TESI DOCTORAL

Doctorat de Medicina

Departament de Medicina. Universitat Autònoma de Barcelona

Management of Early Spondyloarthritis:

from diagnosis to treatment in clinical practice

Alumne: Anna Moltó Revilla

DIRECTORS DE TESI:

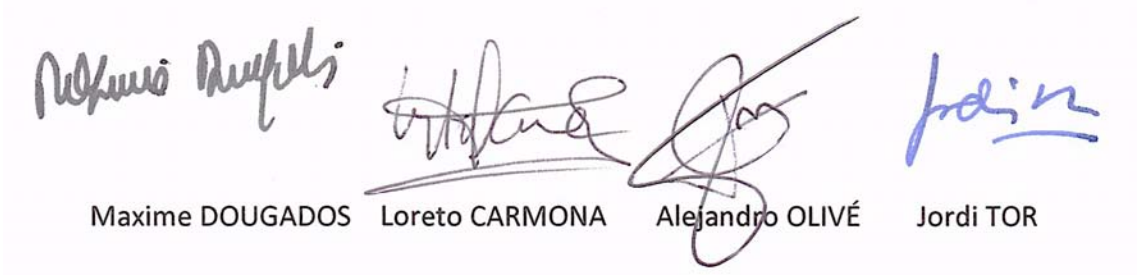
Maxime Dougados MD PhD, *Rheumatology Department, Cochin Hospital Paris, France. Professor de Reumatologia de l'Université René Descartes, Paris, France.*

Loreto Carmona MD PhD, *Instituto de Salud Musculoesquelética Madrid.*

Alejandro Olivé Marquès MD PhD, *Servei de Reumatologia, Hospital Germans Trias i Pujol, Badalona. Professor associat del Departament de Medicina de l'Universitat Autònoma de Barcelona.*

TUTOR DE TESI: Dr Jordi Tor MD, PhD, *Servei de Medicina Interna, Hospital Germans Trias i Pujol, Badalona.Catedràtic del Departament de Medicina de l'Universitat Autònoma de Barcelona.*

Vist i plau:



CONTENT SUMMARY

TABLES SUMMARY

FIGURES SUMMARY

ABBREVIATIONS

AS	*Ankylosing Spondylitis*
ASAS	*Assessment in SpondyloArthritis international Society*
ASDAS	*Ankylosing Spondylitis Disease Activity Score*
AxSpA	*axial SpA*
BASDAI	*Bath Ankylosing Spondylitis Disease Activity Index*
BASFI	*Bath Ankylosing Spondylitis Functional Index*
BASMI	*Bath Ankylosing Spondylitis Metrology Index*
BMI	*Body Mass Index*
CBP	*Chronic Back Pain*
CI	*Confidence Interval*
CRP	*C-reactive protein*
ESR	*Erythrocyte Sedimentation Rate*
ESSG	*European Spondyloarthropathies Study Group*
HAQ-AS	*Health Assessment Questionnaire Ankylosing Spondylitis*
HLA	*Human Leukocyte Antigen*
HR	*Hazard Ratio*
IBD	*Inflammatory Bowel Disease*
IBDRA	*Inflammatory Bowel Disease Related Arthritis*
IL	*Interleukin*
LR+	*Positive Likelihood Ratio*
mNY	*modified New York*
MRI	*Magnetic Resonance Imaging*
mSASSS	*modified Stoke Ankylosing Spondylitis Spine Score*
NS	*Not Significant*
NSAID	*Non-Steroidal Anti-Inflammatory Drugs*
OMERACT	*Outcome Measures in Rheumatoid Arthritis*
OR	*Odds Ratio*
PsA	*Psoriatic Arthritis*
RA	*Rheumatoid Arthritis*
RCT	*Randomized Controlled Trials*

Se	*Sensitivity*
SF36	*Short-From 36*
SIJ	*Sacroiliac Joints*
SpA	*Spondyloarthritis*
Spe	*Specificity*
STIR	*Short Tau Inversion Recovery*
TNF	*Tumor Necrosis Factor*
US	Undifferentiated Spondyloarthritis

INTRODUCTION

Spondyloarthritis: from a concept to a disease

The word spondyloarthritis (SpA) derives from the Greek *spondylo-* (vertebrae) *arthr-* (joint) and *–itis* (inflammation). It encompasses inflammatory rheumatic diseases affecting mainly the axial skeleton. However, other extra-axial manifestations can be seen, namely enthesitic and peripheral articular and extra-articular symptoms, such as psoriasis, uveitis, or inflammatory bowel disease.(1)

Moll and colleagues(2)were, in 1974, the first authors to coin the spondyloarthritis "concept", highlighting a number of signs and symptoms that clustered in diseases that were at that time separated entities: psoriatic arthritis, arthritis related to inflammatory bowel disease, reactive arthritis, undifferentiated spondyloarthritis, and ankylosing spondylitis (AS) —the prototype of SpA. Indeed, the shared characteristics are that they can appear in the same individual (simultaneously or at different time-points) and also in a family member (e.g. a patient can be diagnosed both from psoriasis and his father from Crohn's disease); also, the different clinical manifestations (e.g. psoriasis, eye involvement) are identical regardless the diagnosis.(3)

The same group of authors suggested later that all these diseases shared a common genetic background, by confirming the significantly higher prevalence of Human Leukocyte Antigen (HLA)-B27 in patients suffering from these diseases, especially in those with axial involvement, where the allele is present in 75-90% patients with AS(4,5)(Table 1)

Table 1: Prevalence of HLAB27 in Ankylosing Spondylitis and related diseases

	N° in group	% with HLAB27	Reference
Controls		7-10	
Ankylosing Spondylitis	75	96	Brewerton and others (1973)
	40	83	Schlosstein and others (1973)
Reactive arthritis			
• Peripheral	23	65	Brewerton and others (1973)
• Spinal	10	100	
• Peripheral	19	95	Morris and others (1974)
• Spinal	6	100	
• Peripheral	19	53	McClusky and others (1974)
• Spinal	11	82	
Inflammatory bowel disease			
• Peripheral	8	12.5	Brewerton and others (1974)
• Spinal	20	65	
• Peripheral	14	0	Bluestone and others (1974)
• Spinal	12	67	

Adapted from Lambert JR, et al. Ann Rheum Dis. 1976. (4)

This hypothesis was further confirmed by Hammer et al(6)who demonstrated thatHLA-B27 transgenic rats spontaneously developed inflammatory disease involving the gastrointestinal tract, peripheral and vertebral joints, male genital tract, skin, nails, and heart, strikingly resembling to the clinical manifestations seen in humans suffering from SpA. The exact role of HLAB27 in the SpA pathogenesis has not been clearly established yet, but one of the hypothesis is that, in a predisposing genetic environment (i.e. the presence of HLAB27), an infectious agent might trigger the disease; this hypothesis was supported by findings of the same team(7)on the absence of gut and joint manifestations occurring in transgenic HLAB27 rats when placed in a germfree environment. Recently, other genes have been identified as potential additional risk factors for SpA development (e.g. Interleukin (IL)-23 receptor gene, and the endoplasmic reticulum aminopeptidase 1).(8)

Epidemiology of spondyloarthritis

The prevalence of SpA has been estimated in 0.5–2% in Europe, which is higher than that of rheumatoid arthritis(9); a systematic literature review(10)estimated the prevalence of AS, psoriatic arthritis and SpA related to inflammatory bowel disease (IBD) to be of 0.23%, 0.25% and 0.08%(11–13), respectively. All these figures correspond to estimates in the general population, but some studies have reported SpA prevalence to be as high as 40% in a young (<45 years) population with chronic back pain.(14)

Diagnosis and Classification of spondyloarthritis

No validated diagnostic criteria are available for SpA, but several classification criteria have been proposed. Classification of SpA historically relied on the combination of clinical symptoms plus unequivocal radiographic sacroiliitis according to the modified New York (mNY) criteria presented in 1984 (Tables 2 and 3).(15)However, to fulfil these criteria, patients have to present structural damage (e.g. radiographic sacroiliitis of the sacroiliac joints (SIJ), which can appear several years after disease onset, adding to a mean diagnostic delay of 9 years(16)in patients with clinical symptoms but without such structural damage. Furthermore, these criteria included only axial symptoms, and patients presenting with peripheral symptoms could not be classified as suffering from AS in absence of structural damage.

Table 2: Grading of radiographic sacroiliitis

Grade 0	Normal
Grade 1	Suspicious changes
Grade 2	Minimal abnormality – small localized areas with erosions or sclerosis, without alteration of the joint width
Grade 3	Unequivocal abnormality-moderate or advanced sacroiliitis with one or more of: erosions, evidence of sclerosis, widening, narrowing or partial ankylosis
Grade 4	Severe abnormality-total ankylosis

Adapted from Bennett PH, et al(17)

Table 3: The modified New York Criteria for Ankylosing Spondylitis

Clinical criteria	Low back pain and stiffness for more than 3 months which improves with exercise, but is not relieved by rest Limitation of motion of the lumbar spine bot in the sagittal and frontal planes Limitation of chest expansion relative to normal values correlated for age and sex
Radiological criteria	Sacroiliitis grade ≥ 2 bilaterally or grade 3 or 4 unilaterally
Definite Ankylosing Spondylitis: if the radiological criterion is associated with at least 1 clinical criterion.	

Adapted from van der Linden S, et al. (15)

In order to prevent this diagnostic delay, as well as to incorporate the different clinical presentations of SpA (e.g. peripheral arthritis, uveitis, enthesitis, etc...), others sets of classification criteria were proposed. In the early 1990's Amor and colleagues presented the Amor criteria(18) that included for the first time peripheral features, good response to non-steroidal anti-inflammatory drugs (NSAID), and excluded from mandatory radiological sacroiliitis for the first time, although kept it weighted to a great extent (Table 4).

However, Amor criteria were not broadly accepted in the European SpA community, and a year later, Dougados et al. proposed the European Spondyloarthropathies Study Group (ESSG) set of criteria(19), that also allowed the fulfilment of the criteria without structural damage of the sacroiliac joints and in the absence of axial symptoms (Table 5).

Table 4: The Amor set of criteria for Spondyloarthritis

		Score
Clinical Symptoms/History	Pain at night (spine) or morning stiffness	1
	Asymmetric oligoarthritis	2
	Gluteal (buttock) pain: any or alternating gluteal pain	 1 2
	Sausage-like digit or toe (dactylitis)	2
	Enthesitis (heel)	2
	Uveitis	2
	Urethritis/Cervicitis within 1 month before onset of arthritis	1
	Diarrhoea within 1 month before onset of arthritis	1
	Psoriasis, balanitis or inflammatory bowel disease	2
X-rays	Sacroiliitis grade 2 bilaterally or 3-4 unilaterally	3
Genetic background	HLAB27 positive or positive family history of ankylosing spondylitis, reactive arthritis, uveitis, psoriasis or inflammatory bowel disease	2
Good response to NSAIDs	NSAIDs show a good response within 48 hours, or relapse within 48 hours after NSAID are stopped	2
At least 6 points are necessary to fulfil the Amor criteria		

Adapted from Amor B, et al(18).

Table 5: The European Spondyloarthropathies Study Group set of criteria

Inflammatory back pain **or** **Synovitis (asymmetric or predominantly in the lower limbs)**

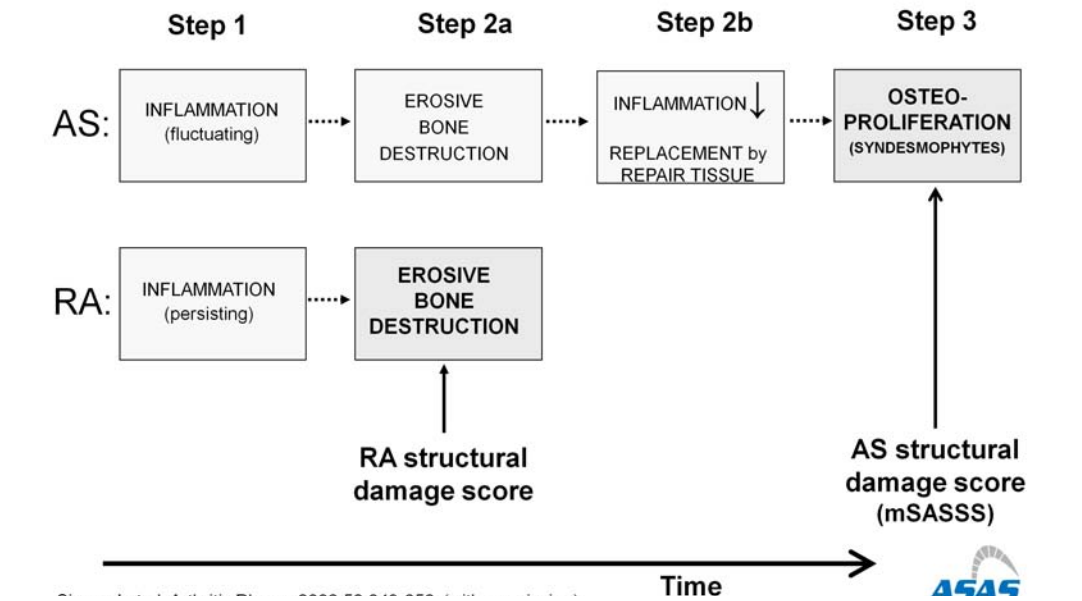
plus one of the following:

- Enthesitis (heel)
- Positive family history
- Psoriasis
- Crohn's disease or ulcerative colitis
- Urethritis/cervicitis or acute diarrhoea within one month before arthritis
- Buttock pain (alternating between right and left gluteal areas)
- Sacroiliitis

Adapted from Dougados et al. (19)

In the late 90's a new imaging modality, magnetic resonance imaging (MRI), allowed for the first time to assess the presence of inflammation in the SIJ and spine. This inflammation could be seen in patients with structural damage (i.e. radiographic sacroiliitis and syndesmophytes), but also in patients without such damage. These findings lead to the idea that inflammation could be the first step in the sequence that would eventually lead to radiographic progression and definitive AS (Figure1).

Figure 1: Proposed sequence of structural damage in ankylosing spondylitis vs. rheumatoid arthritis.



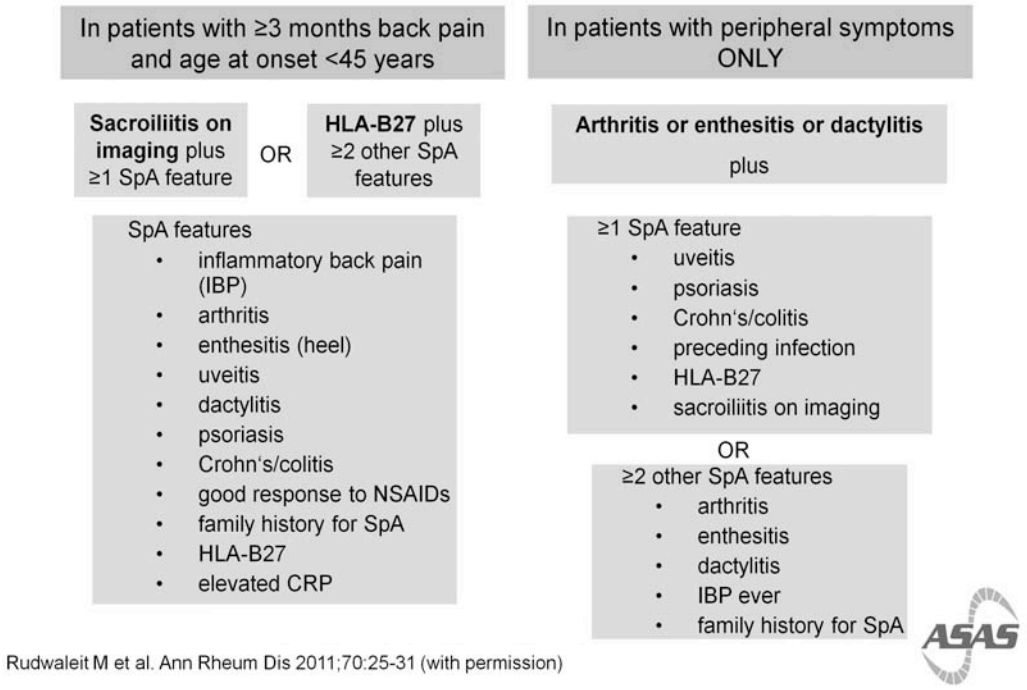
Adapted from Sieper et al. (20)

In 2004, an international group of experts, the Assessment of SpondyloArthritis international Society (ASAS), decided to revise the classification criteria for SpA, to permit an earlier diagnosis, by including MRI findings and an abnormal C-reactive protein (CRP)(i.e. in the presence of CBP absence of other causes that might explain such abnormality) in a set of criteria for the first time (Figure 2). This approach led to the publication in 2009 of the ASAS classification for SpA(20–22) both for axial and peripheral presentations.

If we focus on the axial forms (left side of Figure 2) one patient can fulfil the criteria either by the presence of imaging abnormalities of the SIJ, e.g. radiographic or MRI sacroiliitis (defined in Table 6) and the presence of at least another SpA feature (i.e. one would this patient would fulfil the “imaging” arm of the ASAS criteria), or by the presence of HLAB27 along with at least other two SpA features (i.e. one would this patient would fulfil the “clinical” arm of the ASAS).

In parallel to these new criteria, and reflecting the widespread use of MRI, the concepts of radiographic and non-radiographic SpA forms appeared: former AS were now referred to as radiographic axial SpA (e.g. patients fulfilling the imaging arm due to radiographic sacroiliitis), whereas all other patients fulfilling the ASAS criteria (MRI-imaging arm and clinical arm) encompassed the non-radiographic axial SpA. This nomenclature appeared when non-radiographic forms were believed to correspond to early forms that would all eventually evolve to radiographic forms (Figure 3), leading therefore to the idea that non-radiographic forms were less severe and supposed a lower burden of disease to the patients. However, several studies are starting to report that not all patients with non-radiographic SpA progressed to radiographic SpA after several years(23), suggesting that maybe the non-radiographic and radiographic states are not part of a continuum but only different manifestations of a single disease, as it is the case in rheumatoid arthritis (RA), where that patients may present erosive or non-erosive forms.

Figure 2: The ASAS classification criteria for SpA (axial and peripheral)



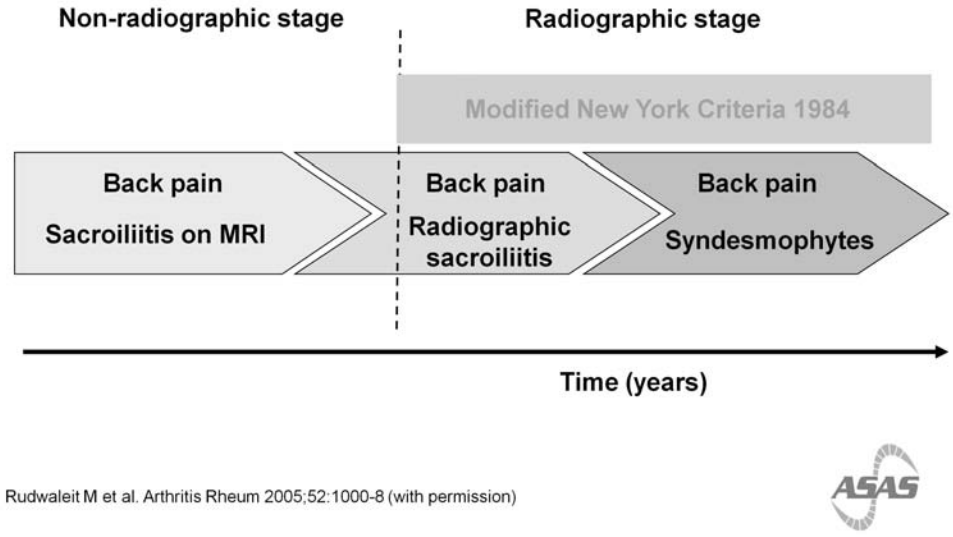
Figures 1 and 2 have been reprinted from the ASAS website(24)

Table 6: Definition of MRI sacroiliitis

Definition of sacroiliitis highly suggestive of SpA (''positive MRI") for application in the ASAS classification criteria
A. Types of findings required for definition of sacroiliitis by MRI
• Active inflammatory lesions of the SI joints (reflecting active sacroiliitis) are required for the definition of ''sacroiliitis on MRI'' as one of the two imaging items in the ASAS classification criteria for axial SpA • BME (STIR) or osteitis (T1 post-gadolinium) highly suggestive of SpA must be clearly present and located in the typical anatomical areas (subchondral or periarticular bone marrow). • The sole presence of other active inflammatory lesions such as synovitis, enthesitis or capsulitis without concomitant BME/osteitis is not sufficient for the definition of sacroiliitis on MRI • Structural lesions such as fat deposition, sclerosis, erosions or bony ankylosis are likely to reflect previous inflammation. At this time, however, the consensus group felt that the sole presence of structural lesions without concomitant BME/osteitis does not suffice for the fulfilment of sacroiliitis on MRI in the ASAS classification criteria for axial SpA.
B. Amount of signal required
• If there is only one signal (lesion) per MRI slice suggesting active inflammation, the lesion should be present on at least two consecutive slices. If there is more than one signal (lesion) on a single slice, one slice may be sufficient

Adapted from Rudwaleit et al. (25)

Figure 3: Non-radiographic and radiographic stages in axial SpA(26)



The ASAS criteria are now the most commonly used criteria, and are widely accepted. However, their performance has been scarcely tested in populations different to the original one.(27) Moreover, as previously exposed, the ASAS criteria for axial SpA can be fulfilled either by the "imaging" arm or by the "clinical" arm and yet the performance of both ASAS criteria's arms has not been compared.

Treatment of Spondyloarthritis

By reducing the diagnostic delay in SpA, rheumatologists aim to lessen the burden of the disease, for the patient, but potentially also at the society level. Since the disease onset usually occurs at a young age, in the patient's most productive years, SpA can have an important socioeconomic impact. Lower employment rates in SpA patients compared to the general population have been reported along with increased SpA-

related work disability.(28)High disease activity and loss of physical function are the most important factors associated with the total costs of SpA.(29)

In this sense, an early recognition and diagnosis of SpA should allow starting an effective treatment promptly, and this may positively alter the course of the disease in these patients.

Non-steroidal anti-inflammatory drugs (NSAID) remain the cornerstone in the treatment of SpA (mainly axial), allowing a symptomatic effect but also by reducing structural damage.(30)Nevertheless, the major advances in the past decades in SpA treatment have been the TNFα blockers.(1,31)

TNFα, an inflammatory cytokine, has been reported to play a major role in the pathophysiology of Spondyloarthritis: several genetic associations of SpA with TNF signalling pathways have been identified (e.g. TNFR1)(32); murine models overexpressing a particular TNF (e.g. TNFdeltaARE mice) present with a destructive polyarthritis , gut inflammation and enthesitis(33) whereas other murine models expressing only the transmembrane TNF (e.g. tmTNF tg mice) do not develop systemic inflammation but both axial and peripheral ankylosis through bone formation.(34) However, the strongest evidence for a key role of TNFα in SpA pathophysiology comes from the in vivo inhibition of TNFα in SpA patients(9) where TNFα blockade leads to a quick reduction of axial inflammatory symptoms and signs(35–45)

According to the 2010 ASAS recommendations for initiation of TNFα blocker in Spa, this treatment should be initiated in patients with definite diagnosis of SpA according to the ASAS criteria for axial SpA (axSpA), with an active disease (i.e. Bath Disease activity Index(46) (BASDAI)≥ 4/10)despite at least 2 NSAID over a 4 weeks period in

total.(47)Notwithstanding these two criteria a positive expert opinion based on parameters, such as a positive MRI, elevated CRP, radiological progression or clinical examination is also needed.

However, we do know that a gap exists between recommendations and their implementation in clinical practice: potentially neither all patients with an indication will be prescribed a TNFα blocker nor all patients in whom a TNFα blocker is prescribed will fully observe these recommendations.

TNFα blocker response has been well evaluated and established for patients meeting NY criteria in many randomized controlled trials (RCT)(38–41,43), but only scarce data(37,42,43) is available for patients without radiological sacroiliitis. Furthermore, inclusion criteria in RCT are very strict, and patients included in these trials may potentially differ from the patients that actually receive the treatment in real practice, with time-changing co-medications, and comorbidities. In this sense, data from observational studies are necessary to evaluate the treatment effect of TNFα blocker in conditions of daily-practice.

The following have been identified in RCT as factors associated to a response to TNFα blocker: age, disease duration and baseline disease characteristics, such as an elevated CRP, BASDAI, Bath Ankylosing Spondylitis Functional Index (BASFI), and human leukocyte antigen (HLA)-B27 positivity.(35,48–50)However, no study has yet identified the phenotype of the patients that would benefit most likely from a TNFα blocker treatment in real practice. This information seems important, especially in view of both the impact on patients' lives and the potential side effects and financial burden these agents bring along.

HYPOTHESIS

This thesis is thus justified in the view of the problems arising and gaps identified in two aspect of the axSpA: 1) the validation of the early diagnosis and classification, and 2) treatment effectiveness in real life conditions.

Diagnosis and Classification of Spondyloarthritis

First, and as previously exposed, the ASAS criteria and in particular its arms (i.e. "clinical" and "imaging") have been validated in few populations besides the one used for its conception, and no data was available concerning their performance (sensitivity, specificity and positive likelihood ratio) in comparison to other sets of criteria. Furthermore, these set of classification criteria are often used in clinical practice as diagnostic criteria by the rheumatologists, and it seemed important to evaluate whether the performances of these criteria for diagnostic purposes were acceptable.

Our hypothesis was that the metric performance of the ASAS criteria both for classification and diagnosis in real life would be adequate, but that the ASAS criteria's arms might present different values. We also anticipated that the ASAS criteria would perform, in terms of likelihood ratio, at least as well as other sets of criteria. We also anticipated a differential weight of individual items of the ASAS criteria in the diagnosis and classification of the patients

Secondly, only sparse data is available concerning the validity of the "clinical" arm. The clinical arm (i.e. the arm of the criteria where a patient can be classified in the absence of imaging abnormalities and even in the absence of raised acute phase reactants) is not well recognised by neither our health authorities (e.g. in some

countries only patients with objective signs of structural damage of the SIJ can receive a TNFα blocker treatment) nor by some physicians (e.g. many are concerned by the possibility to classify HLAB27+ fibromyalgia patients when applying this arm of the ASAS criteria). However, no data has been published comparing the phenotype and disease activity or severity features of patients fulfilling each arm of the ASAS criteria. Our study hypothesis was that the clinical characteristics of the disease in early SpA might be different depending on the arm of the ASAS criteria the patient is fulfilling.

Treatment of early Spondyloarthritis

Regarding treatment in real life, we hypothesised that not all early axSpA patients receiving TNFα blockers would fulfil the ASAS recommendations for initiating such treatment. However, and based on our clinical experience, we anticipated that the treatment effect of such therapy would be comparable to the effect reported in RCT. Furthermore, as suggested also in RCT, we anticipated that some clinical features in early SpA might predict the response to TNFα blockers (e.g. the presence of imaging abnormalities of the SIJ, or the fulfilment of the "imaging" arm of the ASAS criteria).

OBJECTIVES

Primary objectives:

- To analyse the performance of the ASAS criteria in terms of sensitivity (Se), specificity (Spe), and positive likelihood ratio (LR+) to diagnose and to classify SpA in a clinical setting; the performance will be evaluated in total and by arms—namely imaging and clinical arms.
- To compare the phenotype of the patients fulfilling the "imaging" and "clinical" arms of the ASAS criteria for axSpA in an early axSpA population.
- To evaluate the use of TNFα blockers in early axSpA, by:
    - Estimating the frequency of use of TNFα blockers in an early axSpA population in real life,
    - Evaluating the effectiveness (e.g. efficacy in real life) of TNFα blockers in early axSpA, and
    - Identifying the factors associated with response to TNFα blockers in early axSpA.

Secondary objectives

- To identify the item of the ASAS criteria that contribute the most to the diagnosis of SpA
- To evaluate and compare the performance of the ASAS diagnostic and classification criteria to that of other SpA criteria

- To evaluate the performance of the ASAS criteria for SpA in all SpA sub-groups (AS, psoriatic arthritis, reactive arthritis, inflammatory bowel disease related arthritis, undifferentiated spondyloarthritis)
- To evaluate the presence of imaging abnormalities different than those described in the ASAS classification criteria for axSpA in the different subgroups of axSpA patients (e.g. imaging+MRI+Xray- or clinical+CRP-, etc.…)

METHODS

Although the methods of each study included in this thesis are explained in detail in each of the manuscripts that conform this thesis, general aspects of the methods employed and definitions are presented here.

General aspects

The first manuscript included in this thesis aimed to evaluate the performance of the ASAS criteria compared to the other sets of criteria in a chronic back pain setting, while the second part aims to validate the ASAS criteria in an early SpA setting, and compare the phenotype of the patients fulfilling the different arms of the ASAS criteria. For this first part, we used the DECLIC study data, and for the second part, the data from the DESIR cohort.

Finally the third part of the thesis evaluates the treatment effect of TNFα blockers in a daily-practice setting, and for this the data from the DESIR cohort was used.

Patients

The DECLIC study

For the first part of this thesis (ARTICLE 1, page 31), we analysed the patients included in this cross-sectional observational study, performed in 2010. To be included, patients had to present with chronic back pain (CBP) (≥ 3 months) initiating before the age of 45, in a daily-practice outpatient rheumatology setting. In order to be able to evaluate the ASAS criteria optimally, only patients who initiated their CBP after 1995 were included, since MRI was only broadly used in the CBP diagnosis after that date.

Furthermore, patients were consecutively included, regardless of the reason for the consult and regardless the CBP diagnosis, as soon as they fulfilled the inclusion criteria. One thousand rheumatologists were randomly selected from the national database, of which 700 agreed to participate and 384 actively participated in the study. It was anticipated that each rheumatologist would include 4 patients during the 4 months of the study, and in the end, 1379 patients were included. Data on the patients' characteristics as well as all the items permitting the calculation of all the SpA criteria were collected.

The DESIR cohort

For the second and third part of the thesis (ARTICLES 2 and 3, pages 52 and 70) we used the data from the DESIR cohort. DESIR is the acronym *for DEvenir des Spondyloarthrites Indifférenciées Récentes* which is the French wording for outcome of early undifferentiated Spondyloarthritis. This French multicentre cohort was initiated in 2008, and inclusions ended in 2010.

To be included, patients had to aged > 18 but <50 years, present with chronic inflammatory back pain (IBP) according to the Calin or Berlin criteria (Table 7) for > 3 months but less than 3 years, and with a confidence in the SpA diagnosis according to the rheumatologist >5 in a 0 to 10 scale. Seven hundred and eight patients were included. Patients were seen every 6 months for the first 2 years and yearly follow-up is planned for 10 years (follow-up is still ongoing). Data on demographics, disease characteristics, activity, severity, comorbidities and medico-economics, is collected at each study visit. Imaging (Xrays and MRI for the spine and SIJ) and blood testing are performed every 2 and 5 years.

Table 7: Inflammatory back pain according to the Calin and Berlin criteria

Calin et al	Berlin
• Age at onset <40yrs • Duration of back pain > 3 months • Insidious onset • Morning stiffness • Improvement with exercise	• Morning stiffness > 30 min • Improvement with exercise, not with rest • Awakening at 2nd half of the night because of pain • Pain at night (with improvement upon getting up)
Fulfilled if 4/5 items are present	**Fulfilled if 2/4 items are present**

Adapted from Calin et al. (51)and Rudwaleit et al. (52)

Evaluating the metric performance of a set of criteria

In our first study, we evaluated the diagnostic and classification performance of the ASAS criteria. For this, we used the positive likelihood ratio (LR+) and its 95% confidence interval (95%CI). The LR+ is calculated as:

		Gold Standard	
		+	**-**
Test	**+**	a	b
	-	c	d

Sensitivity/1-Specificity, where Sensitivity = a/a+c and Specificity = d/b+d.

The LR+ is the probability of testing positive in patients with the disease/probability of testing positive in patients without the disease. An LR+ greater than 1 indicates that the test result is associated with the disease; however, an LR+ close to 1 have little practical implications, as the post-test probability and pre-test probability will not

differ much. It has been classically proposed to consider a good diagnostic test if its LR+ reaches 10 or more.(53)

Evaluating treatment effect

Dealing with non-randomized longitudinal data

In a longitudinal observational studies, patients receive the treatment according to their treating physicians; therefore it is likely that patients with different characteristics will receive different treatments: e.g. if a treatment "A" is considered to be more aggressive or leading to more side effects than a treatment "B", it is more likely that physicians would rather prescribe "B" in older patients, leading to differences in age in the "A" and "B" treatment groups. This phenomenon is called the prescription bias, reflecting that in observational studies subjects are not randomly assigned in the "treatment" and "non-treatment" groups. For this, it is difficult to estimate an unbiased treatment effect due to these many variables unequally distributed among "treated" and "untreated" groups of patients. Some methodological solutions have been proposed to address this bias, e.g. the propensity score technique (i.e. the probability of a patient of being treated given his covariates). These propensity scores can be used as an adjustment variable in the final model assessing treatment response, but also to match treated and untreated patients according to their propensity score in order to obtain "comparable" groups, creating some sort of "pseudo-randomization".(54,55)

There are at least 2 different techniques to develop a propensity score: either by including all available variables or only those associated to the outcome. These techniques have not been compared so far.

For our third study, we constructed a propensity score for being treated by TNFα blockers and we matched patients who received and not received a TNFα blocker according to the quartiles of this score. Furthermore, in the supplementary (not published) analyses, we compared the two techniques to construct the propensity score.

Evaluating treatment response

For our third analysis, we aimed to evaluate treatment response in early axSpA. Several endpoints could have been used, but the most broadly used is the ASAS40 response,(56,57) which is a binary state (i.e. a patient either is ASAS40 responder or not) and is defined by an improvement of ≥40% and ≥2 units (in a 0-10 scale) in at least 3 out of these 4 domains: patient global, patient pain, function (e.g. Bath Ankylosing Spondylitis Function Index (BASFI)(58)) and inflammation (e.g. mean of BASDAI questions 5 and 6), without worsening in the remaining domains. Since all RCT are currently using this outcome, and since a 40% improvement seems clinically relevant, we decided to use this outcome to evaluate treatment response.

Statistical procedures:

Throughout the dissertation, statistical significance was established for p-values < 0.05, except for the interaction subgroup analysis in ARTICLE 3 (page 70), where the threshold was established at<0.10, in order to increase the power of this analysis. All statistical analysis were performed with the SAS software v.9, except for the propensity score variable selection comparison in the supplementary analysis of ARTICLE 3 (page 92), that was performed with the free software R-CRAN v3.1.1.

Descriptive analysis:

Throughout the dissertation, descriptive analysis of the variables were presented as mean± Standard deviation (SD) or median and interquartile range (IQR) or absolute number and percentage, for continuous and categorical variables, as appropriate.

Univariable analysis:

Continuous variables were tested by T-test in case of normal distribution of the variable or by Wilcoxon test, as appropriate. Identically, categorical variables were tested by Chi2 or Fisher test, as appropriate.

Multivariable analysis:

Only variables with p<0.10 were in the univariable analysis were selected in the multivariable analysis. The quality of the models was assessed by the Area Under the Curve (AUC), and for the nested propensity score models by the Likelihood Ratio test.

ARTICLE 1: PERFORMANCES OF THE ASAS AXIAL SPONDYLOARTHRITIS CRITERIA FOR DIAGNOSIS AND CLASSIFICATION PURPOSES IN PATIENTS VISITING A RHEUMATOLOGIST BECAUSE OF CHRONIC BACK PAIN: THE DECLIC STUDY.

A. Moltó, S. Paternotte, D. Comet, C. Hacquard-Bouder, M. Rudwaleit, P. Claudepierre, D. Van der Hejide, M. Dougados.

Arthritis Care Res (Hoboken). 2013;65:1472-81.

ABSTRACT

Objectives

To evaluate the performances at diagnosis (sensitivity [Se], specificity [Spe], positive and negative predictive values) and study visit (classification purpose) of the ASAS criteria in axial spondyloarthritis (SpA) in patients visiting their rheumatologist for chronic back pain (CBP). Secondary objectives: identifying the most contributive item of to diagnosis/classification of SpA, evaluating the performances of each arm of the ASAS criteria and the other SpA criteria's performances.

Methods

Multi-centric, cross-sectional. *Patients:* history of CBP before the age of 40 visiting a rheumatologist in France. *Data:* a) items of the different sets of criteria, checking if present at diagnosis or at study visit; b) diagnosis of the rheumatologist at study visit. *Statistical analysis:* description of the population. Rheumatologist's diagnosis was considered as the "gold standard" for the estimation of all psychometric properties.

Results

1210 patients were included for our analysis. At diagnosis, Se 0.76 and Spe 0.94 for ASAS axial criteria and Se 0.87 and Spe 0.92 for classification. LR+ of the ASAS axial criteria was 13.6 for diagnosis and 10.30 for classification. The most contributive item to diagnosis and classification was X-ray sacroiliitis, followed by MRI sacroiliitis for diagnosis and history of uveitis for classification. MRI+ imaging ASAS criteria were more sensitive for diagnosis and classification, but as specific as ASAS clinical criteria.

Conclusion

We confirm the validity of the ASAS criteria in diagnosis and classification, in a clinical rheumatological setting of patients with CBP, with good performances compared to the other axial SpA criteria, and for any of their arms.

BACKGROUND:

Spondyloarthritis (SpA) is a frequent disease in patients under 40 year-old presenting with chronic (more than 3 months), back pain but its real prevalence remains unknown.(59)

Classically, axial involvement in SpA was diagnosed upon the modified New York (mNY) criteria(15), by assessing radiological sacroiliitis, resulting in a diagnostic delay in the patients without structural damage of the sacroiliac joint. In order to both prevent this diagnostic delay and also encompass the different clinical presentations of SpA (e.g. peripheral arthritis), other sets of criteria combining both clinical, biological and radiological features, like the Amor criteria(18) and ESSG (European Spondylarthropathy Study Group) criteria(19) were defined, are fairly used and have been validated.

However, these sets of criteria's performances in early diagnosis are rather poor.(60) In 2004, an international group of experts (ASAS for Assessment of SpondyloArthritis international Society) decided to improve the classification criteria of SpA, to permit an earlier diagnosis, including for the first time the magnetic resonance imaging (MRI). This approach led to the publication in 2009 of the ASAS classification in SpA,(21) especially for axial forms, and the proposition of modification of the Amor and ESSG criteria, by including the potential abnormalities in MRI.(61)

However, ASAS criteria's performance (sensitivity, specificity, positive and negative predictive values) have been only scarcely tested in other populations(27) than the one used to define them(22), and only a systematic study of all consecutive patients

presenting with present or past history of chronic back pain occurring before the age of 45 might permit to estimate the axial ASAS criteria's performance in a daily practice.

The primary objective of our study was to evaluate the psychometric properties at diagnosis (sensitivity, specificity, positive and negative predictive values) and at time of the study visit of the ASAS criteria for axial SpA in patients visiting their rheumatologist for any reason but with a history of chronic back pain. Secondary objectives were: a) to identify the most contributive item of these criteria to the diagnosis of SpA, b) to evaluate and compare ASAS criteria's performances to mNY, Amor, ESSG, modified Amor (mAmor) and modified ESSG (mESSG) criteria's performances, c) to evaluate the ASAS criteria's performances in all SpA sub-groups (ankylosing spondylitis, psoriatic arthritis, reactive arthritis, inflammatory bowel disease related arthritis, undifferenciated spondyloarthritis), d) and finally, as ASAS's criteria can be fulfilled either with only clinical features or with the presence of radiological sacroiliitis (either X-ray or MRI) plus only one clinical feature, we aimed to evaluate the performances of each arm of the ASAS criteria (Clinical, MRI radiological and X-ray radiological).

PATIENTS AND METHODS:

Study design:

Multi-centric and cross-sectional observational study. From a national comprehensive file of all 1834 rheumatologists working in private and mixed (half public half private) practice in France(62), 1000 rheumatologists were randomly selected (cluster sampling method). Finally 384 rheumatologists agreed to participate. The study was conducted

in agreement with the good clinical practice: all participating rheumatologists signed a participation agreement, and all patients gave their written consent to participate.

Patients:

Patients had to be over 18 years old, with past or current history of chronic back pain (≥ 3 months) occurring before the age of 45, but after 1995, visiting a rheumatologist in France, and not participating in a clinical trial. Patients presenting with chronic back pain before 1995 were excluded, as only after that date MRI was widely used in the chronic back pain diagnosis in rheumatology daily practice in France. Each rheumatologist had to include 4 consecutive patients responding to the inclusion criteria, but regardless the reason of visit or the definite diagnosis of chronic back pain.

Data collection:

Patient's questionnaires were completed by the rheumatologist from the available data at the time of visit, between 01/2010 and 05/2010 in three chapters that included: a) Checking the inclusion and exclusion criteria, socio-demographical data (age, gender, height, weight), and date and reason for visit, b) Defining the date of onset of chronic back pain, and whether aetiological diagnosis of chronic back pain was definitively established before the visit, and if yes, to choose one between: vertebral fracture, spondylodiscitis, neoplastic disease, mechanical back disorders, spondyloarthritis or other diagnosis, c) Checking, for the presence of the items of ASAS, ESSG, Amor and mNY criteria, and if yes, at when was this item present: before or at diagnosis or after diagnosis. The rheumatologist was not asked to answer whether the patient fulfilled the criteria or not. The items collected were: family

history (SpA, psoriasis, reactive arthritis, inflammatory bowel disease, uveitis), characteristics of chronic back pain (description of chronic back pain [age at first episode, continuous or chronic back pain], inflammatory back pain [yes/no: if "yes": insidious onset, morning stiffness, not improved with rest, awakening at night because of pain with improvement upon getting up], good response to NSAID within 24-48 h), other axial manifestations (cervical and thoracic manifestations, gluteal pain [any or bilateral, alternating gluteal pain]), peripheral clinical characteristics (arthritis, dactylitis, enthesitis), description of extra-articular involvement (uveitis, psoriasis, non-gonococcal urethritis/cervicitis one month prior to arthritis onset, diarrhoea one month prior to arthritis onset, inflammatory bowel disease), HLA B27 status (positive/negative), C-reactive protein (CRP) status (raised:yes/no), and imaging (plain pelvic X-ray [sacroiliitis yes/no] and sacroiliac MRI [sacroiliitis yes/no]).

Statistical analysis:

Description of population at baseline:

All patients without a definite diagnosis were excluded. Population was divided in SpA and Non-SpA groups according to the physician's definite diagnosis. For all measures, physician's diagnosis was considered the "Gold standard". The Non-SpA group was considered the control group. A descriptive analysis of the population was performed at baseline, comparing the demographical data in both groups by Chi square tests and non-parametric Wilcoxon's test as appropriate. Prevalence of SpA in this population was calculated.

Diagnostic versus classification definition:

For any of the items (e.g. psoriasis), in case of positive answer, the rheumatologist was asked whether this information was available at diagnosis (e.g. the psoriasis had appeared 5 years before the diagnosis of chronic back pain) or only after diagnosis but before the study visit (e.g. psoriasis appearing 5 years after the diagnosis of chronic back pain. As explained above, we did not ask the rheumatologists to confirm whether the patient was fulfilling any of the SpA sets of criteria, but by collecting the presence/absence of all the items composing the criteria, we were able to calculate the percentage of patients fulfilling them. If one criteria was fulfilled (e.g. Amor) because of the presence of the items at diagnosis (e.g. presence at diagnosis of inflammatory back pain, alternate buttock pain, family history of SpA, and good NSAID response) the criteria were considered as diagnostic criteria, for diagnosis purposes. When the fulfillment of one criteria (e.g. Amor) was possible because of the presence of the items both at diagnosis or after diagnosis but prior to study visit (e.g. presence at diagnosis of inflammatory back pain, alternate buttock pain, and good NSAID response at diagnosis, with assymetrical oligoarthritis only 4 years after diagnosis, two years before study visit), the criteria were considered as classification criteria, as the diagnosis was already made, and the fulfillment of this criteria would only help to classify the patient.

Evaluation of the performances of the items of the ASAS criteria

For this purpose, frequencies of every item was calculated both at diagnosis and at visit in SpA and Non-SpA groups, and compared by Chi square tests as appropriate.

We also calculated the positive likelihood ratio (LR+ = Sensitivity/1-Specificity) and its 95% confidence interval (CI) for every item, using the rheumatologist's diagnosis as the "gold-standard". LR+ captures both sensitivity and specificity of a given test parameter in a single figure and it is an indicator of the diagnostic value of the respective test: the higher the LR, the better the diagnostic value of the test.(26)

Evaluation of performances of ASAS criteria:

Psychometric performances for the axial ASAS criteria were calculated using the physician's diagnosis as the "Gold standard" as follows: Sensitivity = number of true positive diagnosis/(number of true positives + number false negatives); Specificity = number of true negatives/(number of true negatives + number of false positives); positive predictive value = number of true positive/ (number of true positives + number of false positives); negative predictive value = number of true negatives/(number of true negatives + number of false negatives); LR+ = Sensitivity/1-Specificity. All performances were calculated at diagnosis (diagnostic criteria) and at time of visit (classification criteria).

Comparing performances of all criteria both for diagnosis and classification.

For this purpose we calculated all performances (as defined above) for all set of criteria: mNY criteria, Amor, mAmor, ESSG and mESSG. For the mAmor and mEESG criteria, sacroiliitis was defined either by X-rays or MRI.

Comparing performances of all criteria in each group of SpA.

For this purpose, prevalence of every type of SpA form was calculated. SpA forms were: ankylosing spondylitis, psoriatic arthritis, reactive arthritis, inflammatory bowel

disease related arthritis and undifferenciated spondyloarthritis. Performances of each subgroup were calculated both at diagnosis and at time of study visit.

Comparison of ASAS criteria depending on the availability of imaging

For this purpose we considered 4 different arms in ASAS's criteria: a) clinical ASAS criteria (HLAB27 + other 2 clinical SpA features), b) Imaging ASAS criteria (X-ray or MRI sacroiliitis + 1 other SpA clinical feature) c) MRI+ ASAS criteria (MRI sacroiliitis + 1 other SpA clinical feature) and d) X-ray+ ASAS criteria (mNY sacroiliitis + criteria 1 other SpA clinical feature). We first described the population of each of the ASAS subgroups, comparing the demographical data in both groups by Chi square tests and non-parametric Wilcoxon's test as appropriate. We estimated all performances and LR+ (95%CI) of each of the arms of ASAS criteria.

Statistical analysis were performed with SAS v.9. p values <0,05 were considered significant.

RESULTS:

Baseline characteristics:

Among the 1379 patients recruited by the rheumatologists, 1364 fulfilled the inclusion criteria, but only 1210 patients had a definite diagnosis (SpA or Non-SpA) at the time of visit, and were therefore included in our analysis. Prevalence of SpA was 35.1% (425 patients), and among non SpA patients (785 patients), diagnosis were: mechanical back disorders in 760 patients (62.8%), vertebral fractures in 13 patients (1.1%), infectious spondylodiscitis in 1 patient (0.1%), neoplastic disease in 1 patient (0.1%) and other diagnosis for 10 patients (0.8%). SpA patients were younger (median age 38

years [y] [Q1-Q3 32-45] vs 44 y [36-51], p<0,001), and thinner (median BMI 23.7 [21.7-26.3] vs 24.2 [22-27], p=0,015) with male predominance in both groups (56% men in SpA group vs 52.2% non SpA group, non significant [NS]).

Concerning the missing data, as expected, the most frequent were: HLAB27 typing and imaging of the sacroiliac joints. HLA-B27 typing was available in 393 patients (92.50%) in the SpA group and in 219 patients (28.90%) in the Non-SpA group, plain pelvic radiography was available for 327 patients (76.9%) in the SpA group and in 381 patients (48.53%) in the Non-SPA group; MRI of the sacroiliac joints was available for 131 patients (30.82%) in the SpA and 62 patients (7.90%) in the Non-SpA group. All missing data was considered as negative for the analysis (e.g. in case of no MRI of the sacroiliac joint, the MRI was considered as negative in the analysis).

Performance of the items of the ASAS criteria for diagnosis and classification purposes.

Table 7 represents the frequencies for every item of the composite ASAS criteria, both at diagnosis and study visit (classification purpose). All items were significantly more frequent in the SpA group, except balanitis and family history of reactive arthritis.

Table 7 and Figure 4 resume the LR+ of every item of the ASAS criteria both at diagnosis and time of study visit (classification purpose); Figure 4 resumes as well as the LR+ of every set of criteria.

The most contributive item to diagnosis was X-ray sacroiliitis, followed by MRI sacroiliitis, HLAB27 positivity and raised CRP. Interestingly enough X-ray sacroiliitis was also the most contributive item to the classification, but was followed by history of anterior uveitis, MRI sacroiliitis, and raised CRP.

ASAS criteria's performances at diagnosis and time of study visit.

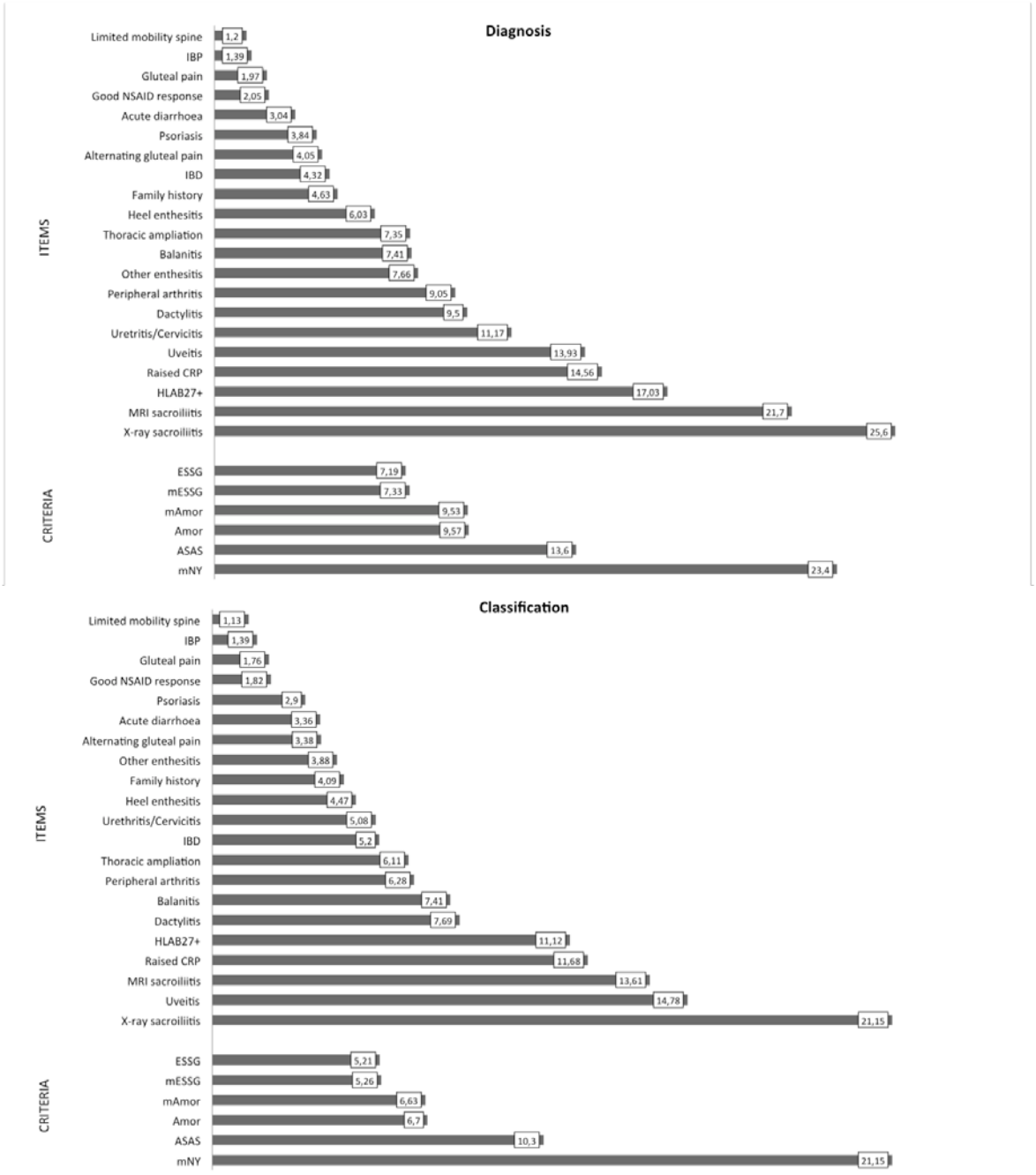
At diagnosis, 324 patients (76.2%) among the SpA group fulfilled the ASAS criteria versus, 368 patients (88.6%) at study visit. LR+ of ASAS criteria both for diagnosis and classification are resumed in Figure 4. Performances of the ASAS criteria for diagnosis and classification are resumed in Figure 5: the ASAS criteria presented high specificity for both purposes, and acceptable sensitivity also for both purposes.

Table 8: Comparison of the frequencies of the different items in SpA vs. Control groups and its Likelihood ratios.

	Diagnosis					Classification				
	SpA	Controls	Chi2	p	LR+	SpA	Controls	Chi2	p	LR+
IBP occurring before 40y	373 (88.6)[a]	492 (63.8)	84.0	0.0	1.39 [1.30 – 1.48][c]	373 (88.6)	492 (63.8)	84.0	<0.001	1.39 [1.30 – 1.48]
IBP according to the physician	413 (98.1)	203 (26.2)	564.0	0.0	3.74 [3.32 – 4.21]	413 (98.1)	203 (26.2)	564.0	<0.001	3.74 [3.32 –4.21]
Limited mobility of spine	251 (65.5)	374 (54.8)	11.6	0.001	1.20 [1.08 – 1.32]	292 (68.9)	477 (60.8%)	7.8	0.005	1.13 [1.04–1.23]
Diminished thoracic ampliation	109 (27.7)	29 (3.8)	142.5	<0.001	7.35 [4.97 –10.86]	139 (32.8)	42 (5,4%)	162.2	<0.001	6.11 [4.42 – 8.45]
NSAID good response	339 (86.7)	296 (42.3)	203.4	<0.001	2.05 [1.87 – 2.25]	370 (87.7)	374 (48.1)	182.2	<0.001	1.82 [1.68 – 1.98]
Gluteal pain	332 (86.9)	307 (44.1)	187.2	<0.001	1.97 [1.80 – 2.16]	369 (88.1)	391 (50.1)	169	<0.001	1.76 [1.62 – 1.90]
Alternating gluteal pain	258 (67.2)	123 (16.6)	289.0	<0.001	4.05 [3.40 – 4.83]	297 (70.2)	162 (20.8)	284	<0.001	3.38 [2.91 – 3.93]
Peripheral arthritis	79 (20.0)	17 (2.2)	109.1	<0.001	9.05 [5.43 –15.06]	109 (25.7)	32 (4.1)	124.4	<0.001	6.28 [4.31 – 9.15]
Dactylitis	35 (8.5)	7 (0.9)	46.0	<0.001	9.50 [4.26 – 21.20]	50 (11.8)	12 (1.5)	59.3	<0.001	7.69 [4.14 – 14.27]
Heel enthesitis	110 (29.6)	37 (4.9)	133.5	<0.001	6.03 [4.24 – 8.56]	169 (38.2)	67 (8.6)	157.6	<0.001	4.47 [3.45 –5.79]
Other enthesitis	42 (10.9)	11 (1.5)	51.3	<0.001	7.66 [3.99–14.70]	75 (17.9)	36 (4.6)	57.4	<0.001	3.88 [2.66 –5.68]
Psoriasis	62 (15)	30 (3.9)	46.0	<0.001	3.84 [2.52 – 5.84]	74 (17.4)	47 (6.0)	39.9	<0.001	2.90 [2.06 –4.11]
Balanitis	4 (1.0)	1 (0.1)	4.46	0.054	7.41 [0.83 – 66.09]	4 (1.0)	1 (0.1)	4.5	0.054	7.41 [0.83–66.09]
IBD	23 (5.5)	10 (1.3)	18.3	<0.001	4.32 [2.08 – 8.99]	31 (7.3)	11 (1.4)	28.5	<0.001	5.20 [2.64–10.24]
Anterior uveitis	36 (8.9)	5 (0.6)	54.5	<0.001	13.92 [5.51 – 35.20]	56 (13.2)	7 (0.9)	84.3	<0.001	14.78 [6.80–32.13]
Urethritis/cervicitis[b]	6 (1.4)	1 (0,1)	8.0	<0.001	11.17 [1.35 – 92.48]	11 (2.6)	4 (0.5)	9.7	<0.001	5.08 [1.63–15.86]
Acute diarrea[b]	13 (3.1)	8 (1.0)	6.9	<0.001	3.04 [1.27 –7.28]	20 (4.7)	11 (1.4)	12.1	<0.001	3.36 [1.63–6.94]
Family history of AS	86 (21.8)	22 (2.8)	112.7	<0.001	7.71 [4.90–12.12]	116 (27.4)	27 (3.5)	150	<0.001	7.92 [5.30–11.84]
Family history of Psoriasis	61(15.1)	35(4.6)	38.7	<0.001	3.29 [2.21 – 4.90]	80 (19.0)	56 (7.2)	38.1	<0.001	2.65 [1.92–3.65]
Family history of Anterior uveitis	10 (2.4)	2 (0.3)	12.7	0.001	9.41 [2.07 – 42.74]	15 (3.6)	3 (0.4)	18.8	<0.001	9.31 [2.71–31.98]
Family history of ReA	6 (1.4)	5 (0.6)	1.9	0.207	2.23 [0.68 –7.26]	6 (1.4)	5 (0.8)	1.9	0.207	2.23 [0.68–7.26]
Family history of IBD	13 (3.2)	4 (0.5)	13.3	0.0003	6.14 [2.01–18.71]	24 (5.7)	6 (0.8)	27.1	<0.001	7.38 [3.04–17.91]
Family history of any of the above	138 (32.6)	55 (7.0)	133.4	<0.001	4.63 [3.47–6.19]	182 (42.8)	82 (10.5)	168.8	<0.001	4.09 [3.24–5.16]
HLAB27 positive	220 (60.1)	27 (3.5)	464.3	<0.001	17.03 [11.65 – 24.90]	289 (68.0)	48 (6.1)	525.5	<0.001	11.12 [8.39–14.74]
Raised CRP	161 (41.3)	22 (2.8)	290.0	<0.001	14.56 [9.49–22.35]	196 (46.1)	31 (4.0)	321.7	<0.001	11.68 [8.15–16.74]
Radiological sacroiliitis	190 (49.2)	15 (1.9)	398.7	<0.001	25.60 [15.35–42.68]	229 (53.9)	20 (2.6)	444.6	<0.001	21.15 [13.60-32.89]
MRI sacroiliitis	96 (25.2)	9 (1,2)	178.7	<0.001	21.70 [11.08–42.49]	140 (32.9)	19 (2.4)	225.0	<0.001	13.61 [8.55–21.66]

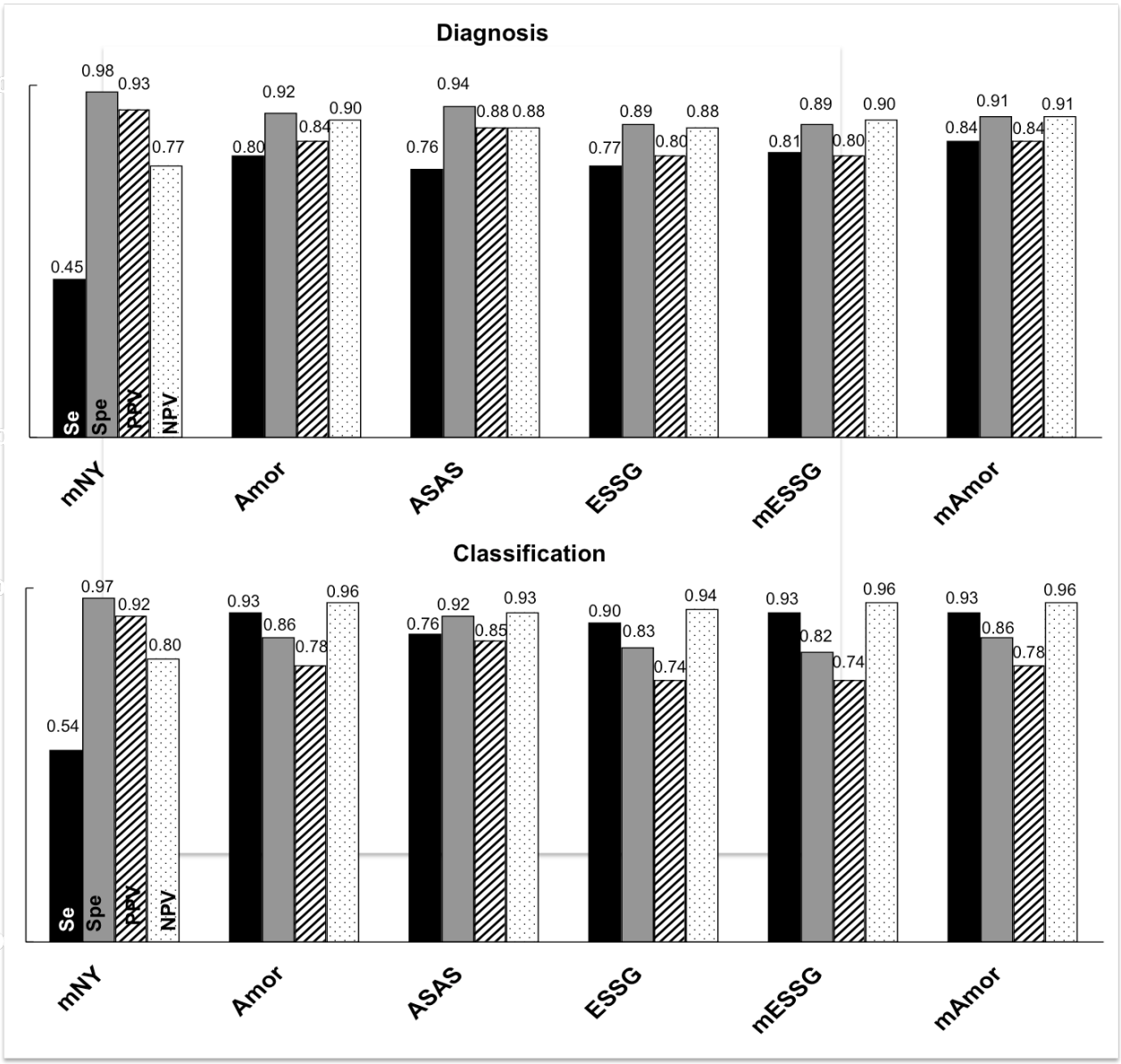
Abreviations: IBP: inflammatory back pain, NSAID: non-steroidal anti-inflammatory drugs, IBD: inflammatory bowel disease, AS: Ankylosing spondylitis, ReA: reactive arthritis, CRP: C-reactive protein. a: N(%) .b: one month prior to diagnosis; c:[CI] ; d:%

Figure 4: Positive likelihood ratio for every item of the ASAS criteria and for all set of criteria (ASAS, Amor, mAmor, ESSG, mESSG and mNY) for diagnosis and classification purposes.



Abreviations: IBP: inflammatory back pain, IBD: inflammatory bowel disease, CRP: C-reactive protein; ASAS: Assessment of SpondyloArthritis international Society; mNY: modified New York criteria; ESSG: European Spondyloarthritis Society Group; mAmor: modified Amor criteria; mESSG: modified ESSG criteria, SpA: spondyloarthritis.

Figure 5: Comparison of the performances of the different set of criteria both for diagnosis and classification.



Abbreviations: Se: Sensitivity, Spe: Specificity, PPV: positive predictive value, NPV: negative predictive value, mNY: modified New York criteria, ESSG: European Spondyloarthritis Society Group, mAmor: modified Amor criteria, mESSG: modified ESSG criteria, SpA: spondyloarthritis

Comparison of criteria performances

At diagnosis, 365 patients (83.8%) fulfilled the mAmor criteria, 345 (81.2%) the mESSG, 342 (80.5%) the Amor, 327 (76.9%) the ESSG and 190 patients (44.7%) the mNY criteria. At study visit 395 (92.9%) fulfilled the mAmor and mESSG criteria, 388 (91.3%) the Amor criteria, 381 (90.0%) the ESSG and 229 patients (53.9%) the mNY criteria.

The performances of the different set of criteria are represented in Figure 5. The mAmor criteria showed the highest sensitivity in both diagnostic and classification purposes, followed by mESSG, ESSG and ASAS. The mNY criteria presented the higher specificity, both in diagnosis and classification, followed by ASAS criteria.

Figure 4 resumes also the LR+ of all sets of criteria: mNY criteria had the highest LR+ both for diagnosis and classification purposes, followed by ASAS also for both purposes. Interestingly enough, only those two sets of criteria showed LR+ above 10 for any of the purposes.

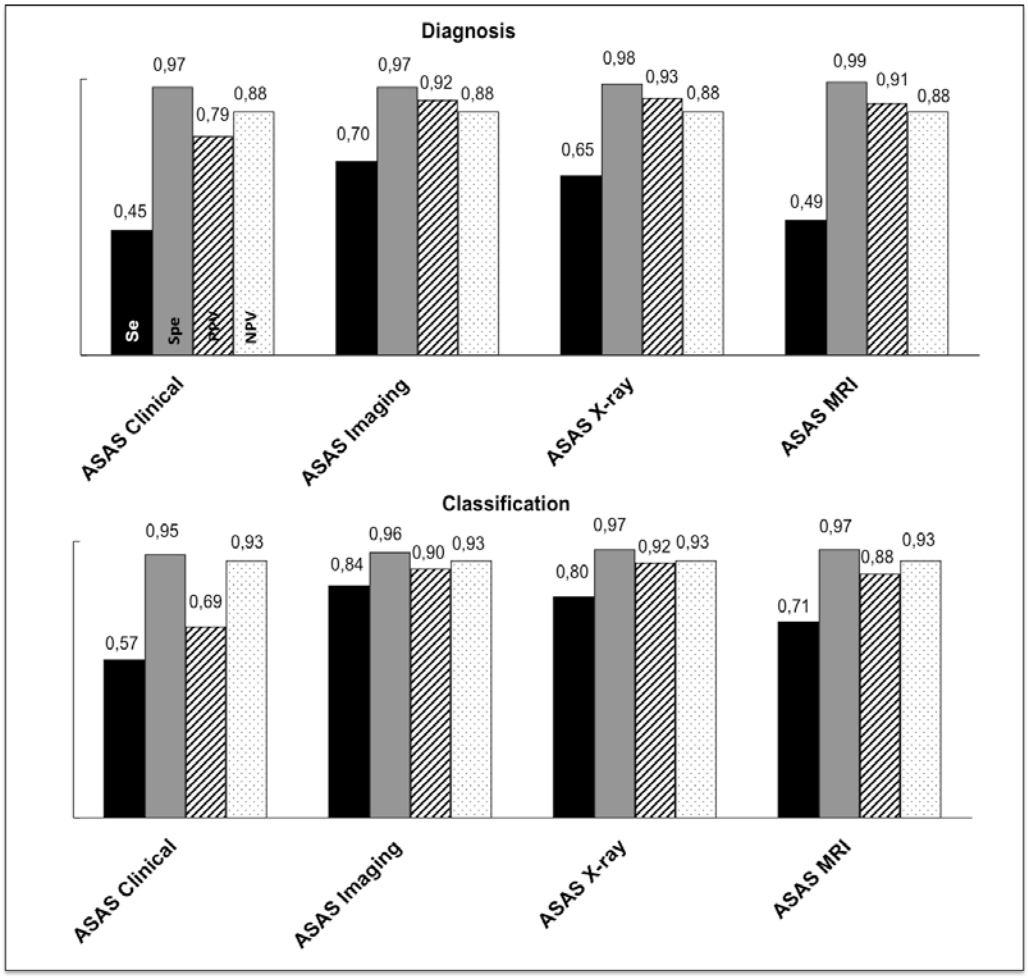
Comparison of the performances of the arms of the ASAS criteria:

For the diagnosis purposes, 107 patients (25.2%) fulfilled the clinical ASAS criteria, 56 patients (13.2%) fulfilled the ASAS MRI+ criteria, and 156 patients (36.7%) with X-ray+ ASAS criteria. Patients fulfilling ASAS clinical criteria at diagnosis were older (median 41y [36-51]) compared to any of the imaging arms (patients from the ASAS MRI+ criteria had a median age of 35y [28-40] and ASAS X-ray+ criteria had a median age of 38y [32-46]), but younger than control patients (median age of 43y [36-51]), and no differences were assessed for gender or BMI.

For classification purposes, 111 patients (26.1%) fulfilled the ASAS clinical criteria, 74 patients (17.4%) fulfilled the ASAS MRI+ criteria, and 164 patients (38.6%) the X-ray+ ASAS criteria. Identically, patients fulfilling ASAS clinical criteria were also older (median 41y [33-49]) compared to any of the imaging arms (patients fulfilling the ASAS MRI+ criteria had a median age of 37y [29-42] and those fulfilling the ASAS X-Ray+ criteria had a median age of 38y [32-46]), but younger than control patients (median age of 43y [36-51]), and no differences were assessed for gender or BMI.

Performances of clinical and imaging arms of the ASAS criteria are represented in Figure 6. Imaging ASAS criteria were more sensitive both for diagnosis and classification purposes, but were nearly as specific as ASAS clinical criteria also in both purposes.

Figure 6: Performance of the clinical and radiological arms of the ASAS criteria for diagnosis and classification.

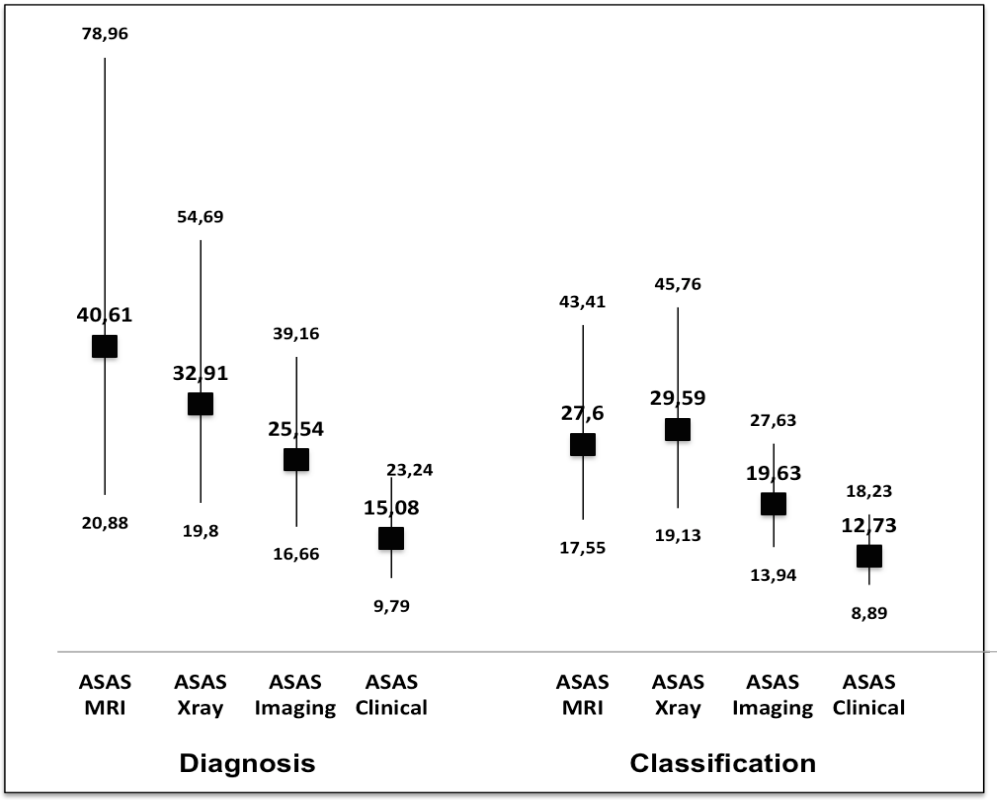


ASAS clinical arm: HLAB27 + 2 SpA features. ASAS imaging arm: X-ray sacroiliitis or MRI sacroiliitis + 1 SpA feature. ASAS-MRI: MRI sacroiliitis + 1 SpA feature. ASAS X-ray: X-ray sacroiliitis + 1 SpA feature

Abbreviations: ASAS: Assessment of SpondyloArthritis international Society; Se: Sensitivity; Spe: Specificity; PPV: positive predictive value; NPV: negative predictive value.

Figure 7 resumes also the likelihood ratios of every arm of the ASAS criteria: for diagnosis, ASAS MRI+ criteria had the highest likelihood ratio, followed by the ASAS X-ray+ criteria; interestingly, for classification, ASAS X-ray+ criteria had the highest likelihood ratio followed by ASAS MRI+ criteria. However, likelihood ratio was above 10 for any of the ASAS criteria arms.

Figure 7: Positive likelihood ratio (and its confidence interval) for every arm of the ASAS criteria for diagnosis and classification purposes.



Abbreviations: ASAS: Assessment of SpondyloArthritis international Society; MRI: Magnetic Resonance Imaging.

Criteria's performances in all SpA sub-groups:

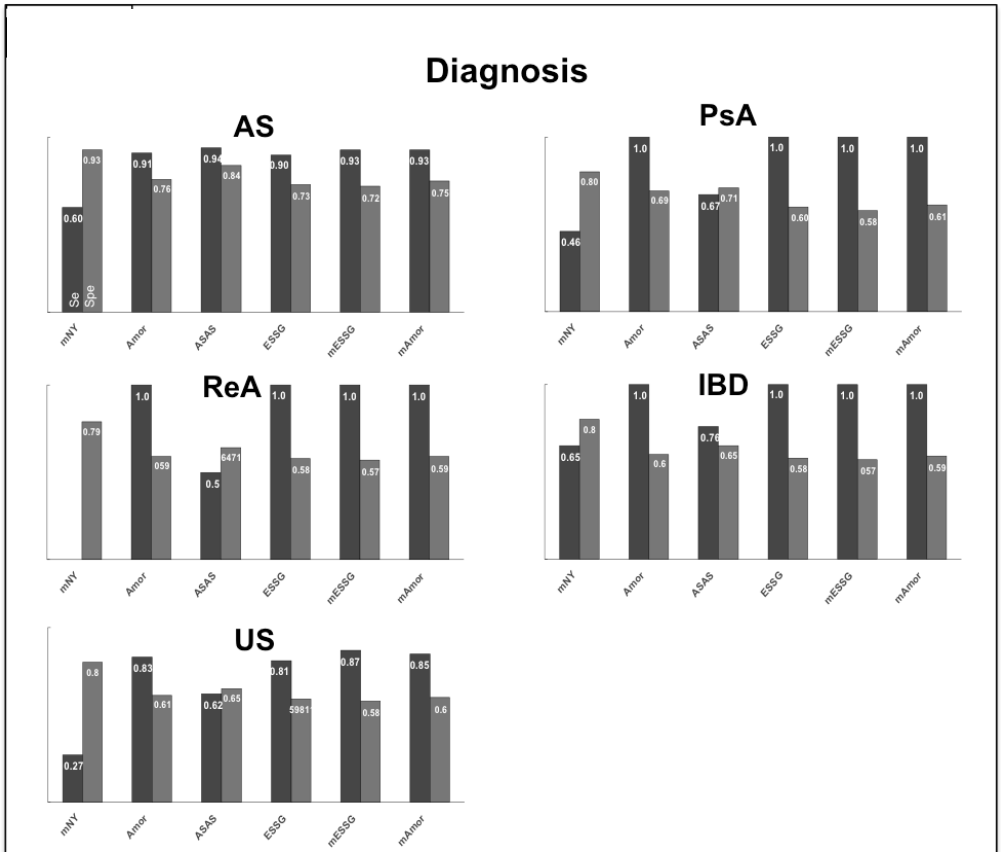
SpA sub-groups were distributed as follows: ankylosing spondylitis in 304 patients (71.5%), psoriatic arthritis in 48 patients (11.3%), reactive arthritis 4 patients (0.9%), inflammatory bowel disease related arthritis in 17 patients (4.0%), and undifferentiated spondyloarthritis in 52 patients (12.2%).

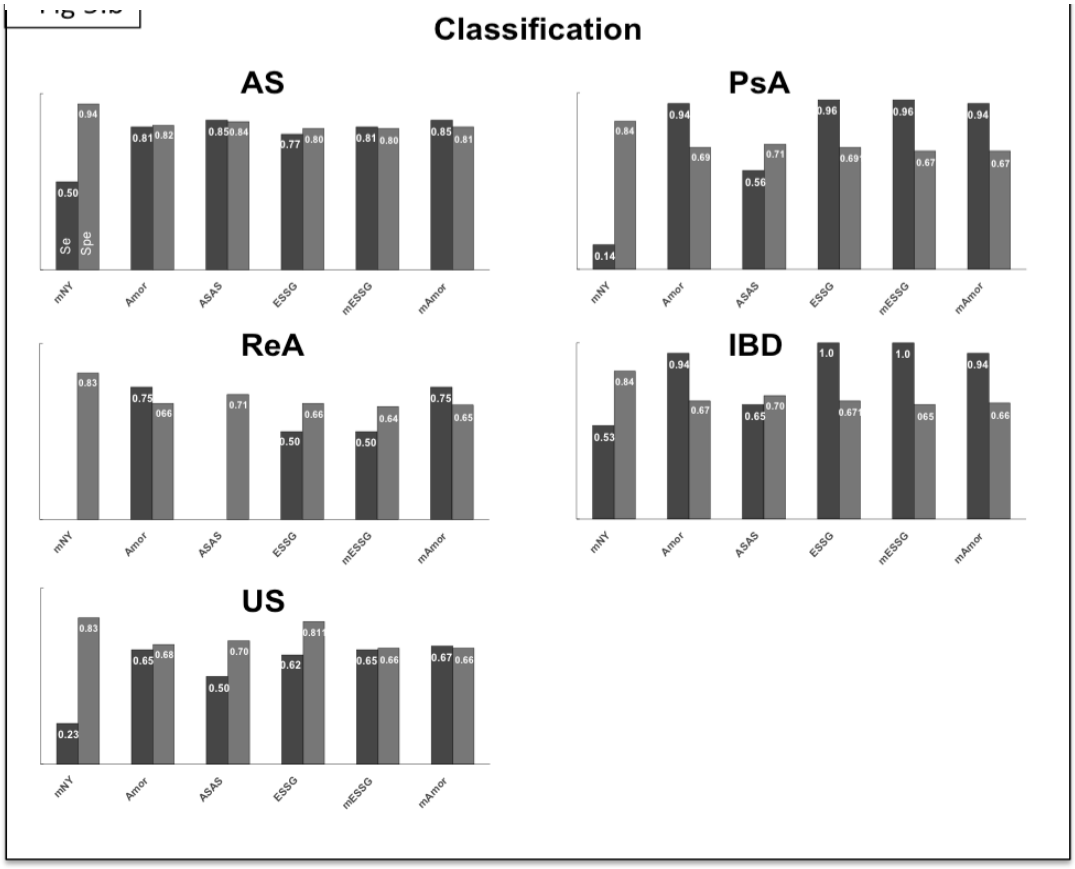
Performances of all criteria in all SpA forms are described in Figure 8. ASAS's criteria had the highest sensitivity in ankylosing spondylitis both for diagnosis and classification, but a rather low sensitivity for the other forms of SpA. As expected, the mNY criteria were the most specific for ankylosing spondylitis diagnosis and

classification, but also in all other subgroups of SpA, followed by the ASAS criteria also in all forms.

ESSG and mESSG's criteria had the highest sensitivity for psoriatic arthritis in both diagnostic and classification purposes, and Amor and mAmor's criteria were the most sensitive for both diagnosis and classification of reactive arthritis and inflammatory bowel disease related arthritis.

Figure 8: Performances of the different sets of criteria (ASAS, Amor, mAmor, ESSG, mESSG and mNY) both for diagnosis and classification in the subgroups of SpA.





Abbreviations: AS: ankylosing spondylitis; PsA: psoriatic arthritis; IBDRA: inflammatory bowel disease related arthritis; ReA: reactive arthritis; US: undifferentiated spondyloarthritis; ASAS: Assessment of SpondyloArthritis international Society; mNY: modified New York criteria; ESSG: European Spondyloarthritis Society Group; mAmor: modified Amor criteria; mESSG: modified ESSG criteria; Se: Sensitivity; Spe: Specificity; PPV: positive predictive value; NPV: negative predictive value.

DISCUSSION:

This study confirms that ASAS criteria are highly specific and have acceptable sensitivity both for diagnosing or classifying SpA patients, with similar performances as described in the original validation study. Tests with very high LR+ (over 10) are considered diagnostic(53) and in this study, the LR+ of the ASAS criteria were higher than 10 for any purpose. Moreover, this study suggests that when analyzed separately, both ASAS criteria arms have high specificity, and good sensitivity, (slightly higher for the imaging arm, and especially for classification); regardless the arm of the ASAS criteria, LR+ was always above 10.

When testing the other SpA criteria's performances, lower sensitivities and higher specificities in the diagnostic purpose for all criteria, and higher sensitivities with similar specificities in the classification purpose compared to the original cohorts that tested those criteria.(15,18,19) mNY were the set of criteria with higher LR+ for any purpose, followed by ASAS, and those two set were the only with LR+ values above 10.

This study has some weaknesses but also some strengths. The main weakness is the retrospective design of the study for the diagnostic purpose, but in the other hand in one same study we were able to collect the necessary data to assess the performances of the criteria both for diagnosis and classification purposes in a single visit. Other weaknesses of this study are the potential bias in the selection of patients, due to the study: participating rheumatologists being aware that our center was participating were more prone to recruit SpA patients, despite well specified in the protocol that patients had to be included consecutively regardless the diagnosis of CBP. This might explain the high prevalence of SpA in this study compared to classical SpA

prevalence.(60,63)Another possible bias might be due to the patient selection, with less established disease than in other studies as by including patients diagnosed only after 1995 we limited the disease duration at 15 years. This might explain the differences in the performances of the different set of criteria, especially the low sensitivity rates, especially in the diagnosis purpose, compared to other trials. However, this reflects the clinical practice regarding MRI prescription in France for the diagnosis of CBP. Finally, regarding imaging modality, no central reading was performed, but here again, this reflects the clinical daily practice.

The study has also some strengths, as the sample size or the patient's recruitment technique that, by recruiting consecutive patient, ensures representativity of the population.

Furthermore, this study provides some information: some items of the ASAS criteria presented with LR+ above 10, suggesting they might be diagnostic or classificatory by themselves (in a CBP under 40years-old population); some of them were expected, like MRI and X-ray sacroiliitis or HLAB27 positive, but others were not, like raised CRP for diagnosis, and raised CRP and anterior uveitis for classification. This confirms the rationale of having added the raised CRP in the ASAS criteria, and the clinical impression that uveitis allows the classification of the disease even after diagnosis of SpA.

In conclusion, we confirm the validity of the ASAS criteria both in diagnosis and classification, in a clinical rheumatological setting of young patients with CBP, for any of its arms, suggesting that those criteria might help not only classifying but also guiding the rheumatologist in diagnosing these patients.

ARTICLE 2: EVALUATION OF THE VALIDITY OF THE DIFFERENT ARMS OF THE ASAS SET OF CRITERIA FOR AXIAL SPONDYLOARTHRITISAND DESCRIPTION OF THE DIFFERENT IMAGING ABNORMALITIES SUGGESTIVE OF SPONDYLOARTHRITIS. DATA FROM THE DESIR COHORT.

A. Moltó, S. Paternotte, D. van der Heijde, P. Claudepierre, M.Rudwaleit, M. Dougados.

Ann Rheum Dis. 2014 Jan 3. doi: 10.1136/annrheumdis-2013-204262.
[Epub ahead of print]

ABSTRACT

Background

The ASAS criteria for axial Spondyloarthritis (SpA) permit to classify a patient with ("imaging" arm) and without ("clinical" arm) imaging abnormalities of the sacroiliac joints.

Objective

To compare the phenotype of early axial SpA with regard to the two arms of the ASAS axial SpA criteria.

Methods

Demographics, SpA clinical and biological features, disease activity and severity parameters, and imaging abnormalities at the sacroiliac and spine levels were compared, in the two arms of the ASAS axial SpA criteria, in the patients of the French cohort of early SpA.

Results

Of the 615 analysed patients, 435 (70.7%) fulfilled the ASAS criteria (262 (60.2%) and 173 (39.8%) in the imaging and clinical arms, respectively. There were no major differences in the characteristics between both groups except for younger patients, more males and higher CRP values in the "imaging" arm. Other imaging abnormalities than the ones permitting the fulfilment if the "imaging" arm of the ASAS criteria (e.g. X-rays structural damage or MRI inflammatory changes of the SIJ) were observed (MRI-SIJ structural damage (55.0% vs. 3.5%), MRI-Spine inflammatory changes (35.1% vs. 12.9%), MRI-spine structural damage (10.3% vs. 5.3%) and X-ray-syndesmophytes (11.8% vs. 5.3%) in the imaging versus clinical arm, respectively.

Conclusion

Our study confirms the external validity of the clinical arm of the ASAS set of criteria. It is remarkable that many patients in the clinical arm showed other imaging changes in SI joints and spine.

BACKGROUND

The Assessment of SpondyloArthritis international Society (ASAS) has recently proposed a set of criteria (21) aiming to recognize patients suffering from early axial spondyloarthritis (SpA). These sets of criteria can be summarized in two main arms, which are both applied to patients with chronic back pain starting before the age of 45 years:

- The "imaging" arm in which a patient can fulfill the set of criteria if there is demonstration of an objective sign of inflammation (MRI)(64)or of structural damage (conventional pelvic X-ray) in the sacroiliac joints (SIJs) together with past history or current symptoms of at least one feature suggestive of SpA (e.g. inflammatory back pain, psoriasis, enthesitis...)
- The "clinical" arm in which a patient can fulfill the set of criteria despite the lack of demonstration of an objective sign of inflammation at MRI or of structural damage in the SIJs. In this case, the patient has to be HLAB27 positive and has to present with past history or current symptoms of at least two features suggestive of SpA (*e.g*. inflammatory back pain, psoriasis, enthesitis...)

Currently, there is still a debate concerning the validity of these criteria both in terms of validity of the clinical arm and also regarding the imaging abnormalities permitting to classify a patient as fulfilling the imaging arm.

The ASAS criteria for axial SpA (ax-SpA) and especially the clinical arm have been validated in different SpA populations as well as in different clinical trials in terms of external validity (36,37,61,65)(*e.g.* by evaluating the clinical presentation, the level of activity and/or severity of the disease, the treatment effect of drugs usually effective

in radiographic spondyloarthritis…) but also in terms of face validity (*e.g.* by evaluating the percentage of patients with histological features suggestive of sacroiliitis despite the lack of imaging [both by X-rays and MRI] evidence of such sacroiliitis).(66) Nevertheless, the clinical arm is not well recognized by the different national and international health care systems; for example, in many countries patients with active, severe and refractory to NSAIDs ax-SpA are not eligible to TNFα blocker treatment if the imaging investigations do not show any sign of sacroiliitis. Moreover, sometimes an elevated CRP is required.

Another aspect of the ASAS criteria for ax-SpA that is currently debated is the definition of the imaging abnormalities. According to the published ASAS criteria, and in order to fulfil the ASAS "imaging" arm, patients have to present with either obvious structural damage of the SIJs observed at pelvic X-rays (*e.g.* bilateral grade 2-4 or unilateral grade 3-4 according to the modified New York criteria)(15) or active (acute) inflammatory lesions of the SIJs observed at pelvic MRI according to the ASAS/OMERACT definition.(25)

However, there is an increasing body of evidence suggesting that other imaging abnormalities might also be of clinical relevance to classify a patient as suffering from SpA. For example, structural damage of the SIJs might be more easily detected by either Computed Tomography Scan (67) or MRI (68,69), but because of the potential long term risk of radiation exposure(70)MRI is the preferred technique. Another example is related to the fact that when these changes (*e.g.* acute inflammation or structural damage) are observed at the spine level, they might be also of relevance in

the classification of a patient suffering from symptoms suggestive of SpA, in particular when using the MRI technology.(71,72)

These preliminary remarks prompted us to conduct an analysis of the data collected in patients suffering from early inflammatory back pain suggestive of spondyloarthritis and participating at the on-going French multi-centric cohort DESIR (acronym which stands in French for outcome of early undifferentiated spondyloarthritis) with the following two main objectives: a) to compare the patient characteristics with regard to the arm ("imaging" *versus* "clinical" arm) of the ASAS criteria they are fulfilling and b) to describe the prevalence of the different imaging abnormalities in the two arms of the ASAS criteria.

PATIENTS AND METHODS

Study design

DESIR is a French prospective, multi-center, longitudinal cohort aiming to study patients with early inflammatory back pain suggestive of SpA (clinicaltrials.gov NCT01648907).(73)

This study fulfilled the current Good Clinical Practices and has obtained the approval of the appropriate ethical committee. Participants at the study gave their written informed consent. The website contains the detailed description of the centers, the organization of the cohort but also the full detailed protocol and case-report form.(74)

A total of 708 patients with early inflammatory back pain (IBP) have been included (inclusion period October 2007 to April 2010). Consecutive patients aged >18 years and <50 years with IBP involving the thoracic, lumbar spine or buttock area for more than 3 months but less than 3 years and symptoms suggestive of diagnosis for SpA

score ≥5 (on a Numerical Rating Scale of 0–10 where 0=not suggestive and 10=very suggestive of SpA) were included in the DESIR cohort. Patients had to fulfil the inflammatory back pain criteria of Calin *et al.* or Berlin *et* al.(51,52) Patients with a definite diagnosis of non-SpA back pain, conditions that might interfere with the validity of the informed consent and/or prevent an optimal compliance (*e.g*. alcoholism, psychiatric disorders) or a history of TNFα blocker treatment were excluded. For this study, analysis included the whole population of the DESIR cohort, and used the data set locked on December 12th 2011.

Collected data

The collected data comprised both patient demographics and clinical presentation of the disease. Demographics included age, gender, and body mass index (BMI). Moreover, all the items permitting to adequately classify a patient according to the ASAS criteria were collected. The activity of the disease was evaluated using the following: BASDAI (Bath Ankylosing Spondylitis Disease Activity Index) (46), and ASDAS-CRP (Ankylosing Spondylitis Disease Activity Score – C Reactive Protein).(75) The severity of the disease was assessed using the following: BASFI (Bath Ankylosing Spondylitis Functional Index),(58) and BASMI (Back Ankylosing Spondylitis Metrology Index).(76)Finally, the quality of life was evaluated according to the short-form 36 (SF 36). (77)

Concerning the imaging modalities, to ensure the quality and standardization of collected images, a written specific procedure was provided to each participating centre. Conventional X-rays of the cervical spine, lumbar spine and pelvis were performed. Radiologists or rheumatologists at each study centre scored each SIJ as

follows: 0=normal; 1=doubtful, 2=obviously abnormal or 3=fused. For this present analysis, SIJs were considered abnormal if at least one SIJ was scored 2 or 3. This scoring method used for the local reading in DESIR is derived from the modified New York criteria for radiographic sacroiliitis changes(15) with one modification: grade 2 and 3 of New York criteria were pooled together in one combined grade.

The modified Stoke Ankylosing Spondylitis Spine Score (mSASSS)(78) was calculated from the conventional X-rays of the cervical and lumbar spine. Definite radiographic damage was defined as an mSASSS score of ≥2 in at least one vertebral edge of each individual patient, representing the appearance of at least one syndesmophyte in that patient.

MRI scans of the SIJs, upper spine (C2 to T10) and lower spine (T8 to S1) were performed using the short-tau inversion recovery (STIR) and T1 fast spin echo (FSE) acquisitions. A contrast product was not used.

Presence of inflammatory and structural damage at the SIJs and spine were assessed by radiologists or rheumatologists at each study centre. Inflammatory changes of the SIJs were defined by the presence of bone oedema the SIJ. Structural damage of the SIJ was defined by the presence of clear characteristic lesions such as sclerosis, erosions, bone bridges or ankylosis in the SIJs. The spine was evaluated at 3 different levels (cervical/thoracic/lumbar), and the presence of either inflammatory (defined by the presence of bone oedema/with contrast enhancement at the entheseal site at vertebral corners or the whole vertebrae, with/without disc involvement) or structural damage (defined by the presence of sclerosis, erosions or vertebral syndesmophytes) was separately assessed at each of these 3 levels. For each of these MRI evaluations,

radiologists or rheumatologists at each study centre scored as follows: 0=normal, 1=doubtful, 2=abnormal. For this present analysis, MRI was considered abnormal only if scored as "abnormal" by the rheumatologist or radiologist.

DESIR definitions for MRI involvement are similar but not identical to the ASAS/OMERACT definitions for MRI sacroiliitis /MRI spinal involvement in SpA because the DESIR study was designed prior to the publication of the ASAS/OMERACT definitions.(25,79)

Statistical analysis:

The first step of the statistical analysis consisted in the classification of each patient according to the ASAS criteria for axial spondyloarthritis resulting in the following 3 categories: patients fulfilling or not the ASAS criteria, and for those fulfilling the ASAS criteria, whether abnormal imaging findings permitted to classify the patient in the "imaging" arm. If not, patients were classified in the "clinical" arm, if HLAB27 was positive and 2 features suggestive of SpA were present.(21) For this purpose, we excluded the patients for whom missing data did not permit to adequately categorize them into a specific arm of the ASAS criteria for axial SpA.

The second step consisted in the comparison of the patient characteristics according to the arm of the criteria they were fulfilling (*e.g.*"imaging" *versus* "clinical"arm). The categorical variables were compared using Chi-square test (or Fisher exact test as applicable), while continuous variables were compared using non-parametric Wilcoxon test.

Because different scenarios can be observed according to the imaging modalities in the imaging arm and according to the CRP status in the clinical arm, we performed a

descriptive analysis in 5 different subgroups (*e.g. a*) X-ray definite SIJ damage and MRI inflammatory changes of the SIJ, *b*) X-ray definite SIJ damage and MRI SIJ normal, c) X-ray SIJ normal and MRI inflammatory changes of the SIJ, d) X-ray SIJ normal and MRI SIJ normal and CRP abnormal, e) X-ray SIJ normal and MRI SIJ normal and CRP normal [where CRP abnormal was defined as > 6mg/L]).

The third step consisted in the evaluation of other (not included in the ASAS criteria for axial SpA) imaging findings suggestive of SpA (e.g. MRI structural damage of the SIJ, MRI inflammatory and structural damage at the spine level, and the presence of at least 1 syndesmophyte at the cervical or lumbar level) observed in the different arms of the ASAS axial SpA criteria. Thereafter, we estimated the concordance in the abnormal imaging findings observed in the X-rays and MRI modalities using a kappa coefficient of concordance (e.g. at the SIJ level between the pelvic X-ray grades 0-1 [normal or doubtful]/2-3 [abnormal or partially fused] vs. MRI grades 0-1 [normal or doubtful]/2 [abnormal] of structural damage; at the spine level between spine X-ray [mSASSS ≥ 2 in at least one vertebral edge yes/no (cervical or lumbar)] vs. spine MRI grades 0-1 [normal or doubtful]/2 [abnormal] of structural damage in at least one of the 3 levels [cervical/thoracic/lumbar]).

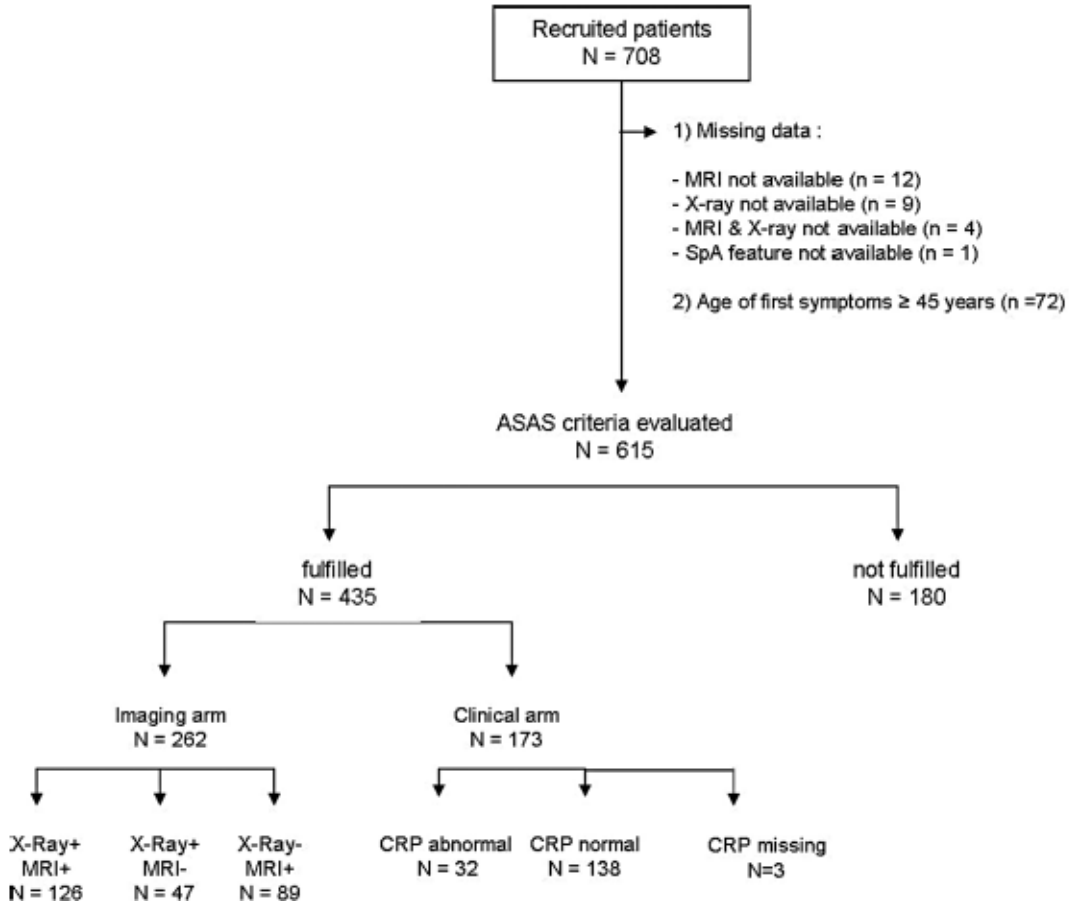
RESULTS

Classification of patients according to the ASAS criteria, and regarding the two arms of the ASAS criteria

Figure 10 summarizes the flowchart of the recruited patients. Because of missing data, the fulfilment (or not) of the ASAS criteria was assessed in 615 patients of the 708 patients included in the DESIR cohort. Of those, 435 patients fulfilled the ASAS criteria

(and 180 did not fulfil the ASAS criteria): 262 and 173 fulfilled the "imaging" and "clinical" arms, respectively. Within the "imaging" arm 126, 47 and 89 patients, belonged to the "X-ray definite SIJ damage and MRI inflammatory changes of the SIJ", "X-ray definite SIJ damage and MRI SIJ normal" and "X-ray SIJ normal and MRI

Figure 9: Distribution of the 708 patients recruited in the DESIR cohort according to the axial ASAS criteria.



inflammatory changes of the SIJ" sub-groups, respectively. Thus, of the patients fulfilling the imaging arm 66.0% could be classified as radiographic ax-SpA and 34.0% as non-radiographic ax-SpA. Within the "clinical" arm (e.g. "X-ray SIJ normal and MRI SIJ normal") 32 (18.5%) and 138 (79.8%) belonged to the "X-ray SIJ normal and MRI SIJ normal and CRP abnormal" and "X-ray SIJ normal, MRI SIJ normal and CRP normal"

sub-groups, respectively. For 3 patients the available data permitted to classify the patient in the clinical arm despite the absence of CRP data; because of this missing data, they were excluded from this subgroup analysis.

The "imaging" versus "clinical" arm

Table 7summarizes the comparison in the patient (age, gender, B27 positivity) and disease (clinical presentation, activity and severity) characteristics fulfilling the two arms of the ASAS criteria for axial SpA. No differences were found between groups except for younger patients, more males and higher CRP values in the "imaging" arm.

Table 7: Comparison of the patients and disease characteristics of early axial spondyloarthritis according to the axial ASAS criteria arm (imaging versus clinical) they are fulfilling.

	Axial ASAS criteria		*p****
	Imaging* arm	**Clinical arm**	
Number	N=262	N=173	
Age (years; mean, ±SD)	30.6 (±7.2)	32.6 (±7.3)	0.005
Female gender, n, (%)	107 (40.8)	101 (58.4)	0.0003
Disease duration in months; mean (±SD)	18.6 (±10.5)	19.2 (±11.2)	0.683
Past history or current symptoms of [n, (%)]			
- Enthesitis,	112 (42.8)	86 (49.7)	0.154
- Peripheral arthritis	56 (41.2)	34 (35.1)	0.344
- Dactylitis	36 (13.7)	20 (11.6)	0.506
- Uveitis	27 (10.3)	12 (6.9)	0.229
- Psoriasis	42 (16.0)	28 (16.2)	0.966
- Inflammatory bowel disease	14 (5.3)	5 (2.9%)	0.221
HLAB27 positivity (n, %)	192 (73.6)	173 (100.0)	<0.0001
Family history of SpA, n (%)	110(44.4)	84 (50.0)	0.257
BASDAI mean (±SD)	41.3 (±20.4)	44.0 (±20.2)	0.169
CRP, mg/L, mean (±SD)	11.6(±15.7)	5.2(±9.3)	<0.0001
Raised CRP** (N=459) (n, %)	111 (44.4)	32 (18.8)	<0.0001
ASDAS-CRP, mean (±SD)	2.6(±1.1)	2.3(±1.0)	0.006
BASFI, mean (±SD) N=473	28.7 (±22.2)	29.3(±22.5)	0.840
BASMI, mean (±SD) N=463	2.4 (±1.0)	2.1(±0.9)	0.020
Mental SF36, mean (±SD)	41.3 (±11.5)	40.5 (±10.8)	0.420
Physical SF36, mean (±SD)	40.9 (±9.0)	39.8 (±9.6)	0.191
Radiological sacroiliitis, n (%)	173 (66.3)	0 (0.0)	<0.0001
MRI acute inflammation of the sacroiliac joint, n (%)	215 (83.7)	0 (0.0)	<0.0001

** imaging = either definite damage of the sacroiliac joints at pelvic X-rays according to the modified New York criteria (see ref. 8) or inflammatory lesion of the sacroiliac joints at MRI as defined in the "methods " section. ** Raised CRP defined as CRP > 6mg/L. *** Statistical significance defined by p<0,05. The categorical variables were compared using Chi-square test (or Fisher exact test when Chi-square test was non applicable), while continuous variables were compared using the Student T-test (or the non-parametric Wilcoxon test when Student test was non applicable).Abbreviations: BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Function Index; BASMI: Bath Ankylosing Spondylitis Metrology Index; SF36: Short Form 36. SpA: Spondyloarthritis.*

The comparison between the 5 sub-groups according to the imaging and/or CRP abnormalities

A descriptive analysis was performed (summarized in Table 8) in the 5 different sub-groups according to the imaging and/or CRP abnormality. Patients from the "X-ray

definite SIJ damage and MRI inflammatory changes of the SIJ" sub-group were younger (29.3years [±6.7]) than the other subgroups. Interestingly, BASDAI was strikingly higher in the "CRP abnormal" subgroup of the "clinical" arm (57.3 [±17.2]), compared to any of the other sub-groups. CRP levels were higher in the "CRP abnormal" sub-group compared to any of the sub-groups of the "imaging" arm.

Concerning the structural damage (presence of structural damage yes/no) of the sacroiliac joints and spine, the concordance between MRI and X-ray findings was very low at both the SIJ (Kappa 0.55 [0.49-0.61]) and spine level (0.18 [0.06-0.31]). (Table 9)

Other imaging abnormalities

Table 10 summarizes the other imaging findings observed in the 5 subgroups as previously described. MRI structural damage of the SIJ were, as expected, more frequently observed in the subgroup of patients with X-ray damage of the SIJ (65.3%), but more interestingly, also in 3.5% patients of the clinical arm. MRI inflammatory changes of the spine were more frequently observed in presence of other markers of inflammation (e.g. local MRI inflammatory changes of the SIJ [38.6%] or CRP abnormality [21.9%]). Furthermore, within the subgroups without X-ray damage of the SIJ, there was evidence of X-ray damage of the spine in as much as 6.7% of the subgroup "X-ray SIJ normal and MRI inflammatory changes of the SIJ" and 9.4% of the subgroup "X-ray SIJ normal and MRI SIJ normal and CRP abnormal".

Table 8: Comparison of the patients and disease characteristics of early axial spondyloarthritis according to the axial ASAS criteria arms ("imaging" *versus* "clinical") and sub-arms (X-rays *versus* CRP) they are fulfilling.

	Axial ASAS criteria				
	Imaging* arm			Clinical arm**	
	X-ray+[a]/ MRI+	X-ray+ [b]/ MRI-	X-ray-[c]/ MRI+	X-ray-/MRI- Abnormal CRP[d]	X-ray-/MRI- Normal CRP
Number	126	47	89	32	138
Age, years, mean (±SD)	29.3 (±6.7)	31.1 (±8.4)	32.3 (±6.8)	31.4 (±6.1)	32.8 (±7.6)
Female gender, n(%)	46 (36.5)	22 (46.8)	39 (43.8)	21 (65.6)	78 (56.5)
Disease duration, months, mean (±SD)	19.1 (±9.9)	19.3 (±11.2)	17.7 (±11.0)	17.1 (±10.0)	19.2 (±10.4)
Past history or current symptoms of : n (%)					
- Enthesitis	45 (35.7)	25(53.2)	42 (47.2)	21 (65.6)	63 (45.7)
- Peripheral arthritis	22 (36.7)	14(46.7)	20 (43.5)	13 (61.9)	20 (27.0)
- Dactylitis	15 (11.9)	8 (17.0)	13 (14.6)	6 (18.8)	14 (10.1)
- Uveitis	15 (11.9)	3 (6.4)	9 (10.1)	5 (15.6)	6 (4.4)
- Psoriasis	16 (12.7)	9 (19.2)	17 (19.1)	5 (15.6)	23 (16.7)
- Bowel disease	9 (7.1)	3 (6.4)	2 (2.3)	1 (3.1)	3 (2.2)
Family history of SpA n (%)	56 (45.9)	18(40.0)	36 (44.4)	16 (50.0)	66 (49.6)
HLAB27 positivity, n (%)	101 (80.2)	29(61.7)	62 (70.5)	32 (100.0)	138 (100.0)
BASDAI mean (±SD)	40.2 (±19.9)	40.7 (±24.0)	43.2 (±19.1)	57.3 (±17.2)	41.5 (±19.5)
CRP mg/L mean (±SD)	10.9 (±13.4)	15.9 (±20.8)	10.5 (±15.7)	15.9 (±17.6)	2.7 (±1.7)
ASDAS-CRP mean (±SD)	2.6 (±1.0)	2.6 (±1.3)	2.6 (±1.1)	3.5 (±0.9)	2.0 (±0.8)
BASFI mean (±SD)	27.4 (±22.5)	31.5 (±23.2)	29.0 (±21.3)	45.0 (±22.8)	26.1 (±21.0)
BASMI mean (±SD)	2.5 (±1.0)	2.5 (±1.0)	2.2 (±0.8)	2.4 (±1.2)	2.1 (±0.8)
Mental SF36, mean (±SD)	41.2 (±11.5)	42.5 (±12.3)	40.7 (±11.0)	39.8 (±11.3)	40.5 (±10.7)
Physical SF36, mean (±SD)	41.9 (±8.4)	40.0 (±10.4)	39.9 (±8.9)	33.5 (±8.3)	41.1 (±9.3)

**imaging = either definite damage of the sacroiliac joints at pelvic X-rays according to the modified New York criteria (see ref. 8) or inflammatory lesion at MRI as defined in the "methods " section. ** clinical= presence of HLA-B27 plus presence of two clinical features of SpA. a: X-ray+MRI+: presence of both structural damage of the sacroiliac joints on pelvic X-rays and MRI inflammatory changes of the sacroiliac joints. b:X-ray+MRI-: presence of structural damage of the sacroiliac joints on pelvic X-rays without MRI inflammatory changes of the sacroiliac joints. c: X-ray-MRI+: presence of MRI inflammatory changes of the sacroiliac joints without structural damage of the sacroiliac joints on pelvic X-rays. d: Abnormal CRP defined as > 6mg/L. Abbreviations: BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Function Index; BASMI: Bath Ankylosing Spondylitis Metrology Index; SF36: Short Form 36. SpA: Spondyloarthritis.*

Table 9: Concordance between a) MRI and X-rays findings concerning the structural/structural damage of the sacroiliac joints and b) between MRI and X-rays findings concerning the structural/structural damage of the spine.

a)

		Structural damage of the sacroiliac joints on conventional pelvic X-rays*	
		YES	**NO**
Structural damage of the sacroiliac joints on pelvic MRI**	**YES**	179	75
	NO	89	515

** Structural damage of the sacroiliac joints on conventional pelvic X-rays defined as grades 2-3 (grades of sacroiliitis in the DESIR cohort are defined as 0=normal; 1= doubtful, 2 = obviously abnormal or 3 = fused): this scoring method used for the local reading in DESIR is derived from the modified New York criteria for radiographic sacroiliitis changes with one modification: grade 2 and 3 of New York criteria were pooled together in one sole grade.*

*** Structural damage of the sacroiliac joints on pelvic MRI: Structural damage of the sacroiliac joints were defined by the definite presence of characteristic lesions such as sclerosis, erosions, bone bridges or ankylosis on the sacroiliac joints. Changes were scored as 0=normal, 1=Doubtful, 2= Abnormal. For this analysis, Structural damage of the sacroiliac joints yes=grade 2. Abbreviations: MRI: Magnetic Resonance Imaging*

b)

		Structural damage of the spine on conventional X-rays*	
		YES	**NO**
Structural damage of the spine on MRI**	**YES**	11	33
	NO	35	532

** Structural damage of the spine on conventional X-rays defined as an mSASSS score of ≥2 (appearance of at least one syndesmophyte) in at least one vertebral edge of each individual patient.*

*** Structural damage of the spine defined as presence of sclerosis, erosions or syndesmophytes on the vertebrae changes by scoring of the MRI scans as normal (grade 0), doubtful (grade 1) and abnormal (grade 2). For this analysis, structural damage of the spine on MRI yes= grade 2. Abbreviations: MRI: Magnetic Resonance Imaging*

Table 10: Spine and sacroiliac joints MRI and X-rays findings (a part from the ones included in the items of the ASAS criteria) in patients suffering from early axial spondyloarthritis.

	Axial ASAS criteria				
	Imaging*			Clinical**	
	X-ray+[a]/ MRI+	X-ray+ [b]/MRI-	X-ray- [c]/MRI+	X-ray-/MRI- Abnormal CRP[d]	X-ray-/MRI- Normal CRP
Number	126	47	89	32	138
MRI Sacroiliac joints structural damage +	92 (73.0%)	21 (50.0%)	31 (34.8%)	3 (9.4%)	3 (2.2%)
MRI Spine inflammatory lesions++	53 (42.7%)	9 (21.4%)	30 (34.1%)	7 (21.9%)	15 (10.9%)
MRI Spine structural damage lesions+++	18 (14.5%)	3 (7.1%)	6 (6.9%)	2 (6.3%)	7 (5.1%)
X-ray spine++++	14 (11.1%)	11 (23.4%)	6 (6.7%)	3 (9.4%)	6 (4.4%)

+ Structural damage of the sacroiliac joints at MRI defined as clear characteristic lesions such as sclerosis, erosions, bone bridges or ankylosis on the sacroiliac joints.

++ Inflammatory changes of the spine defined as bone oedema in or adjacent to the enthesis at the margin of the vertebrae or whole vertebra [with or without disc involvement], compatible with lesions observed in cases of ankylosing spondylitis.

+++ Structural damage of the spine defined as clear characteristic lesions such as sclerosis, erosions or syndesmophytes on the vertebrae.

++++X-ray spine abnormalities defined by an mSASSS with at least one syndesmophyte (Score =2) in at least on vertebral edge (ref 25)

*imaging = either definite damage of the sacroiliac joints at pelvic X-rays according to the modified New York criteria (see ref. 8) or inflammatory lesion of the sacroiliac joints at MRI as defined in the "methods " section.

** clinical= presence of HLA-B27 plus presence of two clinical features of SpA.

a: X-ray+MRI+: presence of both structural damage of the sacroiliac joints on pelvic X-rays and MRI inflammatory changes of the sacroiliac joints.

b:X-ray+MRI-: presence of structural damage of the sacroiliac joints on pelvic X-rays without MRI inflammatory changes of the sacroiliac joints.

c: X-ray-MRI+: presence of MRI inflammatory changes of the sacroiliac joints without structural damage of the sacroiliac joints on pelvic X-rays.

d: Abnormal CRP defined as > 6mg/L.

DISCUSSION

This analysis of the DESIR cohort permitted to adequately evaluate the potential differences in the clinical presentation of the disease with regard to the different arms of the ASAS criteria they are fulfilling, in a population of IBP patients suggestive of SpA. We would like to stress that patients included in the DESIR cohort did not have to fulfil any particular set of criteria, but to present with IBP for more than 3 months and less than 3 years, initiating before the age of 50 and to have a physician's confidence for SpA diagnosis above 50% (e.g. >5 in a 0 to 10 scale, where 0= no SpA and 10= definite diagnosis of SpA). Only after inclusion the different sets of criteria applied based on the data collected during the first visit.

The comparison of the patients between the "imaging" and the "clinical" arms showed that the different clinical features (e.g. peripheral arthritis, dactylitis, uveitis, psoriasis,...) as well as most parameters evaluating the activity (e.g. BASDAI), the severity (e.g. BASFI, BASMI) and the impact of the disease in terms of quality of life (e.g. SF36) were identical between the two groups of patients.

Furthermore, this study confirms the presence of structural damage both at the spine and SIJ level and inflammatory lesions of the spine in a small proportion of patients of the "clinical" arm of the axial ASAS criteria.

This study has some weaknesses but also some strengths. First, because of missing data we were unable to evaluate whether the patients from the DESIR cohort were fulfilling the axial ASAS set of criteria in 26 patients (e.g. 3.7%). This raises the question of handling the missing items of the ASAS criteria, in particular concerning the B27 antigen and the imaging modalities. In some epidemiological studies, such missing

items have been considered as negatives in the evaluation of the ASAS criteria.(80) Because of the main objective of our study, we excluded the patients with missing items that would therefore not permit us to classify the patient according to the different arms of the axial ASAS criteria.

Second, the technique of evaluation of the imaging modalities (e.g. by each local participating investigator and not by a central reader) might be seen as a weakness, but this methodology could also be seen as a strength since it reflects daily practice. Anyway, in the DESIR cohort imaging modalities were standardized both in terms of collection of the imaging (e.g. standardized written protocols) and in terms of the evaluation of the imaging (specific CRF with a reminder of the definition of the abnormalities suggestive of SpA were provided), since the readers (either the rheumatologist and/or the radiologist) had to complete a case-report form as described in the methods section of this manuscript.

Another limitation of the study is that specificity of the ASAS axial criteria of any of its arms can not be evaluated because of the lack of control group: despite a group of patients within the DESIR cohort did not fulfill the ASAS criteria, those patients could not be considered as a control population, as all patients included in the DESIR cohort had to have a physician's confidence for SpA diagnosis above 5 (where 0= no SpA and 10= definite diagnosis of SpA). Another limitation is the cross sectional design of our study, with no gold standard for the diagnosis of SpA.

However, our multicentre study also has some strengths. First, our analyses were performed in a large number of patients (N=682) with IBP suggestive of SpA, ensuring a good representation of an early SpA population from a western European country.

As previously reported, the similarity of the clinical disease manifestations between the patients with regard to the arm of the ASAS criteria they are fulfilling is a strong argument in favour of the validity of such criteria. Yet, there are also differences between the two arms with respect to age, gender, and elevated CRP which may be have relevance for disease progression, for example.

Our findings regarding the prevalence of other imaging abnormalities in the "clinical" arm of the axial ASAS criteria, raises the question of the potential need of revisiting the axial ASAS criteria when conducting clinical epidemiological studies/trials but also the question of the MRI and radiographic investigation of the patients presenting with symptoms suggestive of SpA in daily practice. However, these results have to be carefully interpreted, and will need further validation, as the prevalence of these abnormalities in patients without SpA or in a normal population has not been described so far. Long-term longitudinal evaluation of the patients enrolled in the DESIR cohort and/or other on-going cohorts will permit to confirm or not our findings.

ARTICLE 3: EFFECTIVENESS OF TNF-ALPHA BLOCKERS IN EARLY AXIAL SPONDYLOARTHRITIS: DATA FROM THE DESIR COHORT.

A. Moltó, S. Paternotte, P. Claudepierre, M. Breban, M. Dougados.

Arthritis Rheumatol. 2014 Jul;66(7):1734-44.

ABSTRACT

Objective

To estimate the frequency of use and effectiveness of TNFα blockers in an inflammatory back pain population suggestive of early axial spondyloarthritis (axSpA) in daily practice.

Methods

DESIR is a prospective, multicentre, observational cohort including 708 patients with early (<three years duration) inflammatory back pain suggestive of axSpA. Statistical analysis: The percentage of patients receiving TNFα blockers over the first two years of follow-up was estimated by survival analysis. For the effectiveness evaluation, the outcome (ASAS40 response) was compared in patients with and without TNFα blockers (after a matching procedure based on a propensity score).

Results

Of the 708 patients,30.2% [26.7–33.7] patients received at least one TNFα blocker during the 24 months of follow-up. The percentage of ASAS 40 responders was 62 (31.8%) vs. 26 (13.5%) in the TNFα blocker group vs. usual care groups, respectively (OR=2.99 [1.80–4.99], p=0.0002). This effectiveness was more pronounced in the subgroup of patients with MRI sacroiliitis, with 46% vs. 15%of ASAS40 responders in the TNFα blocker group vs. usual care groups, respectively (OR=4.99 [2.17–11.51]).

Our study shows that TNFα blockers are frequently used in early axSpA in daily practice and confirms the effectiveness of TNFα blockers compared to any other treatment, especially in the subgroup of patients with MRI sacroiliitis.

BACKGROUND

The concept of spondyloarthritis (SpA) as a group of inter-related diseases has been recognized since the early 1970s. The classification and recognition of axial spondyloarthritis (axSpA) relied on the combination of consistent axial skeleton symptoms and unequivocal bilateral grade two or unilateral grade three radiographic sacroiliitis according to the modified New York (mNY) criteria.(15) More recently, the different criteria sets for SpA (e.g. Amor(18), European Spondyloarthropathy Study Group (ESSG)(19) and the ASsessment of SpondyloArthritis International Society (ASAS)(21) criteria) have made it possible to recognize axSpA without chronic X-ray changes to the SIJ.

Tumour necrosis factor alpha (TNFα) blockers have proven to be an effective treatment of SpA, with high levels of research evidence. (81)Many randomized clinical trials (RCTs)(39,40,82)have provided data confirming TNFα blockers' efficacy in axSpA patients meeting the modified NY criteria (e.g. radiographic axial SpA or ankylosing spondylitis (AS) patients), but only limited data are yet available for patients lacking definite radiographic sacroiliitis.(36,37,42)In these and other RCTs, age(48), disease duration(48)and baseline disease characteristics such as elevated C-reactive protein (CRP)(39,48,49), elevated erythrocyte sedimentation rate (ESR)(39,48), presence of enthesitis(50), Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)(46), Bath Ankylosing Spondylitis Functional Index (BASFI)(58), human leukocyte antigen (HLA)-B27 positivity(39) and lack of previous exposure to TNFα blockers(35)have been identified as factors associated with a better response to TNFα blockers.

Furthermore, only very limited data have been published regarding the therapeutic effect of TNFα blockers in axSpA patients in clinical practice.(83)Although RCTs are extremely useful in generating therapeutic response data, the inclusion criteria are very strict, which is not always applicable in daily clinical practice.

Therefore, studies aiming to evaluate the therapeutic response of TNFα blockers in daily practice conditions are required.

Devenir des Spondylarthropathies Indifférenciées Récentes (DESIR) is a prospective longitudinal cohort involving 25 rheumatology centres and 708 patients; its aim is the comprehensive study of the nature and outcome of axSpA from early symptom onset.(74)Evaluating the natural history of the disease in a cohort of patients presenting with recent onset (≤three years) of inflammatory back pain suggestive of axSpA is not an unconventional study. However, the hallmark of this cohort is that patients can receive any treatment at their rheumatologist's discretion(73) and therefore it presents a unique opportunity to evaluate current practice and the effectiveness of TNFα blockers in daily practice.

OBJECTIVES:

Our main objective was to evaluate TNFα blockers in clinical practice in an early axSpA population, by a) estimating the frequency of use of TNFα blockers, b) evaluating their effectiveness and c) exploring the interaction of baseline variables with this effectiveness.

METHODS:

Patients:

A total of 708 patients with early inflammatory back pain were included in the DESIR cohort with scheduled visits every six months. Ten-year follow-up is currently ongoing; our analyses included the first two years of follow-up.

Patients aged between 18 and 50 with IBP involving the thoracic, lumbar spine or buttock area for >three months but <three years and symptoms suggestive of SpA according to the rheumatologists' assessment (score ≥5 on a Numerical Rating Scale of 0–10 where 0=not suggestive and 10=very suggestive of axSpA) were included. Patients were required to fulfil the Calin et al or the Berlin IBP criteria.(51,52) Patients with a definitive diagnosis of non-axSpA back pain, any condition that could affect the validity of the informed consent and/or prevent them achieving optimal compliance (e.g. alcoholism, psychiatric disorders) or a history of previous TNFα blocker use were excluded. All patients (e.g. fulfilling ESSG, Amor or ASAS classification criteria for axial SpA, but also not fulfilling any criteria set) were included in our analysis, except those receiving a biological agent other than TNFα blockers as the first biologic treatment. The data set used for this analysis was locked on December 3, 2012.

Definition of visits:

Patients could be started on TNFα blocker treatment at any time during the follow-up, at their rheumatologist's discretion (e.g. in the interval between the six-monthly scheduled DESIR visits). To evaluate the changes occurring between pre- and post-initiation of TNFα blockers, a definition of "baseline" (before initiation) and "follow-up" (after at least eight weeks of TNFα blocker treatment) visits was required.

"Baseline" visit was defined as the last DESIR cohort visit before the initiation of TNFα blockers or the visit taking place within seven days of initiating such treatment. "Follow-up" visit was defined as the first DESIR cohort visit taking place after at least eight weeks of treatment.

Variables

The variables collected at each DESIR cohort visit have been described in previous studies.(73) Briefly, information on patients' characteristics (age, gender, socio-demographic characteristics, comorbidities, smoking status, alcohol intake and family comorbidities), disease characteristics (axial disease, peripheral disease, criteria fulfilment [Amor(18), ESSG(19) or ASAS(21)], HLAB27 status, mNY sacroiliitis, MRI inflammatory lesions, enthesis ultrasound abnormalities, and bone densitometry status), disease activity (BASDAI(46) and Ankylosing Spondylitis Disease Activity Score (ASDAS)(84)) and disease severity (Bath Ankylosing Spondylitis Metrology Index (BASMI)(85), BASFI(58), Health Assessment Questionnaire for Ankylosing Spondylitis (HAQ-AS)(86) and variables for the calculation of the ASAS-NSAID score (e.g. name, mean dose and % days intake per week during the last 6 months(87)) were collected at each DESIR cohort visit or annually, according to the study protocol (available in English athttp://www.lacohortedesir.fr/desir-in-english/).

Effectiveness endpoints:

Effectiveness is the equivalent of *efficacy* when the treatment effect is observed in routine clinical practice, e.g. in *pragmatic trials*.(88,89)

Primary effectiveness endpoint

The primary effectiveness endpoint was the proportion of patients achieving an ASAS 40 response(57) after at least eight weeks of TNFα blocker therapy. An ASAS 40 response was defined as an improvement of at least 40% and an absolute improvement of at least two units (on a scale of 0 to 10) in three or more of the four areas: BASFI, pain, patient's global disease activity and mean from BASDAI questions five and six, with no worsening in the remaining areas.

Secondary effectiveness endpoints

Secondary effectiveness variables were evaluated at every DESIR cohort visit: BASDAI, BASFI and ASAS-NSAID score for the last week preceding the visit. Moreover, response criteria were evaluated with regard to the changes between the "baseline" and "follow-up" visits: ASDAS-CRP, ASAS 20 response criteria (defined as at least 20% improvement and absolute improvement of at least one unit [on a scale of 0 to 10] compared with baseline in three or more of the four areas: BASFI, pain, patient's global disease activity and mean from BASDAI questions five and six, with no worsening of more than 20% in the remaining area), BASDAI 50 response (defined as reduction of at least 50% or two units compared with baseline), ASDAS-CRP inactive disease (ASDAS-CRP ID) defined as an ASDAS-CRP <1.3, and ASDAS-CRP Clinically Important Improvement (ASDAS-CRP CII) and Major Improvement (ASDAS-CRP MI) defined as a decrease from baseline of ≥1.1 and 2.0, respectively.(90)

Study groups:

Effectiveness was compared in two groups of patients: the "active group" (e.g. the group of patients in DESIR receiving a TNFα blocker during the first two years of follow-

up) and the "control group" (e.g. a selected group of patients within the cohort receiving any other usual treatment and not given TNFα blockers). For this purpose, each patient from the "active group" was matched with a patient from the "control group". This matching process took into account that the baseline characteristics of these two groups were essentially different. To overcome these differences, the matching process was based on a score evaluating the probability of receiving a TNFα blocker. This methodology is called the propensity score method.(91,92) The propensity score is the probability of receiving a particular treatment conditioned on the individual baseline characteristics, here the probability of being treated with TNFα blockers. All baseline data (except for bone mineral density status and ultrasound enthesitis abnormalities, which were not available for all patients) were used to construct the multivariate logistic regression model predicting the probability of being treated with TNFα blockers (data not shown). The model performed adequately, AUC=0.880.

Missing data handling:

If a patient discontinued the treatment between two DESIR cohort visits, the information collected at the last DESIR cohort visit during treatment was collected and carried forward (e.g. using the last observation carried forward technique). If no visit was available between the baseline visit and time of discontinuation, the baseline observation was carried forward.

Statistical analysis:

Estimating the frequency of use of TNFα blockers:

The percentage of patients initiating TNFα blockers during the first two years of follow-up and its 95% confidence interval (95% CI) was estimated using the life table analysis technique according to the Kaplan–Meier method, at each intermediate DESIR cohort visit (e.g. 6, 12, 18 and 24 months).

Evaluating the effectiveness of TNFα blockers:

The primary effectiveness endpoint was estimated by assessing the percentage of patients achieving an ASAS40 response after at least eight weeks of treatment in the group of patients who received TNFα blockers for at least eight weeks and the group of patients who did not receive TNFα blockers but were given any other treatment (e.g. "control" patients). The control patients were matched with the group of patients receiving TNFα blockers not only in terms of baseline characteristics, but also time of evaluation. Therefore, the intervals between the follow-up visit and baseline visit were exactly the same as for the TNFα blockers group. The prediction of the ASAS40 response was computed by a logistic regression model and adjusted by BASDAI, BASFI and ASDAS-CRP values. A two-tailed p-value<0.05 was considered statistically significant.

Secondary effectiveness endpoints were assessed in the same way (e.g. ASAS 20, changes in BASDAI, ASAS-NSAID score, BASFI, etc., as described above).

These outcome measures were also evaluated for the TNFα blocker (e.g. etanercept, adalimumab or infliximab).

Exploring the interaction between baseline variables and the effectiveness of TNFα blockers:

To explore the impact of baseline characteristics on the primary effectiveness endpoint, we first reported the percentage of ASAS40 responders in each subgroup of patients according to the presence of the baseline characteristics: definite X-ray sacroiliitis (e.g. structural damage) (yes/no), MRI sacroiliitis (e.g. SIJ inflammatory lesions) (yes/no), CRP abnormality (yes/no), HLA-B27 positivity (yes/no), history of psoriasis (yes/no), history of peripheral arthritis (yes/no), smoking (yes/no), presence of objective signs of structural damage to the SIJ (X-ray sacroiliitis) or inflammation (MRI or CRP abnormality) (yes/no), and fulfilment of each arm of the ASAS criteria (Imaging/Clinical). For these analyses we considered only the patients for whom all the informations to appropriately classify them according to the ASAS criteria were available (e.g. among the 197 patients who received at least 1 TNFα blocker, data for evaluating the fulfilment of such criteria was available for 194 patients (e.g. for 3 patients data were missing for such evaluation), and among those, 146 fulfilled the ASAS criteria and 48 did not (197 = 146+48+3); among the 146 patients who fulfilled the ASAS criteria 96 patients presented with either structural (X-ray; n=64) and/or inflammatory abnormalities of the SIJ (MRI; n=79), and 50 had no imaging abnormalities of SIJ (e.g. were imaging "negatives") but had B27+ and 2 SpA features.)

Thereafter, we evaluated the possibility of interaction between the presence of these baseline characteristics and the treatment effect. For this purpose we performed a univariate analysis by logistic regression, with a significant interaction defined as $p \leq 0.10$. Finally, we performed a multivariate analysis (including in the model all

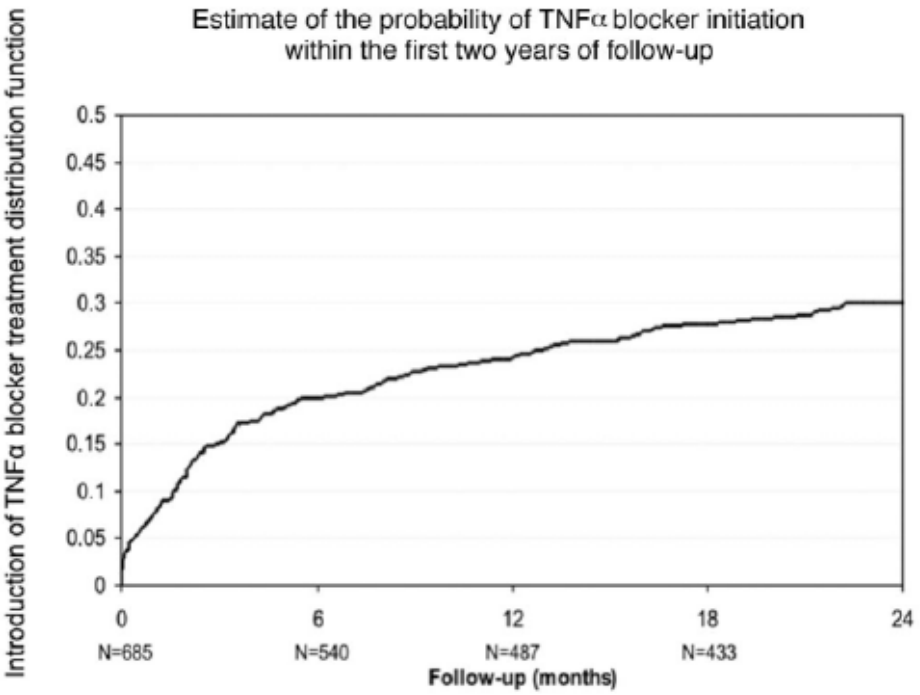
variables with p<0.30 on univariate analysis, but also the interaction model) to explore the most relevant variable.

RESULTS:

Estimating the frequency of use of TNFα blockers:

Of the 708 patients included in our analysis, data for Kaplan–Meier estimates were available for 685. Kaplan–Meier estimates of the proportion of patients initiating a TNFα blocker were 20.0% [17.1–22.9], 24.4% [21.1–27.7], 27.9% [24.6–31.2] and 30.2% [26.7–33.7] at 6, 12, 18 and 24 months after inclusion, respectively (Figure 10).

Figure 10:Estimate of the probability that treatment with a TNF alpha blocker will be initiated within the first 2years of follow-up in a cohort of patients with early axial spondyloarthritis.



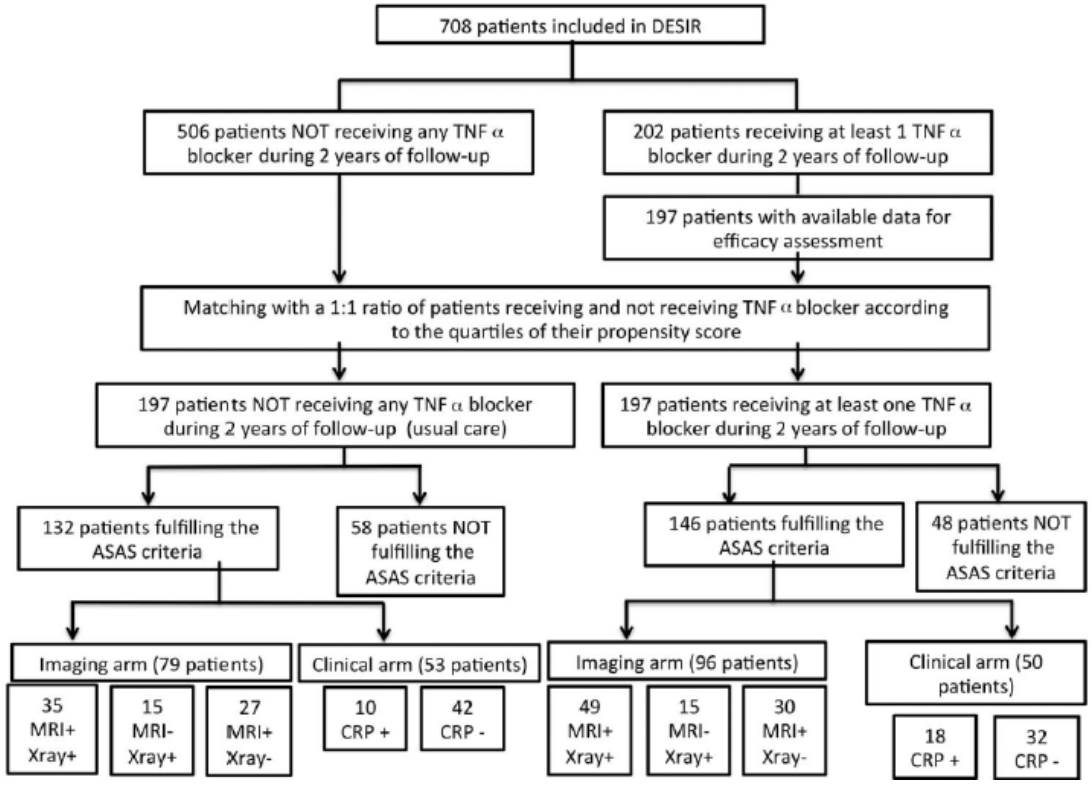
Numbers across the bottom are the number of patients at risk at the beginning of the indicated interval

Overall, a total of 203 patients (28.7% of the initial cohort) received at least one TNFα blocker during follow-up. Figure 11 summarizes the population analysis flowchart.

Among the 203 patients initiating their first TNFα blocker treatment, 103 (51.0%), 79 (39.1%) and 20 (9.9%) patients received Etanercept, Adalimumab and Infliximab, respectively.

.

Figure 11: Flow chart showing the distribution of study patients included in the analysis of the DESIR data.



For the imaging arm, the presence or absence of sacroiliitis on magnetic resonance imaging (MRI) and radiography was determined. For the clinical arm, the presence or absence of an elevated C-reactive protein (CRP) level was determined.

Table 11 summarizes the baseline characteristics of the patients included in our effectiveness analysis after being matched by propensity score. Interestingly, in the group of patients receiving TNFα blockers, 72 patients (35.5%) presented without any objective sign of inflammation or structural damage at baseline, while the other 127

patients (62.6%) presented with either X-ray sacroiliitis (67, 33.1%) or MRI sacroiliitis (82, 40.4%) or CRP abnormality (83, 40.9%).In the group of 197 patients who received a TNFα blocker, 75.3%, 86.3% and 88.3% fulfilled the ASAS, Amor and ESSG criteria, respectively, resulting in 95.9% patients fulfilling at least one criteria set. No significant differences between the groups were found except for BASDAI, BASFI and ASDAS-CRP, with higher values in the TNFα blockers group.

Evaluating the effectiveness of TNFα blockers:

The follow-up visit took place on average after 21.6 +/- 7.3 weeks (Min–Max: 8.4–45.1). A significantly greater percentage of patients achieved the primary endpoint, ASAS40, after at least eight weeks of treatment with TNFα blockers compared to patients receiving any other treatment (62/197 (31.8%) vs. 26/197 (13.5%), OR 2.99 [1.80–4.99], intergroup p-value=0.0002; Figure 12A).Similar results were observed for the secondary effectiveness endpoints, e.g. significantly higher percentages of patients receiving TNFα blockers achieving any other clinical response criteria compared to patients receiving any other usual care (Table 13). It should be emphasized that results on symptomatic outcome variables (e.g. ASAS response criteria) were observed concomitantly to a reduction in NSAID intake during the study (e.g. changes in the ASAS-NSAID score of -35.5±66.3 and -22.5±52.3 in the "active" and "control" groups, respectively, p-value=0.019).No significant differences were observed in the primary effectiveness endpoint (ASAS40) with regard to the TNFα blockers: 25/78 (32.1%), 33/101 (33.0%) and 4/18 (23.5%) for Adalimumab, Etanercept and Infliximab, respectively (interaction p-value=0.529). Similar percentages of response were observed for the secondary endpoints, irrespective of the type of drug (Table 13).

Table 11: Baseline characteristics of the study population

	Matched patients of the DESIR cohort N = 404		
	Patients receiving TNFα blockers N = 202	**Patients receiving usual care N = 202**	***p*****
Age (years) (n=202)*	33.8 [27.1–41.44]	33.7 [27.2–39.7]	0.893
Gender (male) (n=202)	88 (43.6)	80 (39.6)	0.419
Disease duration (months) (n=202)	17.2 [10.0–27.6]	16.3 [8.6–24.2]	0.254
HLA-B27 positive (n=202)	115 (56.9)	114 (56.4)	0.920
mNY sacroiliitis(n=199 vs. 196)	67 (33.5)	59 (30.1)	0.187
MRI sacroiliitis(n=198 vs. 194)	82 (41.4)	64 (33.0)	**0.085**
BASDAI (0–100)(n=201 vs. 201)	57.0 [45.0–67.0]	53.0 [39.0–65.0]	**0.012**
BASFI (0–100)(n=202 vs. 199)	42.0 [25.0–59.0]	35.0 [18.0–55.0]	**0.005**
CRP (mg/L)(n=196 vs. 198)	5.0 [3.0–13.6]	4.4 [2.9–8.0]	**0.070**
ASDAS-CRP	3.1 [2.4–3.7]	2.8 [2.2–3.3]	**0.006**
NSAID intake (yes)(n=192 vs. 187)	174 (88.3)	168 (85.3)	0.372
ASAS-NSAID score (previous week) (n=197 vs. 197)	55.0 [0.0–100.0]	50.0 [0.0–100.0]	0.447

**Data in the table are either number and (%) or median and [interquartile range: Q1–Q3].*

***Statistical differences determined by chi-squared test or T-student tests as appropriate. Statistical significance was established for p<0.05.*

Abbreviations: mNY: Modified New York; MRI: Magnetic Resonance Imaging; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; CRP: C-reactive protein; ASDAS: Ankylosing Spondylitis Disease Activity Score; TNF: Tumour Necrosis Factor;NSAID: Non-Steroidal Anti-Inflammatory Drugs.

Table 12: Effectiveness endpoints

Variable	All patients N = 394	Patients receiving TNFα blockers N = 197	Patients receiving usual care N = 197	OR [95% CI]†	p-value**
ASAS20*	130 (33.5)	90 (46.2)	40 (20.7)	3.23 [2.10–5.13]	<0.0001
ASAS40	**90 (23.1)**	**62 (31.8)**	**28 (13.5)**	**2.99 [1.80–4.99]**	**0.0003**
Δ BASFI	-8.6 ±19.5	-12.9 ±20.3	-4.3±17.3	-	0.014
Δ BASDAI	-11.6 ±20.4	-18.2 ±21.7	-5.0±17.5	-	<.0001
BASDAI50	248 (63.3)	151 (77.0)	97 (49.5)	2.58 [1.62–4.11]	<.0001
ASDAS-CRP	-0.7 ±1.2	-1.0 ±1.3	-0.3 ±1.0	-	0.0002
ASDAS-ID	90 (23.7)	58 (29.4)	35 (17.9)	3.20 [1.79–5.72]	<.0001
ASDAS-MI	51 (13.3)	39 (20.3)	12 (6.3)	3.52 [1.57–7.90]	0.017
ASDAS-CII	125 (32.6)	89 (46.4)	36 (18.9)	3.72 [2.23–6.21]	<.0001
NSAID intake	278(70.6)	141(71.6)	137 (69.5)	-	0.658
Δ ASAS-NSAID score (last week)	-29.0 ±60.0	-35.5 (±66.3)	-22.5 (±52.3)	-	0.019

**Data in the table are either number and (%) or mean and (± standard deviation) **Statistical significance was established for p<0.05 † Adjusted OR (for BASDAI, BASFI and ASDAS-CRP)*

Abbreviations: ASDAS: Ankylosing Spondylitis Disease Activity Score; ASDAS-ID: ASDAS Inactive Disease; ASDAS MI: ASDAS Major Improvement; ASDAS CII: ASDAS Clinically Important Improvement; NSAID: Non-Steroidal Anti-Inflammatory Drugs; Δ: Change From Baseline.

Table 13: Effectiveness endpoints by TNFα blocker

	ADALIMUMAB				ETANERCEPT				INFLIXIMAB			
	Patients receiving Adalimumab N = 78	Patients without TNF blockers matched with the patients receiving Adalimumab N = 78	OR† [95% CI]	*p***	Patients receiving Etanercept N = 101	Patients without TNF blockers matched with the patients receiving Etanercept N = 101	OR [95% CI]	*p*	Patients receiving Infliximab N = 18	Patients without TNF blockers matched with the patients receiving Infliximab N = 18	OR [95% CI]	*p*
ASAS20*	34 (43.6)	12 (15.6)	4.47 [1.90 – 10.52]	0.001	50 (50.0)	22 (22.2)	3.22 [1.66 – 6.23]	0.001	6 (35.3)	6 (35.3)	0.64 [0.09–4.40]	0.649
ASAS40	**25 (32.1)**	**10 (13.0)**	**3.07 [1.22 – 7.73]**	**0.017**	**33 (33.0)**	**14 (14.1)**	**3.42 [1.58 – 7.38]**	**0.002**	**4 (23.5)**	**2 (11.8)**	**1.80 [0.18–18.08]**	**0.619**
Δ BASFI	-14.1 (±23.2)	-2.6 (±16.8)	-	0.027	-13.0 (±19.0)	-5.0 (±18.0)	-	0.027	-7.3 (±12.9)	-7.8 (±15.5)	-	0.864
Δ BASDAI	-16.7 (±22.4)	-3.6 (±17.7)	-	0.001	-19.4 (±21.2)	-5.5 (±18.1)	-	0.002	-18.2 (±22.3)	-8.4 (±13.2)	-	0.838
BASDAI50	59 (75.6)	36 (46.8)	3.46 [1.63–7.36]	0.001	78 (78.0)	50 (49.5%)	3.02 [1.58–5.79]	0.001	14 (77.8)	11 (61.1)	0.64 [0.11–3.91]	0.700
ASDAS CRP	31 (39.7)	16 (20.5)	-	0.0002	40 (39.6)	35 (34.7%)	-	0.012	7 (38.9)	8 (44.4)	-	0.893
ASDAS ID	-0.95 (±1.41)	-0.23 (±0.96)	-0.6 (±0.2)	0.001	-1.06 (±1.28)	-0.36 (±1.07)	-0.6 (±0.2)	0.0002	-1.27 (±1.29)	-0.4 (±0.7)	-0.3 (±0.3)	0.185
ASDAS MI	22 (28.2)	12 (15.4)	5.32 [1.93–14.69]	0.001	31 (30.7)	20 (20.0)	2.69 [1.24–5.82]	0.012	5 (27.8)	3 (16.7)	5.15 [0.25–105.0]	0.912
ASDAS CII	16 (21.3)	6 (7.9)	3.37 [1.08–10.47]	0.036	18 (18.0)	6 (6.2)	2.46 [0.82–7.42]	0.109	5 (29.4)	0 (0.0)	>999 [<.001–>999]	0.959

**Data in the table are either number and (%) or mean and (± standard deviation); **Statistical significance was established for p<0.05; †:Adjusted OR (for BASDAI, BASFI and ASDAS-CRP)Abbreviations: ASDAS: Ankylosing Spondylitis Disease Activity Score; ASDAS-ID: ASDAS Inactive Disease; ASDAS MI: ASDAS Major Improvement; ASDAS CII: ASDAS Clinically Important Improvement*

Exploring the interaction between baseline variables and the effectiveness of the TNFα blockers:

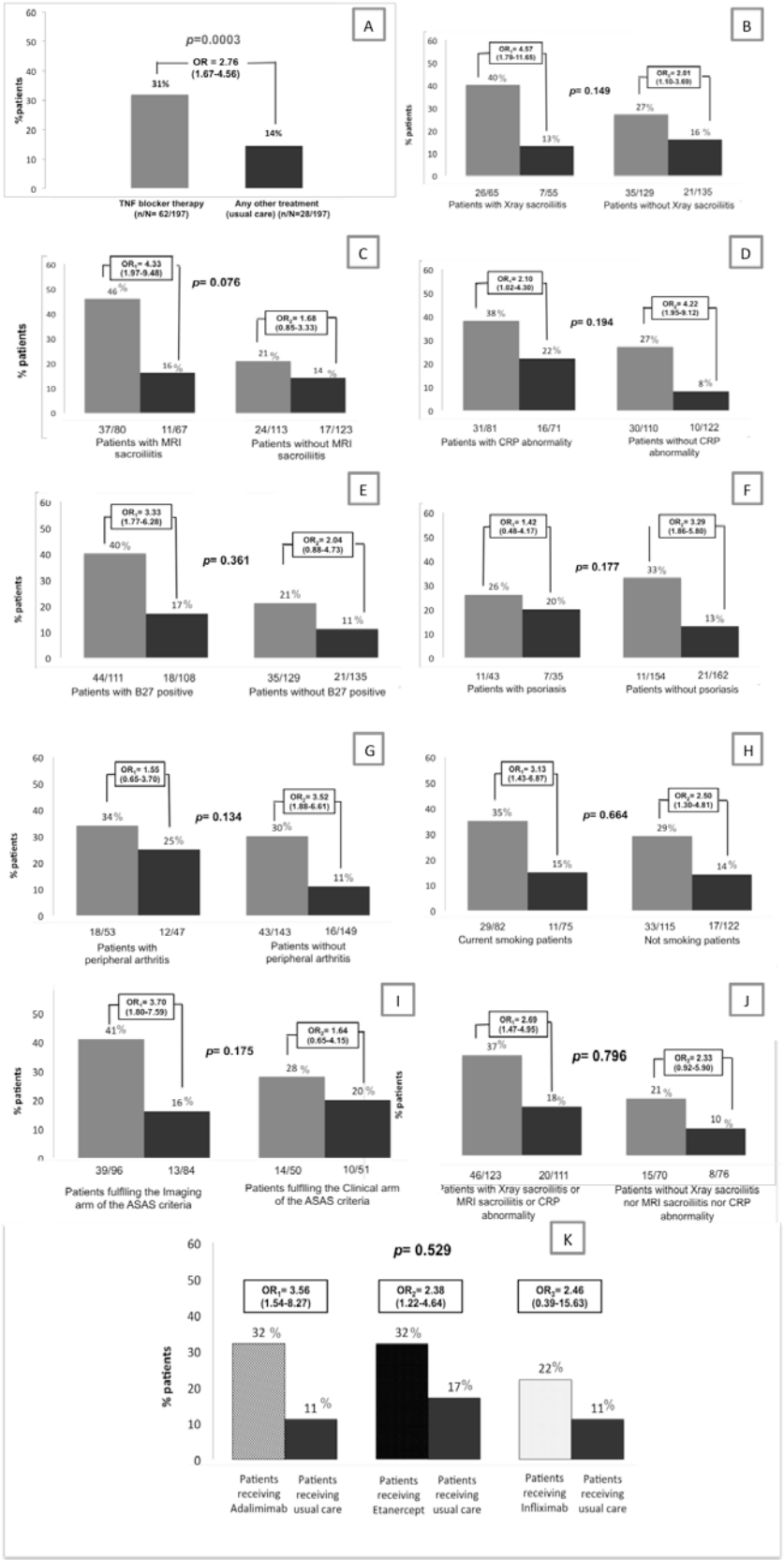
No interactions were found between the ASAS40 response and the presence of X-ray sacroiliitis, CRP abnormality, HLA-B27 positivity, peripheral arthritis, psoriasis or smoking status at baseline (Figure 12).

The only statistically significant interaction found was the presence of MRI sacroiliitis (OR=4.99 [2.17–11.51] vs. OR=1.75 [0.88–3.46] for the presence vs. absence of MRI sacroiliitis, respectively, interaction p-value=0.057) (Figure 12C).

However a trend to a greater treatment effect was observed in the subgroup of patients fulfilling the imaging arm and/or in the subgroup of patients with any objective sign of inflammation (e.g. MRI sacroiliitis or raised CRP) or structural damage of the SIJ (e.g. X-ray sacroiliitis). (Figure 12I and 12J)

The multivariate analysis confirmed the presence of MRI sacroiliitis at baseline as a predisposing factor for the TNFα blockers' treatment effect (OR=3.45 [1.81–6.60], p=0.0002)

Figure 12: Subgroup interaction analysis of ASAS 40 response after at least eight weeks of treatment



DISCUSSION:

Our study confirms that TNFα blockers are widely used in clinical practice in early axSpA patients. As much as 30% of patients initiated this treatment during the first two years of follow-up. Secondly, we confirm the effectiveness of TNFα blockers in "real-life" conditions (with 32% of patients achieving an ASAS40 response after 22 weeks of TNFα blocker treatment on average). Finally, the exploratory subgroup interaction analyses suggest that early axSpA patients presenting with MRI sacroiliitis are more likely to present with an ASAS40 response.

The percentage of patients initiating TNFα blockers was high in our study and a majority received this treatment after only six months of follow-up. This high rate may be explained by the study design (e.g. patients attending their DESIR cohort visits in a tertiary care centre would have the opportunity to meet rheumatology research nurses and other patients, and would obtain information about their disease and the potential therapeutic options (e.g. TNFα blocker treatment) more easily). However, patient recruitment in DESIR was performed in close connection with local community rheumatologists to ensure that the cohort was as representative as possible of patients with inflammatory back pain suggestive of axSpA. Although the percentage of patients initiating TNFα blockers might seem high, similar rates have been reported in previous studies.(43,93)

The results observed in this study in terms of changes to the symptomatic outcome variables should be compared to those reported in RCTs on non-radiographic axSpA. In these trials, the reported magnitude of treatment effect (e.g. the difference in ASAS40 response rates between the active and control groups) was around 22% (21%, 16% and

28%for adalimumab, etanercept and certolizumab, respectively(37,42,43)). In our study this magnitude was 18.3% and therefore comparable to the results observed in such trials.

Moreover, a huge difference between our study and conventional RCTs lies in NSAID intake during the study. In conventional RCTs, it is recommended that NSAID intake remains stable throughout the study.(94) In our study, however, NSAID intake was at the discretion of the patient and his/her rheumatologist. Here the TNFα blocker treatment resulted in a statistically significant reduction in NSAID intake. Therefore, the symptomatic treatment effect of TNFα blockers observed in our study (compared to other RCTs), in a context of reduced NSAID intake, might be considered clinically relevant.

Our study suggests a similar treatment effect in all three evaluated TNFα blockers. Many factors have been suggested to predict a greater therapeutic response to TNFα blockers in axSpA (e.g. age, disease duration, elevated CRP, BASDAI, BASFI and HLA-B27 positivity)(42,48–50); more recently, RCTs(37) have reported that signs of inflammation (MRI sacroiliitis and/or elevated CRP) are predictors of better response in non-radiographic axSpA. In order to address this question in our study, we considered that the appropriate statistical test was the interaction test and not the inter-treatment group test (e.g. TNF- blocker vs. usual care) in each subgroup of patients (e.g. subgroup of patients fulfilling the imaging arm (one test) and thereafter subgroup of patients NOT fulfilling the imaging arm (another test)).

The single baseline characteristic with a significant interaction towards the treatment effect was the presence of an MRI sacroiliitis (the threshold p-value we used in our

subgroup analysis, which is commonly used in subgroup analysis, was 0.10, in order to increase the statistical test power).

However, a greater treatment effect was observed in the subgroups of patients fulfilling the imaging arm (25%) compared to the subgroup of patients fulfilling the clinical arm (9%), suggesting the existence of a difference even in the absence of a statistically significant interaction test (p=0.194). Such conclusion might be reinforced by the fact that the 95% confidence interval (95%CI) of the Odds Ratio to predict an ASAS40 response in the subgroup of patients fulfilling the imaging arm was above 1 ((OR=3.77 [1.80–7.90], resulting in a significant inter-treatment group difference ($p<0.05$) in the imaging arm, while the 95%CI of the Odds Ratio to predict an ASAS40 response in the clinical arm contained 1 (OR=1.72 [0.68–4.34]) resulting in a non-significant inter-treatment group difference ($p>0.05$) in the clinical arm.

Identically, a greater treatment effect was observed in the subgroup of patients presenting with objective signs of either inflammation (MRI sacroiliitis or raised CRP) or structural damage of the SIJ (X-ray sacroiliitis) (16%) compared to the subgroup of patients without any objective sign (10%), here again suggesting a difference, even in the absence of a statistically significant interaction test (p=0.489); identically, the 95%CI of the Odds Ratio to predict an ASAS40 response in the subgroup of patients with objective signs was above 1 (OR=3.46 [1.81-6.60]) (statistically significant difference between treatment groups), while the 95%CI of the Odds Ratio to predict an ASAS40 response in the subgroup of patients without any objective signs contained 1 (OR=2.31 [0.90-5.93]) (non-significant statistical difference between the treatment groups).

However we would like to emphasize that the interpretation of such p values is very hazardous since such analyses are post-hoc analyses performed in sub-groups of patients

CRP abnormality has been reported as a predisposing factor of better response: in our study no statistically significant differences were observed with regard to abnormal CRP (interaction test was non significant; 95%CI of OR to predict an ASAS40 response were above 1 in the subgroup of patients with raised CRP, but also in the subgroup of patients with normal CRP; Figure 12D) . The latter might be explained by the design of our study: the "baseline" parameters in our analysis are those recorded at the DESIR cohort visits, and these parameters might be different to those observed the day before initiating the TNFα blockers as the patient could be stable at the DESIR visit (e.g. M6), present with a flare-up two months later and be started on TNFα blockers (e.g. at M10) between two DESIR visits. Moreover, because of multiple testing, the observed statistically different treatment effect with regard to patients' characteristics might have been obtained by random chance alone. Therefore, our results should be considered with caution (e.g. interaction with the presence of MRI sacroiliitis or absence of interaction with elevated CRP). To avoid these biases, well-powered explanatory trials of TNFα blockers in early axSpA patients should be conducted, particularly in the subgroup of patients with no objective signs of structural damage or inflammation (e.g. patients fulfilling the "clinical arm" of the ASAS axSpA criteria with normal CRP).

This study has some limitations that are worth considering. Firstly, DESIR included patients initiating with chronic IBP, with a likelihood of axSpA diagnosis of more than

50% according to the rheumatologist. Although rheumatologists were asked at the end of each visit whether another diagnosis was more likely in order to exclude the patient from the study, it is not impossible that the cohort included patients with conditions other than axSpA (e.g. IBP linked to degenerative discopathies). However, the majority of patients included in the study (95.9%) fulfilled at least one criteria set for SpA.

Secondly, DESIR is an observational study with potential biases and confounding, mainly due to the lack of randomization. As recommended in this situation(54,55,91,95), we have addressed this potential bias on estimation of treatment effect by matching the "active" and "control" patients based on a propensity score, bringing our study close to a *pragmatic trial* and providing valuable data on real-life prescription rates and therapeutic effects in clinical practice.

This study also has some additional strengths. Firstly, it is the largest prospective observational cohort of IBP patients suggestive of early axSpA patients to describe and analyse the use of TNFα blockers in a "real-life" clinical setting in detail. Secondly, because of the quality of data collection both at inclusion and every six months, data to assess the primary outcome was complete with regard to clinical, physical, laboratory and imaging in the large majority of patients.

Our study confirms the effectiveness of TNFα blockers in an early axSpA "real-life" clinical setting in terms of clinical response, particularly in the subgroup of patients with MRI inflammatory signs at SIJ level. Further explanatory RCTs focusing on different subgroups of patients, and especially the subgroup of patients fulfilling the "clinical" arm of the ASAS criteria with normal CRP, are required to explore the role of TNFα blockers in this context.

SUPPLEMENTARY ANALYSIS OF ARTICLE 3

Due to the limited words allowed for the manuscript in the journal, we could not include the analysis we performed regarding the best propensity score to use in the matching process, or the treatment effect at long term, its retention rate nor the analysis of the TNF alpha switch. These results have been presented at several international meetings, but have not been published yet.

Objectives of the supplementary analysis:

- To evaluate the 2-years effectiveness of TNFα blockers in a cohort of patients with early ax-SpA
- To estimate its retention rate after 2 years of follow-up and the baseline predictors of such continuation and
- To evaluate the percentage of patients switching from fist TNFα blocker to at least another TNFα blocker and the baseline factors predicting such switch

Patients and methods:

Study design: identical to ARTICLE 3 (page 70).

Statistical analysis:

Propensity score modelling:

In order to deal with the lack of randomization in the treatment effect evaluation, and in order to balance the covariates in both groups, we estimated a propensity score for TNFα blockers, e.g. the probability to receive the treatment according to the observed covariates. Despite they are broadly use, there is no consensus regarding the selection

of the covariates to include in the score, and for this we estimated two different propensity score.

First, we estimated by logistic regression a propensity score predicting to receive a TNFα blocker, including in the model all the covariates available at baseline (except for comorbidities), with no variable selection (no univariate testing or correlation testing). This model was called "PScomplete", and was the one used for effectiveness analysis.

Secondly, we estimated by logistic regression another propensity score ("PSselect"): for this estimation, first, the association of each covariate with the outcome was tested by univariate analysis, and only variables with a significant (p<0,10) association with the outcome were selected. Among the selected variables, co-linearity of variables was tested, and among the variables with high correlation, only the most clinically informative were selected. Finally variables with high clinical significance could be added manually to the model.

Model comparison: the two nested models (PSselect as a nested model of PScomplete) were compared by the Likelihood Ratio Test. The area under the curve (AUC) and the Akaike information Criterion (AIC) were also assessed for both models.

Outcome comparison: patients were matched (nearest neighbour technique) according to both PS models, and the outcome was estimated by logistic regression.

Two-years effectiveness:

Patients receiving TNFα blocker were matched with patients not receiving TNFα blocker (but receiving usual care according to their rheumatologist) at a 1:1 ratio according to the "PScomplete" propensity score. The primary endpoint was ASAS40, assessed at the last available visit still on treatment (mean 74±30.9 weeks), was

predicted by conditional logistic regression, including in the model the variables the matching failed to balance. Exploratory subgroup analysis for interactions towards this treatment effect was performed (for this analysis, we established an alpha error of 10%, in order to increase the subgroup analysis power).

Retention rate:

Adherence to treatment was estimated for all TNFα blocker using a survival analysis (Kaplan-Meier). Multivariate Cox regression was performed to identify potential predictors of such treatment adherence.

Switching:

The percentage of patients switching TNFα blocker was described, and both groups (switchers vs. non-switchers were) compared with regard to their baseline characteristics by T-test or Chi-square tests as appropriate (or Wilcoxon or Fisher tests as appropriate in the absence of normal distribution). Finally, predictors of switch were estimated by multivariate logistic regression.

Results:

Propensity modelling*:*

"PScomplete" included 52 variables (Table 15)whereas "PSselect" included only 9 variables. The variables selected in PSselect with a significant association with outcome were: gender (p=0,0024), presence of synovitis (p=0.0297), history of uveitis (p=0.0138), highest education level (p=0.0859), BASGweek (p=0.0174), VAS disease activity (physician) (p=0.0034), HLA-B27 status (p=0.0087), ESR abnormality (p <0.0001), CRP abnormality (p=0.0002), Xray sacroiliitis (p=0.0072) and MRI sacroiliitis

(p=0.0001); among these, a correlation was found between VAS disease activity and BASGweek (r=0.54), MRI and X-ray sacroiliitis (0.49), X-ray and CRP (0.29), ESR and VAS disease activity (0.21). We decided to keep both MRI and Xray sacroiliitis and CRP abnormality in the final model based on a clinical decision. The final PSselect model included 9 covariates: gender, synovitis, global VAS for disease activity, highest level of education, history of uveitis, MRI sacroiliitis, X-ray sacroiliitis, CRP abnormality and HLAB27 status.

Model comparison: The likelihood ratio test showed a statistically significant difference: PScomplete fitted the data significantly better. Moreover, AUC and AIC were also in favour of PScomplete. (Table 16)

Two-years effectiveness:

Of the 708 enrolled patients, 202 patients received at least one TNFα blocker during follow-up, but data for the assessment of the primary endpoint was available in 197 patients receiving TNFα blockers, that were matched according to a the PScomplete propensity score to 197 patients receiving any other usual care (identical matching technique as in ARTICLE 3, page 70).An ASAS40 response at the last available visit (mean 74±30.9 weeks) was found in 62 (31.8%) patients receiving TNFα blocker vs. 31 (16.0%) patients receiving usual care (OR = 2.45 [1.50 – 3.99], p=0,004).

Exploratory subgroup analysis found that males, patients with X-ray sacroiliitis, MRI sacroiliitis, fulfilling the imaging arm, or without psoriasis at baseline were more likely to achieve an ASAS40 response at long-term with TNFα blockers, compared to any other usual care (Table 17). No interaction towards the treatment effect was found for

CRP abnormality, history of peripheral arthritis, history of enthesitis, nor smoking status.

Retention rate*:*

Of the 202 patients included in our analysis, Kaplan-Meier estimates of the proportion of patients still on TNFα blockers over time were 75.2% [69.0 – 81.3], 56.1% [48.8 – 63.3], 50.8% [43.3 – 58.3], 41.2% [33.2 – 49.1], at 6, 12, 18 and 24 months after initiation, respectively. HLA B27 presence (HR = 1.52 [1.01 – 2.27], p=0.044) and X-ray sacroiliitis at baseline (HR = 2.08 [1.28 – 3.33], p=0.003) were associated with continuation of the TNFα blockers over time.

Switching:

Seventy of the 197 patients (35.5%) switched after the first TNFα blocker over the first 2 years of follow-up: 16 and 5 patients switched from Adalimumab to Etanercept and Infliximab, respectively; 42 and 4 patients switched from Etanercept to Adalimumab and Infliximab, respectively; 2 and 1 patients switched from Infliximab to Adalimumab and Etanercept, respectively.

In univariate analysis, switchers were more frequently females (70% vs. 49.6%, p=0.006), more frequently HLAB27 negative (57.1%vs. 36.2%, p=0.005), had less frequently imaging abnormalities of the sacroiliac joints (18.6% vs. 41.9% for radiographic sacroiliitis, p=0.001; 25.7% vs. 50.4%, p=0.001 for MRI sacroiliitis), higher mean BASDAI (59.3 (±13.9) vs. 53.1 (±16.1), p=0.021) and BASFI (49.4 (±21.8) vs. 37.7 (±19.5), p=0.0003) scores, and lower mean CRP values (10.9mg/L (±21.7) vs. 14.3mg/L (±18.4), p=0.005).

The multivariate analysis (including all the significant variables in univariate analysis and the center) confirmed the presence of HLAB27 (OR=0.44 [0.23 – 0.84], p=0.013) and radiographic sacroiliitis (OR=0.34 [0.16 – 0.72], p=0.005) as protective factors towards switch of the first TNFα blocker, but interestingly also the BASFI score as a predictive factor for switch (OR=1.03 [1.01 – 1.05], p=0.01).

Discussion:

These analyses suggest that the effectiveness of TNFα blockers in "real-life" conditions is maintained at long-term (with 32% of patients achieving an ASAS40 response after near 75 weeks of TNFα blocker treatment on average). However, there were more baseline characteristics that interacted with TNFα blockers treatment towards an ASAS40 response, and fulfilling the imaging arm (e.g. having either MRI or radiographic sacroiliitis) was a predisposing factor to achieve an ASAS40 response at long-term. Furthermore, being a male and not having psoriasis were also found to be predisposing factors for an ASAS40 response.

When exploring the retention rate of TNFα blockers, only 50.8% [43.3 – 58.3] patients were still on the first TNFα blocker after 18 months of treatment. These results are significantly lower than those previously reported in ankylosing spondylitis populations(96–98)where retention rates were estimated near 75% after 2 years of treatment. These differences might be explained by the differences in the disease duration, since patients in DESIR present with IBP evolving for less than 3 years, and the other reports, evaluated retention rate in AS, patients, with longer disease evolutions.

Table 14: Variable included in the propensity score "PS COMPLETE"

Centre
Age
Gender (Female vs. Male)
Ethnicity (Caucasian vs. Other)
NSAID sensitivity (Yes/No)
Symptoms delay
Past history of peripheral arthritis (Yes/No)
Past history of enthesitis(Yes/No)
Dactylitis(Yes/No)
Anterior chest pain (Yes/No)
Extra-articular signs(Yes/No)
Infection preceding (3 months) the symptoms(Yes/No)
NSAID score
DMARDs(Yes/No)
Corticoid treatment(Yes/No)
Menopause(Yes/No)
Smoking(Yes/No)
Alcohol (Yes/No)
Family history of SpA(Yes/No)
Family history of psoriasis (Yes/No)
Family history of uveitis (Yes/No)
Family history of de MICI (Yes/No)
College education
Marital status (Married vs. Single)
Parental status (at least one child vs. no child)
Work (Invalidity vs. other)
BASG
PASS
Nocturnal awakening(Yes/No)
BASDAI
BASFI
HAQ-AS
SF36-MCS
SF36-PCS
ASQOL
Pain (spine, night, peripheral)
BMI
Tender joint count
Swollen joint count
Enthesitis score
BASMI
Global evaluation by physician
HLAB27 (Yes/No)
ESR
CRP
Radiographic sacroiliitis (Yes/No)
MRI sacroiliitis (Yes/No)
At least 1 syndesmophyte (Yes/No)

Table 15: Propensity score model comparison

	Likelihood ratio	AUC	AIC

	test		
PS complete(52 variables)		0.8759	698.64
	p<0.0001		
PS select (9variables)		0.7614	724.52

Finally, while age (98)and CRP(97) have been reported as predictive factors for continuations, our analyses found that patients with radiographic sacroiliitis and HLAB27 positive were more likely to continue the first TNFα blocker; identically our study did not confirm enthesitis and peripheral involvement as predictors for switch of the first TNFα blocker, as it was previously reported(96), but we found that BASFI score was a predictor for switch, and HLAB27 and Radiographic sacroiliitis were protective factors for such switch. Again these differences might be explain by the early disease of the patients included in DESIR, compared to the other studies. These analysis have some limitations, mainly that the reason for switch or discontinuation was not collected in the DESIR cohort, and we therefore cannot differentiate discontinuation or switch secondary to inefficacy or adverse events. However, this study represents the largest cohort of early SpA and confirms the effectiveness of TNFα blockers at long-term, furthermore, this study confirms that features from the imaging but also the clinical arm (radiographic sacroiliitis and HLAB27, independently) are predictive factors for continuation, suggesting that not only patients fulfilling the imaging arm are likely to continue treatment, but also those fulfilling the clinical arm. Further exploratory analysis should try to investigate the role of BASFI as a predictive factor to switch.

Table 16: Exploratory subgroup analysis of predisposing factors to TNF alpha effectiveness at long term

Patients characteristics		Percentage of patients achieving an ASAS40 response		Odds Ratio of achieving an ASAS40 response at the last available visit still on treatment	*Interaction test*
		Patients receiving TNFα blockers	Patients receiving usual care		
X-ray sacroiliitis N=384	Yes n=120	31/65 (47.7%)	5/55 5(0.9%)	OR = 9.12 [3.22 – 25.80]	*p*=0.002
	No n=264	30/129 (23.2%)	23/135 (17.0%)	OR = 1.32 [0.73 – 2.41],	
MRI sacroiliitis N=383	Yes n=147	38/80 (47.5%)	10/67 (14.9%)	OR = 5.10 [2.28 – 11.41]	*p*= 0.011
	No n=236	23/113 (20.4%)	20/123 (16.3%)	OR = 1.32 [0.68 – 2.56]	
ASAS criteria N=281	Imaging n=180	39/96 (40.6%)	13/84 (15.5%)	OR = 5.34 [2.51 – 11.33]	*p*= 0.001
	Clinical n=101	14/50 (28.0%)	10/51 (20.0%)	OR = 0.78 [0.32 – 1.89]	
Gender N=394	Male n=169	36/85 (42.4%)	11/84 (13.1%)	OR = 4.84 [2.25 – 10.44]	*p*= 0.016
	Female n=225	26/112 (23.2%)	20/113 (17.7%)	OR = 1.41 [0.73 – 2.70]	
Skin psoriasis N=394	Yes n=78	9/43 (21.0%)	8/35 (22.9%)	OR = 0.92 [0.31 – 2.71]	p=0.048
	No n=316	53/154 (34.4%)	23/162 (14.2%)	OR = 3.13 [1.80 – 5.45]	

GLOBAL RESULTS AND GENERAL DISCUSSION

Validation of the ASAS axSpA criteria and its arms in a real-life clinical setting

The findings in our first study (ARTICLE 1, page 31) confirm the validity of the ASAS classification criteria as a whole, but also the high specificity of each arm of the ASAS criteria for axSpA in a daily-practice setting. Thus, our hypothesis of adequate properties for this set of criteria both for diagnosis and classification in this setting has been verified.

In addition, we also hypothesised that the metric properties of these set of criteria would be comparable to the other sets of criteria, but interestingly, they were (along with the mNY criteria) the only sets of criteria with an LR+ higher than 10, suggesting these set might be the best tool to classify early forms of SpA (as definite sacroiliac damage is not mandatory), with the better balance between sensitivity and specificity. We also anticipated that some symptoms or clinical features would contribute in larger extent to the diagnosis and classification of the patients than others, and we could assess in our study that sacroiliitis (MRI or radiographic)and an elevated CRP were the items with higher LR+, confirming the rationale of including these items in the ASAS axSpA criteria. Furthermore, the performance of these criteria were also adequate for all subgroups of SpA, although maybe to a less extent in the peripheral subgroups, such as psoriatic arthritis and reactive arthritis, but this can be explained by the fact that in our study we aimed to validate only the criteria for axSpA, and we did not test the performance of the peripheral ASAS criteria for SpA. Finally, this study allowed us

also to compare, albeit superficially, the patients fulfilling each arm of the ASAS criteria (i.e. “imaging” and “clinical”) in terms of age, gender, and BMI. The single difference we could find between groups was age, with older patients in the “clinical” group.

These latter results, prompted us to investigate the potential differences in terms of phenotype between the patients fulfilling the different arms of the ASAS criteria for axSpA in an early IBP population (ARTICLE 2, page 52).Our results confirmed that no striking differences in disease characteristics, activity, function or quality of life, except for age and gender, existed between these two groups of patients in our sample. This further supports the validity of the clinical arm of the ASAS criteria, since patients fulfilling either arm seem to be comparable in terms of presentation and disease burden to those fulfilling the “imaging” arm of the ASAS criteria for axSpA. However, as the clinical arm is still very much debated by our health authorities (e.g. in several countries only patients fulfilling the “imaging” arm AND with X-ray abnormalities of the SIJ can receive a TNF), we aimed to explore whether other imaging features besides those included in the “imaging” definition were found in these patients. Our results showed that as much as 22% and 11% of patients presented with MRI inflammatory lesions of the spine in the subgroup switch and without raised CRP, within the clinical arms, respectively. This seems to suggest that potentially the definition of “imaging” of the ASAS might need further discussion. However, these analyses were performed with the imaging reading by the local investigators (not central imaging reading, and an analysis of the concordance between our findings and those from the DESIR central reading is needed before driving any conclusion or revising the “imaging” definition.

Confirmation of the effectiveness of TNF alpha blockers in early axSpA

Our last study (ARTICLE 3, page 70) confirmed our hypothesis, i.e. that TNFα blockers are effective in real life conditions, with a magnitude of effect comparable to what has been reported in RCT. The subgroup of patients presenting with MRI sacroiliitis at baseline were more likely to present a treatment response, but no differences were found with regard to the fulfilment of the imaging and clinical arms of the ASAS criteria after at least 8 weeks of treatment. ASAS 40 response rates of patients fulfilling the clinical arm were greater in the group receiving TNFα blockers than in the group receiving usual care, although not statistically. Furthermore, in the long-term, the fulfilment of the imaging arm (compared to the clinical arm) was a predisposing factor towards an ASAS 40 response, and when looking at the subgroup of patients fulfilling the clinical arm, no treatment effect was observed.

All these findings seem to suggest that whereas the burden of disease seems to be similar regardless the arm of the ASAS criteria for axSpA the patient is fulfilling, the benefits of a TNFα blockers in the patients fulfilling the “clinical” arm seem limited, and need to be further confirmed by trials targeting this population.

CONCLUSIONS AND PERSPECTIVES

ASAS criteria arms: new imaging definitions?

Our findings confirm the validity of the ASAS criteria for axSpA, both in a clinical practice setting and in an early axSpA population, and for both arms.

Nevertheless, the presence of other imaging abnormalities in patients of the clinical arm of the ASAS criteria in DESIR, has prompted us to try to evaluate this time the performance of the lesions so-called "typical of SpA", both in radiographies and MRI (e.g. hyperintense corners in the spine, radiographic sacroiliitis, etc.…). For this, we are now conducting the ILOS project: Imaging in LOw back pain & Spondyloarthritis. This project aims to evaluate the specificity of these so-called "typical SpA" lesions other than radiographic and MRI sacroiliitis by enrolling a "control" group of patients for the DESIR cohort: a group of 100 patients with mechanical back pain initiating before 45 years, with > 3 months but < 3 years of duration. These ILOS patients will undergo the same identical imaging protocol as the DESIR patients, and their images will be randomly mixed with the images of 100 randomly selected patients that will be representative of the global cohort in terms of abnormalities. These 200 images will be read blindly in order to assess the sensitivity and specificity of these lesions and/or their combinations. This project is currently on-going and has the support of the ASAS society.

Further evaluation of the treatment effect in longitudinal observational studies:

Our third study allowed us to evaluate the effectiveness of TNFα blockers, and to identify the baseline factors associated with a favourable outcome after 8 weeks and up to 2 years of treatment. However, these analyses did not take into account the

potential changes in the co-medications taken by the patients in real-life; in this sense, it is difficult to call such analysis as “longitudinal”, since only baseline information is used to predict the outcome.

An alternative methodology that would allow us to evaluate the treatment effect in such setting adjusted by the time-changing prescriptions would be to apply extended Cox models for time-varying covariates. There, for example, information on NSAID consumption or any other co-medication could be included at each study visit. Our on-going analysis, exploring the time-varying effect of NSAID consumption in patients being treated by TNFα blockers should allow us to determine if indeed such variations are relevant, and if including this information in the evaluation of TNFα blockers treatment response might allow us to evaluate this effect more accurately.

ACKNOWLEDGEMENTS/AGRAÏMENTS

This is the end of what at the beginning seemed like a very long journey, but that has finally an exciting adventure, thanks to the people who have surrounded me.

First, I would like to thank all patients and professionals participating in the DECLIC and DESIR studies, since this work would not have been possible without them.

Secondly, I would like to express my deepest gratitude to my thesis directors, Dr Alejandro Olivé, Dr Loreto Carmona and Prof Dr Maxime Dougados, since none of this would have been possible without their unconditional support.

> Dear Alex you were my tutor, the first one to believe in me, to listen to my aspirations and dreams, and to support me in my international adventure. Thank you for all those emails/phone calls/*whatsapp* messages and your wise advice.
>
> Dear Loreto, you were in fact the reason why I decided to get involved in clinical research. The course you taught in Madrid for trainees in Rheumatology and the week I spent in the SER research unit, where you made research seem so fun, while standardized and simple was a life-changing experience, which led me to decide to pursue an academic career. Thank you for your unconditional availability and brilliant counselling.
>
> Dear Prof Dr Dougados, Dear Maxime (if I may), I can hardly find the words to describe what meeting you has meant to me. When I first arrive in Cochin to work as a *Chef de Clinique*, I was already feeling very lucky to work in your department, but I could have never dreamed what was coming. During these past 3 years, you have thought me to be rigorous, to be systematic, to work hard, to speak English with a French accent, and, above all, to love research. You have been, you are and you will always be my mentor, my inspiration and my guidance. Thank you for your generosity, for your patience and for believing that I was capable of things I could have never imagined.

I would like to thank Prof Dr Jordi Tor, for his support as my thesis tutor.

I would like to express my deepest gratitude to Prof Dr Désirée van der Heijde and to Prof Dr Pascal Claudepierre, who have been co-authors in the works I present in this thesis, and have acted as international expert reporters for this work. Your constructive and thoughtful comments have always been an inspiration, and I feel very lucky to have the opportunity to continue working with you within DESIR.

Thank you to Professor Monreal and Dr Cañete for having accepted to be members of my Thesis Jury. Thank you also to Dr Pedro-Botet and Dr Juanola for having accepted to be supplement members of this jury.

I would like to thank Prof Dr Laure Gossec, for having accepted to be part of my Thesis Jury. Laure, thank you for being the living example that is indeed possible to have it all, for your support, for your contagious optimism and your ability to make the difficult seem simple!

A special though for Marta Valero, my CTO Rheumatology teacher, who was the reason why I decided to pursue a Rheumatology training after medical school. Thank you for your enthusiasm and for being able to show how beautiful and challenging our specialty is.

Thank you to Dr Xavier Tena, Dra Susana Holgado and Dra Lourdes Mateo, from Hospital Universitari German Trias i Pujol, who guided me through my first years as a resident. Thank you for your sharing your knowledge, for your common sense and for making my training the incredible journey it was.

I would like to express my deepest gratitude to Prof Dr Frédéric Lioté, Prof Dr Thomas Bardin and Prof Dr Pascal Richette, from Lariboisière Hospital for giving me the first opportunity to work in France. Thank you for always making me feel like any of the other interns, for your words of encouragement, and for your patience when my French was not as fluent as it is now.

Thank you to all the Cochin Hospital Rheumatology B Department Staff, Dr Xavier Ayral and Dr Karine Briot, with whom I have had the opportunity to work very closely, with a special though of gratitude for Prof Dr Roux, who interviewed me for the Chef de Clinique position and gave me this life-changing opportunity that has been working here.

Thank you also to Dr. Séverine Neveu and Dr. Anne Blanchais, for being the best *co-Chefs* ever.

A special though also for all the people I have met thanks to EMEUNET (Fran, Sofia, Pedro, Caroline, Xenofon, Rucsandra, Elena) and that I believe I can now call not international colleagues but friends.

Thank you *Nenes del Ruti* (Vera, Marta, Txell, Elena, Eli, Sonia…) for being the best *R-grans* one could ever have. Melania, you were not my *R-gran*, but it was a pleasure being yours.

A special thought for Esther and Silvia, who followed my struggles when moving abroad, and were always a source of comfort, in front of a cup of coffee or in front of my Mac screen when the former was not possible.

Thank you Alice for being my first friend in Paris, and the best companion I could have ever wished for these first months in a new department/city/country/life. You were so young when I met you and already so wise!

Thank you, *les girls* (Gaetane, Lisa, Marine, Anaïs) for being my family here in Paris, my support in the difficult times, and the best *troupe* to celebrate the good news!

Finally, a huge thank you to my family:

> To my parents, Tomàs i Roser *per ser-hi sempre encara que sigueu lluny, per ser els meus ídols, per ajudar-me sempre en les decisions que semblen impossibles, i haver-me inculcat el gust per la feina ben feta i la cultura de l'esforç.*
>
> To my borthers, Marc i Pau: *a en Marc per no deixar mai d'ensenyar-me que sempre hi ha una manera més feliç de viure la vida, i a en Pau, per ensenyar-me que la disciplina i el rigor et poden portar molt lluny.*
>
> To my companion in life, Agustí, *gràcies per fer-me sempre costat, per veure la vida sempre del costat positiu, per motivar-me sempre i per ajudar-me quan no me'n surto. Sense tu aquesta aventura, aquest viatge, aquesta vida i aquesta tesi no haguessin estat possibles.*

REFERENCES

1. Dougados M, Baeten D. Spondyloarthritis. Lancet. 2011;377(9783):2127-37.

2. Moll JM, Haslock I, Macrae IF, Wright V. Associations between ankylosing spondylitis, psoriatic arthritis, Reiter's disease, the intestinal arthropathies, and Behcet's syndrome. Medicine (Baltimore). 1974;53(5):343-64.

3. The EULAR Textbook. http://www.eular.org/index.cfm?framePage=/edu_compendium.cfm

4. Lambert JR, Wright V, Rajah SM, Moll JM. Histocompatibility antigens in psoriatic arthritis. Ann Rheum Dis. 1976;35(6):526-30.

5. Stolwijk C, Boonen A, van Tubergen A, Reveille JD. Epidemiology of spondyloarthritis. Rheum Dis Clin North Am. 2012;38(3):441-76.

6. Hammer RE, Maika SD, Richardson JA, Tang JP, Taurog JD. Spontaneous inflammatory disease in transgenic rats expressing HLA-B27 and human beta 2m: an animal model of HLA-B27-associated human disorders. Cell. 1990;63(5):1099-112.

7. Taurog JD, Richardson JA, Croft JT, Simmons WA, Zhou M, Fernández-Sueiro JL, et al. The germfree state prevents development of gut and joint inflammatory disease in HLA-B27 transgenic rats. J Exp Med. 1994;180(6):2359-64.

8. Haroon N. Endoplasmic reticulum aminopeptidase 1 and interleukin-23 receptor in ankylosing spondylitis. Curr Rheumatol Rep.2012;14(5):383-9.

9. Braun J, Sieper J. Ankylosing spondylitis. Lancet. 21 de2007;369(9570):1379-90.

10. Dean LE, Jones GT, MacDonald AG, Downham C, Sturrock RD, Macfarlane GJ. Global prevalence of ankylosing spondylitis. Rheumatol Oxf Engl.2014;53(4):650-7.

11. Braun J, Bollow M, Remlinger G, Eggens U, Rudwaleit M, Distler A, et al. Prevalence of spondylarthropathies in HLA-B27 positive and negative blood donors. Arthritis Rheum.1998;41(1):58-67.

12. Hanova P, Pavelka K, Holcatova I, Pikhart H. Incidence and prevalence of psoriatic arthritis, ankylosing spondylitis, and reactive arthritis in the first descriptive population-based study in the Czech Republic. Scand J Rheumatol.2010;39(4):310-7.

13. Palm O, Moum B, Ongre A, Gran JT. Prevalence of ankylosing spondylitis and other spondyloarthropathies among patients with inflammatory bowel disease: a population study (the IBSEN study). J Rheumatol.2002;29(3):511-5.

14. Van den Berg R, de Hooge M, van Gaalen F, Reijnierse M, Huizinga T, van der Heijde D. Percentage of patients with spondyloarthritis in patients referred because of chronic back pain and performance of classification criteria: experience from the Spondyloarthritis Caught Early (SPACE) cohort. Rheumatol Oxf Engl.2013;52(8):1492-9.

15. 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(4):361-8.

16. Feldtkeller E, Bruckel J, Khan MA. Scientific contributions of ankylosing spondylitis patient advocacy groups. Curr Opin Rheumatol.2000;12(4):239-47.

17. Sieper J, Braun J. Ankylosing Spondylitis: In Clinical Practice. Springer Science & Business Media; 2010. 92 p.

18. Amor B, Dougados M, Mijiyawa M. [Criteria of the classification of spondylarthropathies]. Rev Rhum Mal Ostéo-Articul.1990;57(2):85-9.

19. Dougados M, van der Linden S, Juhlin R, Huitfeldt B, Amor B, Calin A, et al. The European Spondylarthropathy Study Group preliminary criteria for the classification of spondylarthropathy. Arthritis Rheum.1991;34(10):1218-27.

20. Rudwaleit M, Landewé R, van der Heijde D, Listing J, Brandt J, Braun J, 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(6):770-6.

21. Rudwaleit M, van der Heijde D, Landewé R, Listing J, Akkoc N, Brandt J, 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(6):777-83.

22. Rudwaleit M, van der Heijde D, Landewé R, Akkoc N, Brandt J, Chou CT, et al. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. Ann Rheum Dis.2011;70(1):25-31.

23. Poddubnyy D, Rudwaleit M, Haibel H, Listing J, Märker-Hermann E, Zeidler H, et al. Rates and predictors of radiographic sacroiliitis progression over 2 years in patients with axial spondyloarthritis. Ann Rheum Dis.2011;70(8):1369-74.

24. ASAS | Slide Library - Results [Internet]. http://slides.asas-group.org/app/slides/search?q=

25. Rudwaleit M, Jurik AG, Hermann K-GA, Landewé R, van der Heijde D, Baraliakos X, 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(10):1520-7.

26. 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(4):1000-8.

27. Cheung PP, Paternotte S, Burki V, Durnez A, Elhai M, Fabreguet I, 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(4):816-21.

28. Boonen A, van der Heijde D, Landewé R, Spoorenberg A, Schouten H, Rutten-van Mölken M, et al. Work status and productivity costs due to ankylosing spondylitis: comparison of three European countries. Ann Rheum Dis.2002;61(5):429-37.

29. Boonen A. A review of work-participation, cost-of-illness and cost-effectiveness studies in ankylosing spondylitis. Nat Clin Pract Rheumatol.2006;2(10):546-53.

30. Wanders A, Heijde D van der, Landewé R, Béhier J-M, Calin A, Olivieri I, et al. Nonsteroidal antiinflammatory drugs reduce radiographic progression in patients with ankylosing spondylitis: a randomized clinical trial. Arthritis Rheum.2005;52(6):1756-65.

31. Van den Bosch F, Kruithof E, Baeten D, De Keyser F, Mielants H, Veys EM. Effects of a loading dose regimen of three infusions of chimeric monoclonal antibody to tumour necrosis factor alpha (infliximab) in spondyloarthropathy: an open pilot study. Ann Rheum Dis.2000;59(6):428-33.

32. Corona-Sanchez EG, Muñoz-Valle JF, Gonzalez-Lopez L, Sanchez-Hernandez JD, Vazquez-Del Mercado M, Ontiveros-Mercado H, et al. -383 A/C tumor necrosis factor receptor 1 polymorphism and ankylosing spondylitis in Mexicans: a preliminary study. Rheumatol Int.2012;32(8):2565-8.

33. Kontoyiannis D, Pasparakis M, Pizarro TT, Cominelli F, Kollias G. Impaired on/off regulation of TNF biosynthesis in mice lacking TNF AU-rich elements: implications for joint and gut-associated immunopathologies. Immunity.1999;10(3):387-98.

34. Van Duivenvoorde LM, Dorris ML, Satumtira N, van Tok MN, Redlich K, Tak PP, et al. Relationship between inflammation, bone destruction, and osteoproliferation in the HLA-B27/human β2 -microglobulin-transgenic rat model of spondylarthritis. Arthritis Rheum.2012;64(10):3210-9.

35. Rudwaleit M, Claudepierre P, Wordsworth P, Cortina EL, Sieper J, Kron M, et al. Effectiveness, safety, and predictors of good clinical response in 1250 patients treated with adalimumab for active ankylosing spondylitis. J Rheumatol.2009;36(4):801-8.

36. Haibel H, Rudwaleit M, Listing J, Heldmann F, Wong RL, Kupper H, et al. Efficacy of adalimumab in the treatment of axial spondylarthritis without radiographically defined sacroiliitis: results of a twelve-week randomized, double-blind, placebo-controlled trial followed by an open-label extension up to week fifty-two. Arthritis Rheum.2008;58(7):1981-91.

37. Sieper J, van der Heijde D, Dougados M, Mease PJ, Maksymowych WP, Brown MA, 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(6):815-22.

38. Van der Heijde D, Dijkmans B, Geusens P, Sieper J, DeWoody K, Williamson P, et al. Efficacy and safety of infliximab in patients with ankylosing spondylitis: results of a randomized, placebo-controlled trial (ASSERT). Arthritis Rheum.2005;52(2):582-91.

39. Braun J, Brandt J, Listing J, Zink A, Alten R, Burmester G, et al. Long-term efficacy and safety of infliximab in the treatment of ankylosing spondylitis: an open, observational, extension study of a three-month, randomized, placebo-controlled trial. Arthritis Rheum.2003;48(8):2224-33.

40. Calin A, Dijkmans B a. C, Emery P, Hakala M, Kalden J, Leirisalo-Repo M, et al. Outcomes of a multicentre randomised clinical trial of etanercept to treat ankylosing spondylitis. Ann Rheum Dis.2004;63(12):1594-600.

41. Davis JC, van der Heijde DM, Braun J, Dougados M, Cush J, Clegg D, et al. Sustained durability and tolerability of etanercept in ankylosing spondylitis for 96 weeks. Ann Rheum Dis.2005;64(11):1557-62.

42. Dougados M, van der Heijde D, Sieper J, Braun J, Maksymowych WP, Citera G, et al. Symptomatic efficacy of etanercept and its effects on objective signs of inflammation in early nonradiographic axial spondyloarthritis: a multicenter, randomized, double-blind, placebo-controlled trial. Arthritis Rheumatol Hoboken NJ.2014;66(8):2091-102.

43. Landewé R, Braun J, Deodhar A, Dougados M, Maksymowych WP, Mease PJ, et al. Efficacy of certolizumab pegol on signs and symptoms of axial spondyloarthritis including ankylosing spondylitis: 24-week results of a double-blind randomised placebo-controlled Phase 3 study. Ann Rheum Dis.2014;73(1):39-47.

44. Sieper J, Landewé R, Rudwaleit M, van der Heijde D, Dougados M, Mease PJ, et al. Effect of Certolizumab Pegol over 96 Weeks in Patients with Axial Spondyloarthritis: Results from a Phase 3 Randomized Trial. Arthritis Rheumatol Hoboken NJ. 2 de2014;

45. Inman RD, Davis JC, Heijde D van der, Diekman L, Sieper J, Kim SI, et al. Efficacy and safety of golimumab in patients with ankylosing spondylitis: results of a randomized, double-blind, placebo-controlled, phase III trial. Arthritis Rheum.2008;58(11):3402-12.

46. Garrett S, Jenkinson T, Kennedy LG, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol.1994;21(12):2286-91.

47. Van der Heijde D, Sieper J, Maksymowych WP, Dougados M, Burgos-Vargas R, Landewé R, et al. 2010 Update of the international ASAS recommendations for the use of anti-TNF agents in patients with axial spondyloarthritis. Ann Rheum Dis.2011;70(6):905-8.

48. Arends S, Brouwer E, van der Veer E, Groen H, Leijsma MK, Houtman PM, et al. Baseline predictors of response and discontinuation of tumor necrosis factor-alpha blocking therapy in ankylosing spondylitis: a prospective longitudinal observational cohort study. Arthritis Res Ther. 2011;13(3):R94.

49. Lord PAC, Farragher TM, Lunt M, Watson KD, Symmons DPM, Hyrich KL, et al. Predictors of response to anti-TNF therapy in ankylosing spondylitis: results from the British Society for Rheumatology Biologics Register. Rheumatol Oxf Engl.2010;49(3):563-70.

50. Vastesaeger N, van der Heijde D, Inman RD, Wang Y, Deodhar A, Hsu B, et al. Predicting the outcome of ankylosing spondylitis therapy. Ann Rheum Dis.2011;70(6):973-81.

51. Calin A, Porta J, Fries JF, Schurman DJ. Clinical history as a screening test for ankylosing spondylitis. JAMA. 13 de1977;237(24):2613-4.

52. Rudwaleit M, Metter A, Listing J, Sieper J, Braun J. Inflammatory back pain in ankylosing spondylitis: a reassessment of the clinical history for application as classification and diagnostic criteria. Arthritis Rheum.2006;54(2):569-78.

53. McGee S. Simplifying likelihood ratios. J Gen Intern Med.2002;17(8):646-9.

54. Soubrier M, Lukas C, Sibilia J, Fautrel B, Roux F, Gossec L, et al. Disease activity score-driven therapy versus routine care in patients with recent-onset active rheumatoid arthritis: data from the GUEPARD trial and ESPOIR cohort. Ann Rheum Dis.2011;70(4):611-5.

55. Gomez-Reino JJ, Maneiro JR, Ruiz J, Roselló R, Sanmarti R, Romero AB, et al. Comparative effectiveness of switching to alternative tumour necrosis factor (TNF) antagonists versus switching to rituximab in patients with rheumatoid arthritis who failed previous TNF antagonists: the MIRAR Study. Ann Rheum Dis.2012;71(11):1861-4.

56. Stone MA, Inman RD, Wright JG, Maetzel A. Validation exercise of the Ankylosing Spondylitis Assessment Study (ASAS) group response criteria in ankylosing spondylitis patients treated with biologics. Arthritis Rheum. 15 de2004;51(3):316-20.

57. Van der Heijde D, Dougados M, Davis J, Weisman MH, Maksymowych W, Braun J, et al. ASsessment in Ankylosing Spondylitis International Working Group/Spondylitis Association of America recommendations for conducting clinical trials in ankylosing spondylitis. Arthritis Rheum.2005;52(2):386-94.

58. Calin A, Garrett S, Whitelock H, Kennedy LG, O'Hea J, Mallorie P, 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(12):2281-5.

59. Dillon CF, Hirsch R. The United States National Health and Nutrition Examination Survey and the epidemiology of ankylosing spondylitis. Am J Med Sci.2011;341(4):281-3.

60. Rudwaleit M, van der Heijde D, Khan MA, Braun J, Sieper J. How to diagnose axial spondyloarthritis early. Ann Rheum Dis.2004;63(5):535-43.

61. Aydin SZ, Maksymowych WP, Bennett AN, McGonagle D, Emery P, Marzo-Ortega H. Validation of the ASAS criteria and definition of a positive MRI of the sacroiliac joint in an inception cohort of axial spondyloarthritis followed up for 8 years. Ann Rheum Dis.2012;71(1):56-60.

62. Les médecins - Estimations au 1er janvier 2008 - Drees - Ministère des Affaires sociales et de la Santé [Internet]. [citado 31 de2015]. Recuperado a partir de: http://www.drees.sante.gouv.fr/les-medecins-estimations-au-1er-janvier-2008,5251.html

63. Underwood MR, Dawes P. Inflammatory back pain in primary care. Br J Rheumatol.1995;34(11):1074-7.

64. Sieper J, Rudwaleit M, Baraliakos X, Brandt J, Braun J, Burgos-Vargas R, et al. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. Ann Rheum Dis.2009;68 Suppl 2:ii1-44.

65. Rudwaleit M, Haibel H, Baraliakos X, Listing J, Märker-Hermann E, Zeidler H, et al. The early disease stage in axial spondylarthritis: results from the German Spondyloarthritis Inception Cohort. Arthritis Rheum.2009;60(3):717-27.

66. Gong Y, Zheng N, Chen S-B, Xiao Z-Y, Wu M-Y, Liu Y, et al. Ten years' experience with needle biopsy in the early diagnosis of sacroiliitis. Arthritis Rheum.2012;64(5):1399-406.

67. Devauchelle-Pensec V, D'Agostino MA, Marion J, Lapierre M, Jousse-Joulin S, Colin D, et al. Computed tomography scanning facilitates the diagnosis of sacroiliitis in patients with suspected spondylarthritis: results of a prospective multicenter French cohort study. Arthritis Rheum.2012;64(5):1412-9.

68. Weber U, Maksymowych WP. Sensitivity and specificity of magnetic resonance imaging for axial spondyloarthritis. Am J Med Sci.2011;341(4):272-7.

69. Weber U, Lambert RGW, Pedersen SJ, Hodler J, Østergaard M, Maksymowych WP. Assessment of structural lesions in sacroiliac joints enhances diagnostic utility of magnetic resonance imaging in early spondylarthritis. Arthritis Care Res.2010;62(12):1763-71.

70. Smith-Bindman R, Lipson J, Marcus R, Kim K-P, Mahesh M, Gould R, et al. Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. Arch Intern Med. 14 de2009;169(22):2078-86.

71. Hermann K-GA, Althoff CE, Schneider U, Zühlsdorf S, Lembcke A, Hamm B, et al. Spinal changes in patients with spondyloarthritis: comparison of MR imaging and radiographic appearances. Radiogr Rev Publ Radiol Soc N Am Inc.2005;25(3):559-69; discussion 569-70.

72. Bollow M, Enzweiler C, Taupitz M, Golder W, Hamm B, Sieper J, et al. Use of contrast enhanced magnetic resonance imaging to detect spinal inflammation in patients with spondyloarthritides. Clin Exp Rheumatol.2002;20(6 Suppl 28):S167-74.

73. Dougados M, d' Agostino M-A, Benessiano J, Berenbaum F, Breban M, Claudepierre P, 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 Rev Rhum.2011;78(6):598-603.

74. La cohorte DESIR [Internet]: http://www.lacohortedesir.fr/

75. Van der Heijde D, Lie E, Kvien TK, Sieper J, Van den Bosch F, Listing J, et al. ASDAS, a highly discriminatory ASAS-endorsed disease activity score in patients with ankylosing spondylitis. Ann Rheum Dis.2009;68(12):1811-8.

76. Jenkinson TR, Mallorie PA, Whitelock HC, Kennedy LG, Garrett SL, Calin A. Defining spinal mobility in ankylosing spondylitis (AS). The Bath AS Metrology Index. J Rheumatol.1994;21(9):1694-8.

77. Jenkinson C, Coulter A, Wright L. Short form 36 (SF36) health survey questionnaire: normative data for adults of working age. BMJ. 29 de1993;306(6890):1437-40.

78. Creemers MCW, Franssen MJ a. M, van't Hof MA, Gribnau FWJ, van de Putte LBA, van Riel PLCM. Assessment of outcome in ankylosing spondylitis: an extended radiographic scoring system. Ann Rheum Dis.2005;64(1):127-9.

79. Hermann K-GA, Baraliakos X, van der Heijde DMFM, Jurik A-G, Landewé R, Marzo-Ortega H, 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(8):1278-88.

80. Moltó A, Paternotte S, Comet D, Thibout E, Rudwaleit M, Claudepierre P, et al. Performances of the Assessment of SpondyloArthritis International Society axial spondyloarthritis criteria for diagnostic and classification purposes in patients visiting a rheumatologist because of chronic back pain: results from a multicenter, cross-sectional study. Arthritis Care Res.2013;65(9):1472-81.

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

82. Sieper J, van der Heijde D, Dougados M, Brown LS, Lavie F, Pangan AL. Early response to adalimumab predicts long-term remission through 5 years of treatment in patients with ankylosing spondylitis. Ann Rheum Dis.2012;71(5):700-6.

83. Ciurea A, Scherer A, Exer P, Bernhard J, Dudler J, Beyeler B, et al. Tumor necrosis factor α inhibition in radiographic and nonradiographic axial spondyloarthritis: results from a large observational cohort. Arthritis Rheum.2013;65(12):3096-106.

84. Lukas C, Landewé R, Sieper J, Dougados M, Davis J, Braun J, et al. Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. Ann Rheum Dis.2009;68(1):18-24.

85. Jones SD, Porter J, Garrett SL, Kennedy LG, Whitelock H, Calin A. A new scoring system for the Bath Ankylosing Spondylitis Metrology Index (BASMI). J Rheumatol.1995;22(8):1609.

86. Daltroy LH, Larson MG, Roberts NW, Liang MH. A modification of the Health Assessment Questionnaire for the spondyloarthropathies. J Rheumatol.1990;17(7):946-50.

87. Dougados M, Simon P, Braun J, Burgos-Vargas R, Maksymowych WP, Sieper 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(2):249-51.

88. Price D, Hillyer EV, van der Molen T. Efficacy versus effectiveness trials: informing guidelines for asthma management. Curr Opin Allergy Clin Immunol.2013;13(1):50-7.

89. Charlton BG. Understanding randomized controlled trials: explanatory or pragmatic? Fam Pract.1994;11(3):243-4.

90. Machado P, Landewé R, Lie E, Kvien TK, Braun J, Baker D, et al. Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores. Ann Rheum Dis.2011;70(1):47-53.

91. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 4 de1983;70(1):41-55.

92. Rubin DB. Estimating causal effects from large data sets using propensity scores. Ann Intern Med. 15 de1997;127(8 Pt 2):757-63.

93. Pham T, Landewé R, van der Linden S, Dougados M, Sieper J, Braun J, et al. An international study on starting tumour necrosis factor-blocking agents in ankylosing spondylitis. Ann Rheum Dis.2006;65(12):1620-5.

94. Dougados M, Braun J, Szanto S, Combe B, Geher P, Leblanc V, et al. Nonsteroidal antiinflammatory drug intake according to the Assessment of SpondyloArthritis International Society Score in clinical trials evaluating tumor necrosis factor blockers: example of etanercept in advanced ankylosing spondylitis. Arthritis Care Res.2012;64(2):290-4.

95. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. N Engl J Med. 22 de2000;342(25):1887-92.

96. Dadoun S, Geri G, Paternotte S, Dougados M, Gossec L. Switching between tumour necrosis factor blockers in spondyloarthritis: a retrospective monocentre study of 222 patients. Clin Exp Rheumatol.2011;29(6):1010-3.

97. Luc M, Gossec L, Ruyssen-Witrand A, Salliot C, Duclos M, Guignard S, et al. C-reactive protein predicts tumor necrosis factor-alpha blocker retention rate in axial ankylosing spondylitis. J Rheumatol.2007;34(10):2078-81.

98. Busquets N, Tomero E, Descalzo MÁ, Ponce A, Ortiz-Santamaría V, Surís X, et al. Age at treatment predicts reason for discontinuation of TNF antagonists: data from the BIOBADASER 2.0 registry. Rheumatol Oxf Engl.2011;50(11):1999-2004.