor criteria van origine de MRI niet meenemen zou men kunnen zeggen dat het niet correct is om de prestatie van de ASAS criteria te vergelijken met de prestaties van de ESSG/Amor criteria. Maar wanneer we de definitie ‘positieve MRI’ zoals omschreven in de ASAS criteria toevoegen als eigenschap aan de ESSG en Amor criteria, zien we nog steeds dat de prestaties van de ASAS criteria beter zijn.

Verder bestond het idee dat patiënten die geen (radiografische) sacroiliitis hebben maar wel geclassificeerd worden als axSpA, omdat ze aan de klinische arm van de ASAS criteria voldoen, misschien niet dezelfde ziekte hebben. De klinische arm van de criteria zou uit patiënten kunnen bestaan die eigenlijk niet geclassificeerd zouden mogen worden als axSpA omdat de kenmerken teveel zouden verschillen van patiënten met (radiografische) sacroiliitis. Maar **hoofdstuk 2** laat zien dat dit niet het geval is. Patiënten met (radiografische) sacroiliitis en patiënten die aan de klinische arm van de ASAS criteria voldoen, tonen dezelfde ziekteactiviteit en ook de duur van de rugklachten en de frequentie van SpA kenmerken zijn gelijk. Er was wel één kenmerk dat vaker voorkwam bij patiënten die aan de klinische arm voldeden, namelijk een positieve familieanamnese. Dit ligt echter in de lijn der verwachting aangezien dit kenmerk sterk geassocieerd is met de aanwezigheid van HLA-B27. En de aanwezigheid van HLA-B27 is weer een kenmerk dat verplicht is om aan de klinische arm van de ASAS criteria te kunnen voldoen.

De aanwezigheid van SpA kenmerken bij axSpA en niet-axSpA patiënten is onderzocht in **hoofdstuk 2**. Hieruit blijkt onder meer dat axSpA patiënten vaker inflammatoire rugpijn hebben dan niet-axSpA patiënten. Echter het blijkt ook dat niet alle axSpA patiënten inflammatoire rugpijn hebben: 17% van de axSpA patiënten vertoont dit kenmerk niet. Inflammatoire rugpijn wordt gezien als een belangrijk kenmerk en huisartsen wordt aanbevolen patiënten met inflammatoire rugpijn door te verwijzen naar de reumatoloog. Desalniettemin komen de resultaten uit **hoofdstuk 2** overeen met eerdere literatuur waarin beschreven wordt dat er axSpA patiënten zijn zonder inflammatoire rugpijn en patiënten met inflammatoire rugpijn die geen axSpA

hebben. Dit leidde tot de vraag of het Berlijn algoritme in de huidige vorm, waarbij inflammatoire rugpijn een vereiste is, wel de optimale vorm van het Berlijn algoritme is. Zoals hierboven beschreven is het Berlijn algoritme een instrument dat de reumatoloog kan gebruiken om haar/hem te begeleiden in het diagnostische proces van (vroeg) axSpA. In **hoofdstuk 3** wordt onderzocht of een verandering aan het Berlijn algoritme de prestatie verbetert. Deze studie vindt plaats in 2 cohorten; het SPACE cohort en het ASAS cohort. Het ASAS cohort is opgesteld ter validatie van de ASAS classificatiecriteria en bestaat uit patiënten met chronische rugpijn die 3 maanden of langer aanhoudt en begonnen is voor het 45^{ste} levensjaar met een verdenking op axSpA.

In **hoofdstuk 3** worden 2 veranderingen voorgesteld die allebei betrekking hebben op inflammatoire rugpijn. In het originele Berlijn algoritme is het hebben van inflammatoire rugpijn een vereiste. Een patiënt zonder inflammatoire rugpijn zal het algoritme niet doorlopen en de reumatoloog wordt voorgesteld een andere diagnose dan axSpA te overwegen. Beide voorgestelde veranderingen hebben betrekking op inflammatoire rugpijn. Met de eerste verandering worden de criteria versoepeld waaraan voldaan moet worden om inflammatoire rugpijn te hebben, maar inflammatoire rugpijn blijft een vereiste stap in het algoritme. De andere verandering elimineert inflammatoire rugpijn als verplichte stap. Inflammatoire rugpijn wordt toegevoegd bij de rest van de SpA kenmerken die in de daaropvolgende stap in het algoritme aan bod komen. Beide veranderingen blijken de prestatie van het algoritme te verbeteren, maar de verbetering is het grootst bij de modificatie waarbij inflammatoire rugklachten worden verwijderd als verplicht kenmerk. In beide cohorten wordt deze verandering als de beste verbetering gezien. Aan de hand van de bevindingen in **hoofdstuk 3** heeft de ASAS werkgroep de verandering, waarbij inflammatoire rugpijn geen vereiste meer is, geaccepteerd en stelt de werkgroep voor de gemodificeerde versie van het Berlijn algoritme te gebruiken in plaats van de originele versie.

Beeldvormend onderzoek en de interpretatie van beeldvorming

De ASAS werkgroep heeft een handboek ontwikkeld waarin de condities beschreven staan waaraan een MRI en het protocol van het maken van een MRI moeten voldoen. In het handboek worden de STIR en T1 post gadolinium (T1/Gd) opnames aanbevolen om ontsteking op MRI in beeld te brengen. Er kleven een aantal nadelen aan het toedienen van een intraveneuze contrastvloeistof, zoals gadolinium-DTPA, bij het maken van een MRI: het is tijdrovend, erg duur, er kunnen ernstige bijwerkingen ontstaan bij patiënten waarvan de nieren niet meer goed functioneren en omdat er een

inфуus moet worden ingebracht kan het oncomfortabel zijn voor de patiënt. Echter, met de T1/Gd opname kunnen meer anatomische structuren met ontsteking zichtbaar worden gemaakt, in vergelijking met de STIR opname. **Hoofdstuk 4** gaat in op de vraag of deze invasieve methode van het verkrijgen van MRI's van de SI-gewrichten cruciaal is om axSpA patiënten beter te kunnen herkennen. In **hoofdstuk 4** is te zien dat ontsteking van het heiligbeen en/of de darmbeenderen die gezien wordt op de T1/Gd opname in alle gevallen óók gezien wordt op de STIR opname. Met de T1/Gd opname worden ontstekingen van de peesschede en het gewrichtskapsel wel beter afgebeeld dan met de STIR. Dit soort ontsteking op zichzelf is niet voldoende om een positieve MRI te hebben volgens de definitie van ASAS werkgroep; er moet ontsteking van het heiligbeen en/of de darmbeenderen aanwezig zijn. In het SPACE cohort, waar dit onderzoek in is afgenomen, hadden alle patiënten met ontstekingen van de peesschede of het gewrichtskapsel ook ontsteking van het heiligbeen en/of de darmbeenderen. Dit betekent dat het gebruik van de contrastvloeistof gadolinium geen toegevoegde waarde heeft voor het detecteren van ontsteking die van belang is in de classificatie van axSpA patiënten. Als direct gevolg van deze resultaten is men sinds april 2012 gestopt met het toedienen van gadolinium bij MRI's die ten behoeve van het SPACE onderzoek werden gemaakt.

De definitie van een 'positieve MRI' die door de ASAS werkgroep is opgesteld en die is opgenomen in de ASAS classificatiecriteria voor axSpA, beschrijft alleen afwijkingen die betrekking hebben op de ontsteking van de SI-gewrichten. Met MRI kan men niet alleen ontsteking maar ook structurele afwijkingen visualiseren. Toch zijn deze afwijkingen niet meegenomen in de definitie van een positieve MRI. Hetzelfde kan gezegd worden voor MRI-afwijkingen (van inflammatoire of structurele aard) in de rug; ook deze afwijkingen zijn niet meegenomen in de ASAS classificatiecriteria. De reden hiervoor was het gebrek aan bewijs dat dit soort afwijkingen daadwerkelijk behulpzaam zouden kunnen zijn bij de classificatie van axSpA patiënten. Voordat men kan onderzoeken of de definitie van een positieve MRI verbetert door de toevoeging van structurele MRI-afwijkingen in de SI-gewrichten of MRI-afwijkingen in de rug, is het van belang om te weten of er daadwerkelijk afwijkingen zijn die regelmatig voorkomen in axSpA patiënten. In **hoofdstuk 5** wordt niet alleen gekeken of er afwijkingen zijn die regelmatig voorkomen maar ook hoeveel afwijkingen er bij axSpA patiënten voorkomen. Het betreft zowel structurele MRI-afwijkingen van de SI-gewrichten als MRI-afwijkingen in de rug (ontsteking en structurele schade). In de literatuur zijn er wel gegevens bekend over het aantal afwijkingen in axSpA patiënten, maar omdat deze studies geen controlegroep bevatten is het moeilijk om te zeggen of vergelijkbare aantallen in niet-

axSpA patiënten voorkomen. Door het ontwerp van het SPACE cohort is er MRI data aanwezig van patiënten die geen axSpA hebben maar wel chronische rugpijn. Deze groep kan daardoor als controlegroep fungeren. Door de aanwezigheid van de controlegroep is het mogelijk om het aantal afwijkingen op MRI van de SI-gewrichten (MRI-SI) en de rug (MRI-rug) te bepalen in axSpA en niet-axSpA patiënten. We hebben in **hoofdstuk 5** gekeken of er bepaalde afwijkingen zijn die substantieel vaker voorkomen bij axSpA patiënten. Uit de analyses van dit hoofdstuk is de regel van vijf (*the rule of five*) voortgekomen die de afwijkingen en het aantal afwijkingen beschrijft dat kenmerkend is voor axSpA. Het blijkt dat ‘5 of meer vetafwijkingen en/of erosies op MRI-SI’, net als ‘5 of meer inflammatoire afwijkingen op MRI-rug’ en ‘5 of meer vetafwijkingen op MRI-rug’ veel voorkomen bij axSpA patiënten terwijl dit aantal afwijkingen slechts bij een paar niet-axSpA patiënten (<5%) wordt gezien.

Niet alleen het type en het aantal afwijkingen is van belang in het beeldvormend onderzoek naar axSpA, maar er moet ook rekening gehouden worden met de correcte beoordeling van de beeldvorming. In de dagelijkse praktijk worden beelden beoordeeld door een lokale radioloog en/of reumatoloog. Onder andere op basis van dit oordeel neemt de reumatoloog een beslissing over de diagnose van de patiënt. In een onderzoek setting is er echter een andere procedure. Hier worden de beelden doorgaans beoordeeld door meerdere ‘lezers’, die centraal zijn getraind. Vaak hebben ze naast een training ook minimaal één kalibratiesessie gehad waarin de leesprocedure en valkuilen zijn besproken. Het is van belang om te weten of de oordelen over de beeldvorming die verkregen zijn via deze twee verschillende methodes (de lokale lezing en de centrale lezing) overeenkomen met elkaar. Met andere woorden, komen de bevindingen gegeven in de dagelijkse praktijk overeen met die uit het onderzoek? Het DESIR cohort biedt de juiste factoren om dit te onderzoeken omdat dit cohort het unieke ontwerp heeft waarin zowel de bevindingen op röntgenfoto’s en MRI’s van lokale lezers als centrale lezers beschikbaar zijn. Uit eerder onderzoek dat uitgevoerd is binnen het DESIR cohort bleek, dat de overeenstemming tussen de lokale en centrale lezing van de röntgenfoto’s en MRI’s van de SI-gewrichten matig tot substantieel was. Het verschil tussen beide lezingen leidde in sommige gevallen tot misclassificatie van patiënten waarbij lokale lezers, vooral betreffende de röntgenfoto’s, de classificatie axSpA/AS overschatten.

In **hoofdstuk 6** en **hoofdstuk 7** wordt het oordeel van lokale lezers vergeleken met het oordeel van centrale lezers voor röntgenfoto’s en MRI van de rug. De overeenstemming tussen centrale lezers is hoger dan tussen de lokale en centrale lezing, maar het niveau van overeenstemming is matig. Bij het beoordelen van de

röntgenfoto's van de rug zien lokale lezers afwijkingen die niet gezien worden door de centrale lezers. Dezelfde trend zien we bij eerder onderzoek van de vergelijking tussen lokale en centrale lezing van de SI-gewrichten. Voor de bevindingen op de MRI-rug zien we echter dat de centrale lezers meer afwijkingen zien dan de lokale lezers. Dit kan verklaard worden door de verschillende methodes van het verzamelen van de gegevens. De centrale lezers hebben voor elke individuele MRI-rug afwijking (inflammatie, vet, erosies en syndesmofyten) per wervel aangegeven of deze aanwezig/afwezig was. Dit is een veel gedetailleerder oordeel dan dat wat de lokale lezers hebben gegeven: aanwezigheid van enige inflammatoire afwijking in de rug en aanwezigheid van enige structurele afwijkingen (vet, erosies en syndesmofyten) in de rug. Daarnaast hadden de lokale lezers de mogelijkheid om hun onzekerheid over de aanwezigheid van afwijkingen te uiten. Dus waar de centrale lezers bij twijfel gedwongen waren een keuze te maken over de aanwezigheid van afwijkingen, hadden lokale lezers de antwoordoptie 'twijfel' waardoor een concreet antwoord uitbleef. Voor de vergelijking tussen de lezingen zijn alleen de antwoordopties 'aanwezig' en 'afwezig' gebruikt en is de optie 'twijfel' niet meegenomen omdat het onmogelijk is om te bepalen of de lokale lezer in dat geval naar de antwoordoptie 'aanwezig' of naar de optie 'afwezig' neigt. In **hoofdstuk 7** is te lezen dat bij een groot deel van de patiënten de lokale lezer twijfelt over de aanwezigheid van afwijkingen op MRI-rug. Dit beïnvloedt de resultaten en met deze discrepantie in antwoordmethodes moet dan ook zeker rekening gehouden worden bij de interpretatie van de resultaten.

In **hoofdstuk 8** gaan we verder in op de interpretatie van beeldvorming. In dit hoofdstuk onderzoeken we de vraag of het hebben van MRI-afwijkingen (zowel MRI-SI als MRI-rug) op een bepaalde locatie ook betekent dat de patiënt pijn heeft op diezelfde locatie. Uit de resultaten blijkt dat pijn gerelateerd is aan verscheidene afwijkingen. In de SI-gewrichten zien we dat de kans dat een patiënt last heeft van bilpijn met 6% toeneemt voor elke inflammatoire MRI-afwijking die hij/zij heeft in de SI-gewrichten. En met een maximum van 53 inflammatoire MRI-SI afwijkingen kan die kans heel hoog worden. Ook het hebben van erosies op MRI-SI blijkt gerelateerd te zijn aan bilpijn maar alleen in patiënten die jonger dan 25 jaar zijn; in deze groep neemt de kans dat een patiënt aangeeft last te hebben van bilpijn met 16% per erosie toe.

In deze analyses is de rug onderverdeeld in 2 locaties; het thoracale gedeelte, dat verwijst naar het midden van de rug, en het lumbale gedeelte dat verwijst naar het onderste deel van de rug. De cervicale wervelkolom, het gedeelte tussen de nek en de schouderbladen, is niet meegenomen in deze analyses. In de analyses wordt er

rekening gehouden met de onderlinge relatie tussen thoracale en lumbale segmenten binnen 1 patiënt. Want wanneer een patiënt pijn (of afwijkingen) in een segment heeft, is de waarschijnlijkheid dat bij die patiënt ook pijn (of afwijkingen) in het andere segment aanwezig is hoger dan de aanwezigheid van pijn (of afwijkingen) tussen patiënten. Uit **hoofdstuk 8** blijkt dat ook afwijkingen in de rug gerelateerd zijn aan pijn op dezelfde locatie in de rug. Maar hier betreft het uitsluitend afwijkingen van degeneratieve aard, de axSpA afwijkingen op MRI-rug zijn niet geassocieerd met rugpijn op dezelfde locatie. Er zijn uiteenlopende degeneratieve afwijkingen die geassocieerd zijn met rugpijn op dezelfde locatie, maar het is interessant om te zien dat deze associaties alleen gevonden worden in patiënten die 35 jaar of ouder zijn en in patiënten die niet aan de ASAS axSpA classificatiecriteria voldoen, en dus een andere oorzaak voor hun rugpijn hebben.

Toekomstperspectieven

Nu de gemodificeerde versie van het algoritme officieel geaccepteerd is, is het zaak om te zorgen dat reumatologen en andere artsen (bijvoorbeeld huisartsen) het algoritme ook daadwerkelijk gaan gebruiken. Om het algoritme te implementeren in de dagelijkse praktijk is het van belang om informatie te verzamelen over het huidige gebruik van het algoritme door artsen, of ze van het bestaan van het algoritme weten en waarom ze het wel -of juist niet- zouden gebruiken in de kliniek. De informatie over het gebruik van het algoritme door reumatologen kan bovendien ook worden toegepast om de prestaties van het algoritme verder te verbeteren. Daarnaast zou de kosteneffectiviteit van beeldvormend onderzoek een goede focus voor de toekomst zijn. Het algoritme is namelijk zo opgebouwd dat het maken van een MRI van de SI-gewrichten beperkt blijft. Maar het maken van een MRI-SI zou weldegelijk nuttig kunnen zijn bij patiënten die HLA-B27 negatief zijn maar wel 2 tot 3 andere SpA kenmerken hebben. In het gemodificeerde algoritme wordt voor deze patiënten een andere diagnose gesuggereerd. Echter uit **hoofdstuk 3** blijkt dat er wel degelijk HLA-B27 negatieve patiënten zijn met 2 tot 3 SpA kenmerken waarbij de reumatoloog de diagnose axSpA heeft gegeven. De vraag is of de kosten die het maken van een MRI met zich meebrengen opwegen tegen de voordelen van het verder zoeken naar axSpA afwijkingen in de SI-gewrichten.

Door de regel van vijf (**hoofdstuk 5**) toe te passen kan men een onderscheid maken tussen axSpA en niet-axSpA patiënten met betrekking tot de beeldvorming. Deze regel is ook bruikbaar bij het beantwoorden van de vraag of de huidige definitie van een 'positieve MRI' die gebruikt wordt in de ASAS axSpA criteria, de optimale is.

Vanaf de ontwikkeling van de criteria is er een discussie gaande omtrent deze vraag. Sommigen vinden namelijk dat de definitie niet volledig genoeg is doordat deze alleen maar ontsteking van de SI-gewrichten in beschouwing neemt. De ASAS werkgroep heeft ten tijde van het opstellen van de criteria besloten (voorlopig) alleen inflammatoire MRI-SI afwijkingen te includeren in de definitie omdat er op dat moment bewijs ontbrak dat aangaf welke andere MRI-afwijkingen typerend zijn voor axSpA. Door de resultaten uit **hoofdstuk 5** is er nu bewijs volgens de regel van vijf. Echter, deze regel geldt voor patiënten met een korte klachtenduur toen ze begonnen met deelname aan de studie (≤ 2 jaar rugklachten) en moet bevestigd worden in een cohort met patiënten met langdurige rugklachten. Alvorens een beslissing te nemen over het aanpassen of behouden van de huidige definitie van een positieve MRI in de ASAS criteria is het van belang om te kijken of patiënten met afwijkingen zoals omschreven in deze regel ook volgens de reumatoloog axSpA hebben. Daarnaast zal niet alleen de klinische toegevoegde waarde maar ook het economische belang van een eventuele aanpassing van de definitie intensief moeten worden onderzocht, want het uitvoeren van MRI's is een kostbare zaak. Bovendien beschikken niet alle klinische centra over een MRI-scanner.

We hebben geconstateerd dat de lokale lezers weinig vertrouwen hebben in hun besluit over de aan- of afwezigheid van afwijkingen op MRI-rug. Omdat er door de lokale lezers m.b.t. zoveel patiënten wordt getwijfeld zou men zich kunnen afvragen of het vertrouwen van een lezer verbeterd kan worden door training. Uit eerder onderzoek is gebleken dat de betrouwbaarheid van beoordelingen van (radiografische) sacroiliitis niet verbetert met training. Dit hoeft echter niet het geval te zijn voor axSpA afwijkingen in de rug. Voor lezers zonder specifieke training blijkt het namelijk lastig om een onderscheid te maken tussen axSpA afwijkingen en afwijkingen in de rug met een andere oorzaak zoals degeneratie. Degeneratieve afwijkingen zijn afwijkingen in verschillende onderdelen van de rug zoals de tussenwervelschijf die smaller kan worden en gaat uitpuilen -hernia- waarbij de tussenwervelschijf tegen een zenuw in de rug duwt. Uit een onderzoek in het SPACE cohort dat nog niet gepubliceerd is, blijkt dat lezers die wel een specifieke training hebben gehad heel goed het onderscheid kunnen maken tussen afwijkingen typerend voor axSpA en degeneratieve afwijkingen. Dit pleit voor de stelling dat het mogelijk is om een onderscheid te maken en dat de betrouwbaarheid van de beoordeling van MRI-rug afwijkingen wellicht verbetert wanneer lezers een specifieke training krijgen.

Conclusies

Als laatste vat ik de belangrijkste resultaten van dit proefschrift samen. Zo blijken de ASAS classificatie criteria voor axSpA erg robuust en zijn ze goed toepasbaar in een setting waarbij de waarschijnlijkheid op axSpA lager is dan de setting waarin de criteria ontwikkeld zijn. Daarnaast is er op basis van bevindingen in dit proefschrift een belangrijke aanpassing doorgevoerd in een hulpmiddel dat reumatologen begeleidt in het diagnostische proces van (vroege) axSpA. Tevens is het MRI protocol van het SPACE cohort aangepast nadat de toegevoegde waarde van het toedienen van een contrastvloeistof bij het maken van MRI's nihil bleek te zijn, met als resultaat dat er sinds april 2012 geen MRI's meer worden gemaakt met de toediening van gadolinium. Het SPACE protocol beperkt zich tot STIR en T1-gewogen opnames sedert april 2012.

Dit proefschrift biedt ook adequate handvatten voor verder onderzoek naar de optimale definitie van een positieve MRI. Bovendien worden discrepanties aangaande de beoordeling door verschillende soorten lezers van afwijkingen in de rug (op röntgenfoto's en MRI's) aan het licht gebracht waardoor de vraag naar voren komt of training in het onderscheiden van axSpA afwijkingen en afwijkingen die niet gerelateerd zijn aan axSpA van potentiële belang is op de betrouwbaarheid van het beoordelen van afwijkingen in de rug.

De onderzoeken in dit proefschrift zullen hopelijk een bijdrage leveren aan beter herkenning en daardoor beter behandeling van axSpA.

List of publications

1. **de Hooge M**, van den Berg R, Navarro-Compán V, van Gaalen F, van der Heijde D, Huizinga T, Reijnerse M. Magnetic resonance imaging of the sacroiliac joints in the early detection of spondyloarthritis: no added value of gadolinium compared with short tau inversion recovery sequence. *Rheumatology (Oxford)* 2013;52(7):1220-4.
2. van den Berg R*, **de Hooge M***, Rudwaleit M, Sieper J, van Gaalen F, Reijnerse 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(10):1646-53.
3. van den Berg R, **de Hooge M**, van Gaalen F, Reijnerse 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(8):1492-9.
4. van Gaalen F, Bakker P, **de Hooge M**, Schoones J, van der Heijde D. Assessment of sacroiliitis by radiographs and MRI: where are we now? *Curr Opin Rheumatol* 2014;26(4):384-8.
5. van den Berg R, **de Hooge M**, Bakker PA, van Gaalen F, Navarro-Compán V, Fagerli KM, Landewé R, van Oosterhout M, Ramonda R, Reijnerse M, van der Heijde D. Metric properties of the SPARCC score of the sacroiliac joints - data from baseline, 3-month, and 12-month follow up in the SPACE Cohort. *J Rheumatol* 2015;42(7):1186-93.
6. Lorenzin M, Ortolan A, **de Hooge M**, Frallonardo P, Piccoli A, Cozzi F, Oliviero F, Punzi L, Ramonda R. Lengthening the time intervals between doses of biological agents in psoriatic arthritis patients: a single-center retrospective study. *Int J Immunopathol Pharmacol* 2015;28(4):479-87.

-
7. de Bruin F, ter Horst S, Bloem JL, van den Berg R, **de Hooze M**, van Gaalen F, Fagerli K, Landewé R, van Oosterhout M, Ramonda R, van der Heijde D, Reijnierse M. Prevalence of degenerative changes of the spine on MR and radiographs in patients between 16 and 45 years with chronic back pain of short duration in the SPondyloArthritis Caught Early (SPACE) cohort. *Rheumatology (Oxford)* 2016;55(1):56-65.
 8. **de Hooze M**, van den Berg R, Navarro-Compán V, Reijnierse M, van Gaalen F, Fagerli K, Landewé R, van Oosterhout M, Ramonda R, Huizinga T, van der Heijde D. Patients with chronic back pain of short duration from the SPACE-cohort: which MRI structural lesions in the sacroiliac joints and inflammatory and/or structural lesions in the spine are most specific for axial spondyloarthritis? *Ann Rheum Dis* 2016;75(7):1308-14.
 9. de Bruin F, Ter Horst S, van den Berg R, **de Hooze M**, van Gaalen F, Fagerli KM, Landewé R, van Oosterhout M, Bloem JL, van der Heijde D, Reijnierse M. Signal intensity loss of the intervertebral discs in the cervical spine of young patients on STIR sequences. *Skeletal Radiol* 2016 Mar;45(3):375-81.
 10. Bakker P, Ez-Zaitouni Z, van Lunteren M, van den Berg R, **de Hooze M**, Fagerli K, Landewé R, van Oosterhout M, Ramonda R, Reijnierse M, van der Heijde D, van Gaalen F. Patients aged 16-45 years with chronic back pain of a short duration and maximally 1 SpA-feature: are additional tests needed to rule out axSpA? *Arthritis Care Res (Hoboken)* 2016 Mar 18. doi: 10.1002/acr.22883. [Epub ahead of print].
 11. **de Hooze M***, de Bruin F*, de Beer L, Bakker P, van den Berg R, Ramiro S, van Gaalen F, Fagerli K, Landewé R, van Oosterhout M, Ramonda R, Huizinga T, Bloem L, Reijnierse M, van der Heijde D. Is site of back pain related to location of MRI lesion in patients with chronic back pain included in the SPACE-cohort included in the SPACE cohort? *Arthritis Care Res (Hoboken)*. (Accepted for publication).
 12. Claudepierre P*, **de Hooze M***, Feydy A, Reijnierse M, Saraux A, Dougados M, van der Heijde D. Reliability of mSASSS scoring in everyday practice in DESIR

cohort study centres: cross-sectional study of agreement with trained readers.
Ann Rheum Dis (Accepted for publication).

13. **de Hooge M**, Pialat J-B, Reijnierse M, van der Heijde D, Claudepierre P, Saraux A, Dougados M, Feydy A. Assessment of typical SpA lesions on MRI of the spine: do local readers and central readers agree in the DESIR-cohort at baseline? (Submitted).
14. **de Hooge M**, Ramonda R, Lorenzin M, Ortolan A, Frallonardo P, Punzi L, Doria A. Work productivity is associated with disease activity and functional ability in Italian patients with axial spondyloarthritis: Results for the SPACE-cohort. (Submitted).
15. Pellegrom S, **de Hooge M**, Huizinga T, van der Heijde D, van Gaalen F. Usage and performance of referral parameters for axial spondyloarthritis: data from the SPondyloarthritis Caught Early (SPACE) Cohort. (Submitted).
16. de Bruin F, Treyvaud MO, Feydy A, **de Hooge M**, Pialat J, Dougados M, Gossec L, Bloem JL, van der Heijde D, Reijnierse M. Prevalence of degenerative changes and overlap with spondyloarthritis-associated lesions in the spine of patients from the desir cohort. (Submitted).
17. Bakker P, van den Berg R, **de Hooge M**, van Lunteren M, Ez-Zaitouni Z, Fagerli K, Landewé R, van Oosterhout M, Ramonda R, Reijnierse M, van Gaalen F, van der Heijde D. Replacing sacroiliitis on radiographs by structural lesions seen on MRI in the SPACE cohort: impact on the classification of patients according to the ASAS axial SPondyloArthritis criteria. (Submitted).

*Both authors contributed equally

Curriculum vitae

Manouk Sumudu Madushani de Hooge werd geboren op 15 september 1986 in Negombo (Sri Lanka). In 2004 behaalde zij haar vwo diploma aan het Aloysius College in Den Haag, waarna ze een jaar Biomedische Wetenschappen studeerde aan de Universiteit van Leiden. Vervolgens maakte ze de overstap naar de bachelor Gezondheidswetenschappen aan de Universiteit van Maastricht. Tijdens haar studie was ze ook geïnteresseerd in andere disciplines en voltooide ze vakken in Beleid & Management, Moleculaire Levenswetenschappen en Geneeskunde. In 2010 sloot ze haar bachelor opleiding af.

In 2009 was ze al begonnen aan de master Human Movement Sciences aan de Vrije Universiteit in Amsterdam. Ze koos voor de richting 'Health' waarbij ze ook extra-curriculaire vakken afrondde: 'entrepreneurship in Human Movement Sciences' en 'spierfysiologie'.

In 2011 behaalde ze haar master diploma. In april van datzelfde jaar begon ze aan haar promotietraject in het Leids Universitair Medisch Centrum onder begeleiding van prof.dr. D. van der Heijde, dr. F. van Gaalen en dr. M. Reijnierse. Begin mei 2014 verhuisde ze naar Milaan (Italië), om vanuit daar verder te werken aan haar promotie. Vanaf dat moment werkt ze ook als zzp'er voor het beoordelen van röntgenfoto's en MRI's in verschillende projecten. Sinds half februari 2016 doet ze dit vanuit Brussel in België.

Dankwoord

In de periode waarin dit proefschrift is geschreven heb ik zowel op wetenschappelijk vlak als op persoonlijk vlak veel geleerd. Mijn dank gaat uit naar iedereen die mij geholpen en gesteund heeft.

Graag bedank ik mijn promotor prof.dr. van der Heijde en mijn co-promotoren dr. van Gaalen en dr. Reijnierse. Beste Désirée, toen ik begon aan dit traject had ik geen idee wat voor een geweldige kans ik van jou heb gekregen. Dank je wel voor je vertrouwen in mij en alle tijd en moeite die je in mij steekt! Jouw afwezigheid in het LUMC was nooit een barrière voor jouw goede begeleiding en ook dat waardeer en bewonder ik heel erg.

Beste Floris, jij hebt mij wegwijs gemaakt in de klinische wereld van SpA. Jouw enthousiasme voor SpA kwam daarin duidelijk naar voren en dit werkte aanstekelijk voor mij. Bedankt voor alle stimulerende (klinische en wetenschappelijke) overleggen die wij mochten voeren.

Beste Monique, bedankt voor de vele momenten op de radiologie afdeling waar jij uitleg gaf over de MRI's en röntgenfoto's. Met jouw hulp heb ik letterlijk en figuurlijk een duidelijk beeld gekregen van axSpA laesies.

Op de Reumatologie afdeling in het Leids Universitair Medisch Centrum ben ik aan dit proefschrift begonnen. Beste prof. Huizinga, beste Tom, het was fijn om jou te leren kennen en ik ben dankbaar voor alle mogelijkheden die jij en jouw afdeling mij hebben gegeven in de onderzoekswereld. Lieve kamergenootjes (C1-46) en 'de andere kamer'-genootjes (C1-45) dank jullie wel voor het creëren van een aangename en leergierige werksfeer.

Mijn promotietijd is onlosmakelijk verbonden met het SPACE project. Daarom wil ik alle patiënten bedanken die deelnemen en deel hebben genomen aan het SPACE onderzoek. Dank aan alle arts-assistenten, stafleden en artsen buiten het LUMC die patiënten hebben aangedragen en geïncludeerd voor SPACE. Ook het secretariaat (van de poli en de afdeling) alsmede Jozé en Cedric van datamanagement wil ik graag bedanken, want jullie hulp en inzet zorgen voor een gewichtige bijdrage aan de opbouw en het voortbestaan van dit mooie cohort. Een speciale vermelding voor Paul de Bruin, zonder jouw programma was het anonimiseren van de te scoren beelden een stuk lastiger geweest. Dank voor de tijd die je hierin hebt gestoken.

Lieve Rosaline, onze tijd samen als SPACE-team koester ik. We waren als team goed op elkaar ingespeeld. Dank je wel voor alle fijne momenten, zowel op werk als privé.

Lieve dames van het SPACE-team, Pauline, Zineb en Miranda, het was erg leuk om jullie te leren kennen en met jullie samen te werken. Ik wens jullie, en nieuwkomer Anoek, veel succes in de toekomst. Lieve Freek, superFreak ;), bedankt voor de samenwerking en onze bevorderlijke besprekingen op C6, en later via Skype.

To the SPACE-teams outside the LUMC; Amsterdam, Gouda, Oslo and Padua; dank, takk, grazie, thank you very much for the great collaboration we had together.

I am thankful to the people involved in the DESIR cohort for the nice collaboration, with special thanks to prof.dr. Dougados and prof.dr. Claudepierre. Dear Maxime and Pascal, it is truly inspiring to work with you.

To all other colleagues whom I met during meetings and conferences/congresses; thank you very much for the stimulating conversations and discussions.

Tijdens mijn promotietraject heb ik een aantal dierbare vriendschappen gesloten waarbij onze liefde voor de wetenschap de primaire, maar zeker niet de enige, gemene deler was.

Dear Victoria, thank you for all the (work related) heart-to-hearts that we have had. I admire how you combine work and private life. Lieve Sofia, ik ben heel blij met al jouw adviezen en bedankt voor het aanmoedigen en je steun in de laatste fases van mijn promotie. Lieve Annemarie, je energieke persoonlijkheid verblijdt mij en bedankt voor de leuke tijd in o.a. de borrelcommissie. Dear Cesar and dear Felice, I am happy with our special friendship and would like to thank you very much for your support during the challenging last phase in Leiden.

Lieve Emilia, met recht noem je mij je onbetwiste maatje gedurende onze LUMC tijd en dat is wederzijds. Heel erg bedankt voor de enorm plezante en leerzame tijd.

Lieve Pien, onze vriendschap gaat al zo ver terug en we gaan letterlijk over bergen en door zee voor elkaar. Bedankt voor deze sterke en pure vriendschap.

Aan mijn dierbare vrienden uit mijn studietijd in Den Haag, Maastricht en Amsterdam, dank jullie wel voor de belangstelling die jullie hebben getoond in mijn promotietraject en voor de hechte band die we hebben.

Lieve familie (met een speciale vermelding voor Katja) en schoonfamilie, ik ben dankbaar voor jullie interesse in mijn werk en dat jullie deel uit maken van mijn leven. Lieve oma, jouw liefde en verwennerij zijn het beste dat een kleinkind zich kan wensen, dank je wel daarvoor. Lief klein zusje (Kim), lieve grote zus (Annemiek) en lieve grote broers (Joris en Menno), het is een genoegen om jullie grote/kleine zus te

zijn. Lief broertje, lieve Rouwan, jouw gevatte commentaar en je steun zijn onmisbaar in mijn leven. Het is fijn dat jij na 3 jaar in mijn leven bent gekomen!

Hele speciale dank gaat natuurlijk uit naar mijn ouders die mij normen, waarden en wijze lessen hebben meegegeven, welke ik in mijn leven probeer na te streven.

Lieve mama, het beste dat mij is overkomen is dat jij en papa mij hebben geadopteerd. Jouw karakter als sterke, onafhankelijke vrouw en de devotie die je hebt voor je werk zijn een groot voorbeeld voor mij. Dank je wel voor alle liefde die je geeft en de trots die je uit. Lieve papa Henk, wat fijn dat jij besloot om deel uit te maken van mijn (en Rouwan's) leven. Op medisch wetenschappelijk gebied hebben we een connectie maar ook op persoonlijk vlak klikte het direct. Dank je wel, ik had me geen betere 'nep-papa' kunnen wensen.

Er zijn geen woorden die mijn dankbaarheid naar mijn vader hadden kunnen uitdrukken. Mijn lieve papa, zoals The Beach Boys zo mooi verwoorden in een van hun songteksten: 'God only knows what I'd be without you'!

Tot slot, mijn lieve Paul, mijn liefde voor jou is letterlijk grensverleggend; tijdens mijn promotieperiode begonnen we samen in de regio Amsterdam, vervolgens een tijdje apart (alleen qua locatie) en uiteindelijk weer samen in Milano en nu in Bruxelles. De uitdaging die jij mij biedt verrijkt mijn leven, jouw kritische maar stimulerende visie leert mij veel en de liefde die jij voor mij hebt maakt mij gelukkig. Daarnaast waardeer ik het enorm dat mijn ambities geen bedreiging maar een stimulans voor jou zijn.

Voor alles dat is en is geweest; dank je wel, grazie mille, merci beaucoup!

