







ORIGINAL RESEARCH

Heterogeneity of axial
spondyloarthritis: genetics, sex and
structural damage matter

Zhixiu Li ^{1,2} Sjeff M van der Linden ^{3,4} Muhammad Asim Khan ⁵
Heinz Baumberger,⁶ Hermine van Zandwijk,⁷ Mohammad Kazim Khan ⁸
Peter M Villiger ^{9,10} Matthew A Brown ^{11,12}

To cite: Li Z, van der Linden SM, Khan MA, *et al.* Heterogeneity of axial spondyloarthritis: genetics, sex and structural damage matter. *RMD Open* 2022;**8**:e002302. doi:10.1136/rmdopen-2022-002302

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/rmdopen-2022-002302>).

Deceased

Received 22 February 2022
Accepted 6 April 2022

ABSTRACT

Objective Axial spondyloarthritis (axSpA) comprises both radiographic and non-radiographic disease. However, the paucity of specific objective measures for the disease and current classification criteria showing suboptimal specificity contribute to disease heterogeneity observed in clinical practice and research. We used a historical cohort of patients with axSpA to assess sources of heterogeneity. **Methods** The study involved 363 axSpA probands recruited from membership of the Swiss Ankylosing Spondylitis Patient Society. Participants underwent examination by a rheumatologist, completed questionnaires and provided blood samples for HLA typing. Patients underwent radiography of sacroiliac joints and were categorised according to the New York (NY) criteria (ankylosing spondylitis (AS) or non-radiographic axSpA (nr-axSpA)) and HLA-B27 status. Genetic characterisation by single nucleotide polymorphism microarray was performed and AS polygenic risk scores (PRS) were calculated.

Results Considerable heterogeneity was observed. The male to female ratio for AS (NY+) was 3:1, but 1:1 for nr-axSpA. For HLA-27(+) AS, the ratio was 2.5:1, but nearly 1:1 for HLA-B27(−) disease. Women with nr-axSpA had strikingly lower mean PRS and lower HLA-B27 prevalence than men with nr-axSpA or NY(+) male and female patients with AS. PRS was able to distinguish male but not female patients with nr-axSpA from related healthy first-degree relatives. Radiographic sacroiliitis was strongly associated with HLA-B27, especially in men.

Conclusion Women clinically diagnosed with axSpA but without radiographic sacroiliitis as a group have a disease that is distinct from AS by the modified New York criteria overall and from nr-axSpA in men. Given the high degree of heterogeneity, stratified or adjusted analysis of effectiveness studies is indicated, taking genetics, sex and radiographic damage (sacroiliitis) into account.

INTRODUCTION

Ankylosing spondylitis (AS) is quite a common and frequently familial inflammatory rheumatic disorder that is largely genetically determined and strongly associated with the HLA-B27 allele. The estimates of the prevalence of ‘radiographic AS’, as defined by

Key messages

What is already known about this subject?

- ⇒ Ankylosing spondylitis (AS), now called axial spondyloarthritis (axSpA), comprises both radiographic and non-radiographic disease (nr-axSpA).
- ⇒ The new broader axSpA concept encompasses a heterogeneous group of conditions.

What does this study add?

- ⇒ Women with nr-axSpA as a group have a disease that is distinct from AS by the modified New York criteria and from men with nr-axSpA.

How might this impact on clinical practice or further developments?

- ⇒ To be meaningful, clinical axSpA studies should be analysed taking genetics, sex and structural damage into consideration.
- ⇒ Findings are also relevant for patient education and counselling.

the modified New York (mNY) criteria, range between 0.1% and 0.4%, mostly influenced by the frequency of this genetic factor in the population.^{1–8}

During the last few decades, the concept of AS has widened. It has been realised that radiographic sacroiliitis by no means is an early or obligatory manifestation of the disease.^{9–10} This has led to the concept of axial spondyloarthritis (axSpA), comprising both radiographic AS and non-radiographic axSpA (nr-axSpA). Nowadays, the disease is often called axSpA, a term we use to clearly refer to the full spectrum of the disease. The ASAS has developed the criteria for axSpA as a single disorder.^{11–14} These criteria are primarily intended for classification purposes and were not developed or validated for use in the diagnosis of individual patients. However, in daily practice the distinction might be less clear. Current criteria lack sensitivity and in



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Professor Matthew A Brown;
matt.brown@genomicsengland.co.uk

particular specificity, although the ongoing CLASSIC study, which aims to improve the performance of the ASAS criteria, is underway.¹³ In the absence of a gold standard for a disease, look-alike conditions may give rise to false-positive diagnoses and induce or increase heterogeneity among patients diagnosed with conditions such as axSpA. Clinical heterogeneity might cause unwanted disparities in aetiopathogenesis, prognosis, outcome and response to treatment with conventional medications or treatment with biologicals.^{15 16} In this study we aim to assess potential sources of heterogeneity by using data from a family study that originated in 1985, that is, long before the concept of axSpA was coined. This implies that our data cannot be confounded by the introduction of the more recent classification criteria.

METHODS

The study presented here is based on the findings of the 1985 baseline study and the 2019 follow-up study.

The baseline study required five steps. First, in 1985, all members of Schweizerische Vereinigung Morbus Bechterew, the nationwide Swiss Ankylosing Spondylitis Patient Society, and their first-degree relatives (FDR) were invited to participate in a family study that was performed in centres spread all over Switzerland. Informed consent of participants was obtained. Before becoming a member of the patient society, for all 363 probands the diagnosis AS (now axSpA) had been established by a Swiss rheumatologist. A total of 1178 persons consented to participate and completed questionnaires on disease manifestations. Second, the clinical diagnosis was established based on the clinical history and examination by the rheumatologist at the study centre who was blinded to any radiographic findings. All participants underwent physical examination of their axial and peripheral joints. Third, blood samples were drawn for HLA-A, HLA-B and HLA-C typing and peripheral blood nucleated cells (PBNCs) were stored in liquid nitrogen. Fourth, pelvic radiographs were taken to assess the presence of sacroiliitis. Consenting non-pregnant participants aged ≥ 18 years underwent pelvic radiography unless a recent radiograph was available. Fifth, sacroiliac (SI) joints were scored according to the mNY criteria. All 1081 pelvic radiographs of 360 probands and 713 FDR and 8 spouses were assessed twice by each of four experienced readers, that is, a total of eight (sometimes nine) blinded readings for each SI joint. This could be performed only once for 46% of the 360 radiographs of the probands and 3% of the 713 radiographs of the FDR because these radiographs were only available on-site for a few hours at the time of participants' physical examination at the local hospital. All the readers were unaware of participants' clinical findings and HLA-B27 status. Overall, 17.2% of 1081 radiographs were read once, 0.4% two to four times, 3.2% five to seven times and 79.2% eight to nine times. The sacroiliitis score ranged from 0 (normal) to 4 (ankylosis) for each SI joint assessment by a reader as per the

mNY scoring system.¹⁷ Scores for a single SI joint were added and divided by the number of assessments (range 1–9). Scores below bilateral grade 2.0 were considered not fulfilling the mNY criteria, as did below unilateral grade 3.0 sacroiliitis. Therefore, probands were categorised according to the mNY criteria as AS if these criteria were met and as nr-axSpA if the radiographic mNY criteria were not fulfilled.¹⁷ Please note that this classification was done long before the introduction of the ASAS criteria or the availability of MRI scanning.^{11 12} Furthermore, at the group level, mean SI grades were calculated by sex, HLA-B27 and radiographic status. Interobserver and intraobserver reliability were assessed for five observers by evaluating a subset of 243 pelvic films. Observers read the films twice in sets of 40–50 radiographs. The interval between both readings was ≥ 7 days.¹⁸

The 2019 follow-up study comprised three additional steps. First, in January 2018, the ethics committee approved the follow-up study. Second, participants (n=485, including 125 probands) who had provided written informed consent to use their PBNCs for genetic analysis were mailed a postal questionnaire on their health status. Third, calculating polygenic risk scores (PRS), the DNA of 679 participants (226 probands and 453 healthy FDR) of the baseline study was extracted from PBNCs (stored since 1985) and genotyping was performed using the Illumina CoreExome single nucleotide polymorphism (SNP) microarray, as previously reported.^{19 20} The SNP genotype data were then used to calculate individual PRS, using a model developed in European-descent AS case and healthy control cohorts.²⁰ In brief, the PRS is a numeric score that reflects an individual's estimated genetic predisposition for a given trait and can be used as a predictor or diagnostic biomarker of the disease of interest (here AS/axSpA). SNP data from 14 337 unrelated European-descent healthy controls also genotyped using Illumina CoreExome arrays as previously reported were used as *unrelated* healthy control data,²⁰ and unaffected FDR (after 35 years of follow-up) were used as *related* healthy controls. SPSS was used to perform area under the curve (AUC) calculation and statistical analysis.

Patient and public involvement

Two patients/coauthors were fully involved in the study.

RESULTS

Altogether 363 axSpA probands participated in the family study (table 1 and figure 1). We report on heterogeneity among probands due to genetics (HLA-B27 status and PRS), sex and severity, defined as structural damage to the SI joints (by the mNY criteria). Occurrence of axSpA among FDR is reported elsewhere.²¹ The mean age of the 249 male and 114 female probands was 44.19 \pm 11.1 years. Table 1 shows separately for men and women the number of probands by presence or absence of radiographic damage of the SI joints (by the mNY criteria)

Table 1 Demographic data and clinical features of 363 axSpA probands by radiographic status of sacroiliac joints and HLA-B27 status of men and women

Probands	Total (n)	Age, 1985, mean (SD)	CIBP (%)	Grade of single SI joint, mean (SE)	Male:female ratio
axSpA total	363*	44.2 (11.1)	72.8	2.7 (0.07)	2.18
Male	249	45.2 (11.7)	74.5	3.0 (0.07)	
Female	114	42.1 (9.4)	69.0	2.1 (0.12)	
Radiographic axSpA (AS)	273	44.6 (11.4)	73.4	3.3 (0.04)	2.96
Male	204	45.6 (11.7)	75.7	3.4 (0.05)	
Female	69	41.8 (10.1)	66.7	3.0 (0.08)	
Non-radiographic axSpA	87	42.9 (10.1)	69.8	0.9 (0.07)	0.98
Male	43	42.9 (11.9)	67.4	1.0 (0.10)	
Female	44	42.9 (8.1)	72.1	0.8 (0.09)	
HLA-B27(+) axSpA	308	44.0 (11.1)	72.9	2.9 (0.06)	2.46
Male	219	45.2 (11.5)	74.7	3.1 (0.07)	
Female	89	41.1 (9.5)	67.4	2.5 (0.12)	
AS	247	44.6 (11.6)	73.1	3.3 (0.05)	2.98
Male	185	45.7 (11.8)	75.4	3.4 (0.05)	
Female	62	41.3 (10.2)	66.1	3.0 (0.09)	
nr-axSpA	58	41.7 (8.5)	69.0	1.1 (0.08)	1.23
Male	32	42.3 (9.4)	68.8	1.0 (0.11)	
Female	26	41.0 (7.4)	69.2	1.1 (0.11)	
HLA-B27(-) axial axSpA	50	46.1 (11.5)	69.4	1.7 (0.21)	1.08
Male	26	46.3 (14.0)	73.1	2.3 (0.29)	
Female	24	46.0 (8.4)	77.1	0.9 (0.23)	
AS	22	46.9 (10.1)	72.7	3.1 (0.18)	2.67
Male	16	46.7 (11.0)	75.0	3.3 (0.22)	
Female	6	47.3 (8.4)	66.7	2.7 (0.25)	
nr-axSpA	28	45.6 (12.7)	74.1	0.5 (0.10)	0.56
Male	10	45.6 (18.5)	70.0	0.8 (0.24)	
Female	18	45.6 (8.6)	75.5	0.3 (0.08)	

The prevalence of CIBP by the Calin criteria and the mean single SI joint grade are provided at the group level.

*Unknown radiographic status: 3 probands; unknown HLA-B27 status: 5 probands, including 1 nr-axSpA proband.

AS, ankylosing spondylitis; axSpA, axial spondyloarthritis; CIBP, chronic inflammatory back pain; nr-axSpA, non-radiographic axial spondyloarthritis.

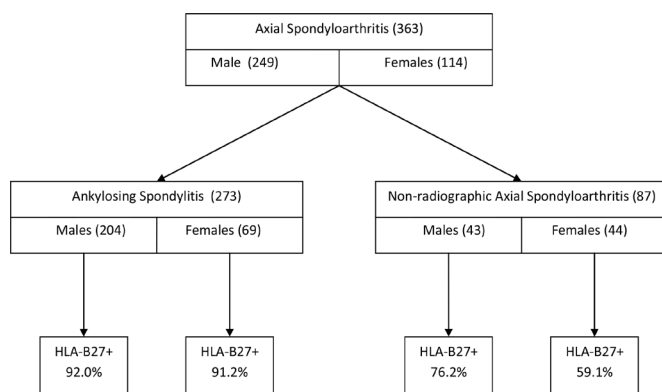


Figure 1 Flow chart of 363 probands with axial spondyloarthritis by radiographic status, gender and HLA-B27 prevalence. The radiographic status is unknown for three probands.

and by HLA-B27 status. The table also shows for each group the prevalence of chronic inflammatory back pain (by the Calin criteria²²) and the mean grade of sacroiliitis (mNY grading). Considering reading of the pelvic radiographs, the interobserver and intraobserver reliability coefficients were 0.865 and 0.903, respectively.

There are remarkable differences in the male to female ratios (figures 1 and 2 and table 1). Among those with AS, the sex ratio is about 3:1, but for probands categorised as nr-axSpA it is about 1:1 (AS 204men/69 women vs nr-axSpA 43men/44 women; $p=0.00001$, $OR=3.03$, 95% CI 1.83 to 4.99) (table 1). Therefore, there is a significant association between male sex and presence of sacroiliitis by the mNY criteria. Within the mNY(+) group there is no significant association between sex and HLA-B27 status (HLA-B27(+) 185men/62 women vs HLA-B27(-) 16men/6 women; $p=0.82$, $OR=1.12$, 95% CI 0.42

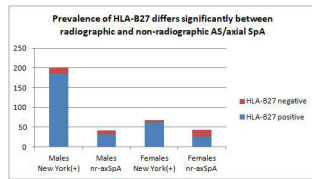


Figure 2 The figure shows that a higher proportion of patients with AS, defined by the radiographic New York criteria for sacroiliitis, are men. For nr-axSpA the ratio is about equal. Among women with nr-axSpA the prevalence of HLA-B27(–) disease is considerably higher compared with men and women with AS (by the New York criteria) and men with nr-axSpA. AS, ankylosing spondylitis; nr-axSpA, non-radiographic axial spondyloarthritis; SpA, spondyloarthritis.

to 2.99). The same holds for the nr-axSpA group (HLA-B27(+) 32 men/26 women vs HLA-B27(–) 10 men/18 women; $p=0.09$, $OR=2.22$, 95% CI 0.87 to 5.62) (table 1). In the nr-axSpA group the sex ratio for HLA-B27(+) men and women is 1.2:1 (32 men and 26 women), whereas it is 0.6:1 (10 men and 18 women) among HLA-B27(–) patients (table 1 and figures 1 and 2). Furthermore, among the probands with nr-axSpA, the prevalence of the HLA-B27 allele is non-significantly lower among women compared with men (26/44 or 59.1% vs 32/43 or 74.4%, $p=0.13$).

Irrespective of the radiographic status, there is a significant association between the presence of HLA-B27 and male sex, as 217 of 243 (89.3%) male probands are HLA-B27(+) compared with 88 of 112 (78.6%) female probands ($p=0.007$, $OR=2.28$, 95% CI 1.24 to 4.18).

The sex ratio is 2.5:1 for all 308 HLA-B27(+) probands (men/women 219/89) and about equal (1.1:1) for all 50 HLA-B27(–) axSpA probands (men/women 26/24) (table 1). Among all 355 probands with known HLA-B27 and radiographic status, women relatively more often have nr-axSpA than men (44/112 or 39.3% vs 42/243 or 17.3%; $p=0.00001$, $OR=3.10$, 95% CI 1.87 to 5.13).

Table 2 provides the HLA-B27 carriage status and the mean PRS for axSpA, by sex and radiographic mNY status. Table 3, figures 3 and 4, and online supplemental figures 1 and 2 provide a comparison of HLA-B27 status and PRS values and discriminatory performance in terms of AUC in receiver operator characteristic analyses. Considering initially the clinical diagnosis of axSpA, the PRS was significantly lower for women than for men (0.293 vs 0.372, $p=0.043$). In contrast, no difference was observed among AS cases (0.418 vs 0.377, $p=0.28$; table 2). In patients with nr-axSpA, however, a higher proportion of women were HLA-B27(–) (women 16/32 vs men 5/24, $p=0.030$ Fisher's exact test), and women had a lower PRS (0.104 vs 0.322, $p=0.029$) (table 2). Overall, PRS was lower in patients with nr-axSpA than those with AS (0.198 vs 0.387, $p=0.000004$), but this difference was restricted to women with nr-axSpA (0.104 vs 0.418 in women with AS, $p=0.00021$), with the PRS of male patients with nr-axSpA being no different from those with AS (0.322 vs 0.377 in those with AS, $p=0.29$).

Consistent with this, the PRS was able to distinguish female AS from female nr-axSpA cases ($AUC=0.711$, $p=0.00085$), but not male AS from male nr-axSpA cases

Table 2 PRS and HLA-B27 status of male and female patients with axSpA and AS (by the modified New York criteria) and patients with non-radiographic axSpA, their unaffected healthy FDR and unrelated healthy controls

		n (female/male)	PRS		
			Female	Male	Overall
axSpA	Overall	226 (75/151)	0.293 (0.341)	0.372 (0.236)	0.346 (0.277)
	HLA-B27(+)	190 (56/134)	0.455 (0.176)	0.440 (0.135)	0.444 (0.148)
	HLA-B27(–)	36 (19/17)	–0.187 (0.245)	–0.162 (0.177)	–0.175 (0.213)
AS	Overall	168 (42/126)	0.418 (0.205)	0.377 (0.210)	0.387 (0.209)
	HLA-B27(+)	153 (39/114)	0.450 (0.161)	0.429 (0.126)	0.435 (0.136)
	HLA-B27(–)	15 (3/12)	–0.001 (0.285)	–0.121 (0.195)	–0.097 (0.209)
nr-axSpA	Overall	56 (32/24)	0.104 (0.38)	0.322 (0.331)	0.198 (0.373)
	HLA-B27(+)	35 (16/19)	0.43 (0.14)	0.476 (0.141)	0.455 (0.14)
	HLA-B27(–)	21 (16/5)	–0.221 (0.23)	–0.263 (0.056)	–0.231 (0.202)
Related healthy controls	Overall	453 (243/210)	0.087 (0.330)	0.054 (0.320)	0.072 (0.326)
	HLA-B27(+)	187 (106/81)	0.426 (0.157)	0.417 (0.145)	0.422 (0.152)
	HLA-B27(–)	266 (137/129)	–0.175 (0.129)	–0.173 (0.138)	–0.174 (0.133)
Unrelated healthy controls	Overall	14261 (6278/7983)	–0.222 (0.196)	–0.222 (0.194)	–0.221 (0.195)
	HLA-B27(+)	1197 (506/691)	0.298 (0.159)	0.304 (0.159)	0.301 (0.159)
	HLA-B27(–)	12901 (5694/7207)	–0.272 (0.106)	–0.269 (0.109)	–0.271 (0.108)

Note that gender for unrelated healthy controls and HLA-B27 status for all participants are imputed from genotype information. Overall sample sizes for different clinical subgroups are thus greater than the total of HLA-B27 positive and negative subsets.

AS, ankylosing spondylitis; axSpA, axial spondyloarthritis; FDR, first-degree relative; nr-axSpA, non-radiographic axial spondyloarthritis; PRS, polygenic risk score.

Table 3 Comparison of HLA-B27 and PRS of different axSpA groups, overall and separately for men and women

Cases	Gender	Versus nr-axSpA						Versus related healthy controls						Versus unrelated healthy controls					
		HLA-B27			PRS			HLA-B27			PRS			HLA-B27			PRS		
		AUC	P-AUC	P-PRS	AUC	P-AUC	P-PRS	AUC	P-AUC	P-PRS	AUC	P-AUC	P-PRS	AUC	P-AUC	P-PRS	AUC	P-AUC	P-PRS
axSpA	Overall	-	-	-	0.714	0	2.73×10 ⁻²⁵	0.878	0	0.920	0	0	0.878	0	0.920	0	0	0	0
	Female	-	-	-	0.655	0.000010	0.0000004	0.833	0	0.869	0	0	0.833	0	0.869	0	0	0	0
	Male	-	-	-	0.751	0	4.69×10 ⁻²²	0.900	0	0.946	0	0	0.900	0	0.946	0	0	0	0
AS	Overall	0.636	0.0028	0.00010	0.749	0	1.22×10 ⁻²⁸	0.913	0	0.960	0	0	0.913	0	0.960	0	0	0	0
	Female	0.707	0.0011	0.00085	0.746	7.63×10 ⁻¹³	0.923	0	0.969	0	0	0.923	0	0.969	0	0	0	0	0
	Male	0.552	0.428	0.477	0.75	0.55	4.74×10 ⁻²¹	0.909	0	0.958	0	0	0.909	0	0.958	0	0	0	0
nr-axSpA	Overall	-	-	-	0.606	0.0076	0.0077	0.770	4.53×10 ⁻¹²	0.797	1.09×10 ⁻¹³	8.11×10 ⁻⁵⁷	0.770	4.53×10 ⁻¹²	0.797	1.09×10 ⁻¹³	8.11×10 ⁻⁵⁷	0.797	1.09×10 ⁻¹³
	Female	-	-	-	0.532	0.56	0.79	0.709	0.00016	0.734	0.000072	7.27×10 ⁻²¹	0.709	0.00016	0.734	0.000072	7.27×10 ⁻²¹	0.734	0.000072
	Male	-	-	-	0.703	0.00013	0.00053	0.852	3.58×10 ⁻¹³	0.881	0	2.91×10 ⁻⁴¹	0.852	3.58×10 ⁻¹³	0.881	0	2.91×10 ⁻⁴¹	0.881	0

AUC and p values for comparisons of AUC between groups are provided for both HLA-B27 and PRS (P-AUC), and for PRS p values from independent sample t-tests between groups (P-PRS). P=0 refers to P<10⁻¹⁰⁰. AS, ankylosing spondylitis; AUC, area under the curve; axSpA, axial spondyloarthritis; nr-axSpA, non-radiographic axial spondyloarthritis; PRS, polygenic risk score.

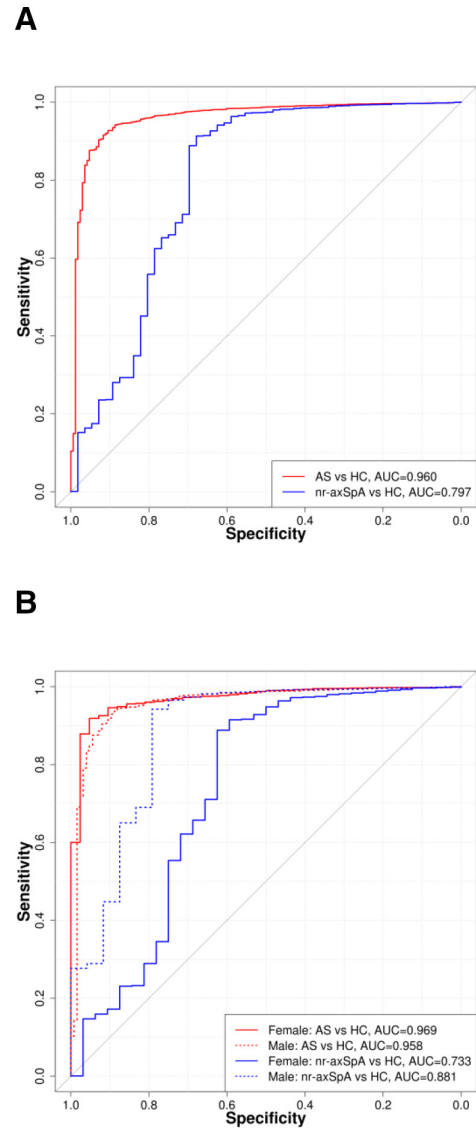


Figure 3 (A) Discriminatory performance of PRS among probands with AS and nr-axSpA compared with HC. (B) Discriminatory performance of PRS among female and male probands with AS and nr-axSpA compared with HC. AS, ankylosing spondylitis; AUC, area under the curve; HC, healthy unrelated control; nr-axSpA, non-radiographic axial spondyloarthritis; PRS, polygenic risk score.

(AUC=0.477, p=0.75). The PRS was not able to distinguish female nr-axSpA cases from female healthy FDR (AUC=0.498, p=0.98), whereas the PRS had good discriminatory capacity considering male nr-axSpA cases and male healthy FDR (AUC=0.717, p=0.00053). The AUC for female nr-axSpA compared with either related or unrelated female healthy controls was significantly lower than for male nr-axSpA (p=0.013 and p=0.049, respectively), whereas for AS the scores were not significantly different. Among patients with AS, the PRS had good discriminatory capacity in relation to healthy FDR (AUC of 0.772 overall, 0.774 in women and 0.786 in men) and excellent discrimination compared with unrelated healthy controls (AUC of 0.960 overall, 0.969 in women and 0.958 in men). This indicates that women

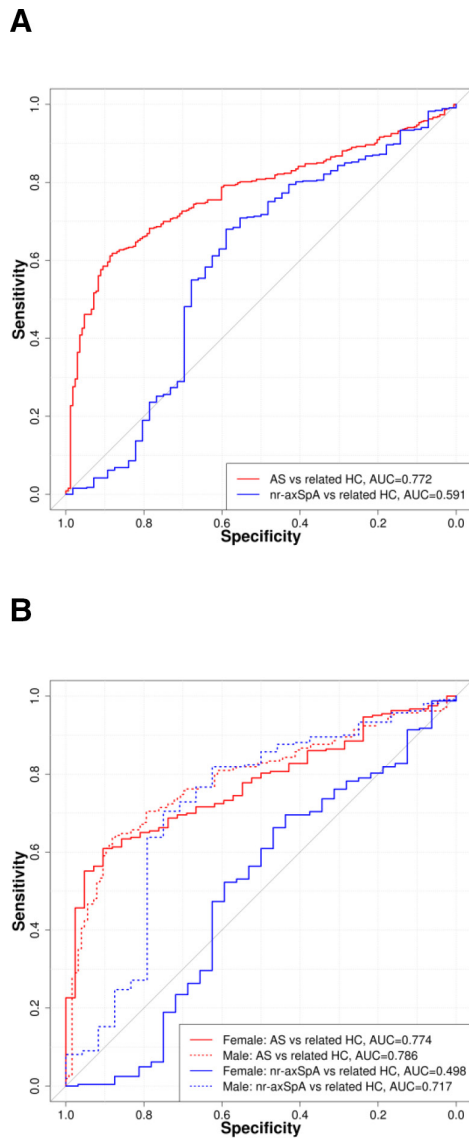


Figure 4 (A) Discriminatory performance of PRS among probands with AS and nr-axSpA compared with related healthy controls (FDR). (B) Discriminatory performance of PRS among female and male probands with AS and nr-axSpA compared with related healthy controls (FDR). AS, ankylosing spondylitis; AUC, area under the curve; FDR, first-degree relative; HC, healthy unrelated control; nr-axSpA, non-radiographic axial spondyloarthritis; PRS, polygenic risk score.

clinically diagnosed with axSpA but who do not have radiographic sacroiliitis as a *group* have a disease that is distinct from AS by the mNY criteria overall and from nr-axSpA in men. In comparisons with unrelated healthy controls, the PRS performed significantly better than HLA-B27 for axSpA and AS (axSpA PRS AUC=0.869, HLA-B27 AUC=0.833, $p=0.028$; AS-PRS AUC 0.969, HLA-B27 AUC=0.923, $p=0.0022$), but not for nr-axSpA (PRS AUC=0.734, HLA-B27 AUC=0.709, $p=0.63$) (table 3). This indicates that the non-HLA-B27 component of the PRS was greater in patients with AS than in patients with nr-axSpA.

DISCUSSION

The new and wider concept of axSpA, comprising both AS and nr-axSpA, is important and clinically relevant. Radiographic sacroiliitis is by no means an early or obligatory manifestation of the disease. There is, however, a drawback that the wider concept increases heterogeneity of what we might consider as a single disease. Radiographic and nr-axSpA are reportedly not different regarding health status, disease activity and physical function, but they did differ in signs of inflammation; all these signs were found to be higher in patients with AS.²³

Our study, starting long before the new concept was coined, shows a sizeable (~25%) proportion of probands (index patients with the disease) who have what we now call nr-axSpA (table 1 and figure 1). Of note, at baseline the diagnosis axSpA (in 1985 'AS') was made clinically by the rheumatologist at the study centre. The findings on the pelvic radiograph of the participants were used for classification according to the mNY criteria, not for diagnosis. Although the concept axSpA was not yet known in 1985, the notion of 'AS without radiographic sacroiliitis' already existed.⁹ Establishing the diagnosis has always been quite possible without definite radiographic findings, for example, according to the Rome criteria (if four of the five clinical criteria are met).²⁴ We strongly feel that the low mean SI scores in the several nr-axSpA groups (table 1) truly represent non-radiographic disease.

Our findings clearly show heterogeneity regarding the sex ratio, with a male to female ratio of 3:1 for AS by the mNY criteria, but a 1:1 ratio for nr-axSpA. We also noted considerable differences in PRS and HLA-B27 association between radiographic and nr-axSpA. The observed heterogeneity has potentially important clinical consequences. For example, due to heterogeneity it might be difficult, if not impossible, to develop appropriately performing classification criteria for the whole group with sufficient sensitivity and high specificity. Our findings suggest that including genetic analyses, such as PRS, provides a potential solution to this issue.

The observed genetic differences (PRS and association with HLA-B27) are also certainly relevant regarding differences in heredity. Recurrence of the disease is high among the offspring of HLA-B27(+) parents with AS, but rare in HLA-B27(-) families.^{8 21 25} In particular, disease recurrence among the offspring of female HLA-B27(+) probands is substantial.²⁵ According to the threshold model of polygenic inheritance, the mechanism by which a continuous distribution of genetic risk leads to dichotomous trait or disease states,²⁶ the genetic threshold for women to get the disease is increased relative to men, and women would be predicted to have higher PRS. In the current study women with AS have higher PRS than men (0.418 vs 0.377), although this does not reach statistical significance. This may in turn translate to a higher proportion of affected children of mothers with AS.

In addition to demonstrating genetic heterogeneity, particularly involving nr-axSpA, this study shows for the first time the high discriminatory capacity of an AS-PRS

for men with nr-axSpA (AUC=0.881 compared with unrelated healthy controls, $p < 10^{-100}$). While this AUC was lower than that observed for men with AS also compared with healthy controls, it indicates that some of nr-axSpA probands who have increased PRS, which was developed using genotype data from patients with AS, will proceed to meet the radiographic mNY criteria. The study also provides further confirmation of the high discriminatory performance of AS-PRS in European-descent populations and confirms that it performs better than HLA-B27 testing alone (for AS overall, AS-PRS AUC=0.96, HLA-B27 AUC=0.913).

Heterogeneity of cohorts defined by classification criteria can significantly affect both basic and clinical research projects and, when those criteria are also employed as diagnostic criteria, adversely influence the accuracy of prediction of prognosis and treatment responses. There is considerable evidence that genetics and sex influence the clinical features of the disease. Male patients with AS have more extensive radiographic change,^{27–29} whereas female patients with AS have higher self-reported disease activity, similar functional incapacity and lower C-reactive protein (CRP) levels (reviewed in Rusman *et al.*³⁰). In cohorts with nr-axSpA, women have been shown to have lower prevalence of objective MRI evidence of SI inflammation.³¹ HLA-B27(+) patients have earlier disease onset and are more likely to develop acute anterior uveitis (reviewed in Akkoç *et al.*³² and Brown *et al.*³³). Evidence from the literature indicates that the heterogeneity of what we call one disease affects the efficacy of response to treatment with biologicals. In real-world observational studies of treatment of nr-axSpA with biologicals, significantly lower response rates are found among women than among men³⁴ and shorter retention on biological treatment consistent with lower efficacy.^{35–37} The results are in line with randomised controlled trials of adalimumab, golimumab and certolizumab in nr-axSpA, showing lower response rates in women compared with men.^{38–40} Our data here, demonstrating genetic differences between axSpA patient subsets, suggest that genetics might also be used as a tool to assess and possibly predict the efficacy of treatment of axSpA, in particular of nr-axSpA, consistent with the known lower response rates of tumour necrosis factor inhibitor therapy in HLA-B27(–) AS.⁴¹ Among women with clinical features of axSpA but negative SI radiographs, the ability of the AS-PRS to distinguish between women with nr-axSpA and AS suggests that PRS may be particularly helpful in distinguishing those with true inflammatory back pain from those with other non-axSpA conditions causing similar symptoms.

Our study design has certain limitations. We address (1) the validity and confidence of clinical diagnosis, (2) the possible consequences of time frame, and (3) the accuracy of distinguishing AS (radiographic axSpA from nr-axSpA).

1. Validity and confidence of the diagnosis AS (now axSpA). The proband's diagnosis was confirmed twice,

initially by their local rheumatologist and subsequently again at the study centre, where the rheumatologist confirmed the diagnosis blinded to radiographic and HLA-B27 status. Based on the SI scores, probands were then categorised as AS or nr-axSpA. Objective documentation of inflammation of the SI joints is lacking for the nr-axSpA cases (MRI was not yet available in 1985). This might cause lower diagnostic confidence for nr-axSpA probands compared with AS cases. Indeed, the *likelihood* of a false-positive diagnosis of nr-axSpA seems higher than of AS. This translates to increased heterogeneity among nr-axSpA. However, in our view, it is unlikely that MRI of SI joints would have eliminated this source of heterogeneity. As we and others have shown, this technique lacks as yet—certainly outside experts' centres—sensitivity and specificity. A substantial proportion of healthy individuals without current or past back pain have an MRI positive for sacroiliitis according to the ASAS definition.^{42–43} Further, in recreational and elite athletes, MRI revealed bone marrow edema changes meeting the ASAS definition of active sacroiliitis in 30%–41% of subjects.^{44–45}

2. Possible consequences of time frame. The current spectrum of axSpA might differ from that in 1985. Although AS without radiographic sacroiliitis (now nr-axSpA) was already known⁹ at that time, awareness of the condition has risen considerably. This might have widened the spectrum of the disease and increased the likelihood of inclusion of look-alike conditions. This again tends to increase heterogeneity. Therefore, we may conclude that both lower diagnostic confidence for nr-axSpA cases as well as increased awareness of the disease might contribute to observed heterogeneity. This underpins the need to adjust the analysis of axSpA outcomes.

3. The radiographic heterogeneity of axSpA is associated with genetic differences and differences in heredity (discussed in the Discussion above). Therefore, accurate distinction between radiographic disease and nr-axSpA is important. However, there is as yet no standardised method to assess pelvic radiographs for the presence of sacroiliitis. For example, it is unclear how many readings by how many readers are needed to obtain high-quality results or how does one reach consensus avoiding *His Master's Voice* bias. Intraobserver and interobserver reliability do not guarantee sufficient sensitivity and specificity. In the past, we have reported moderate sensitivity and specificity in reading pelvic radiographs for sacroiliitis: sensitivity (84.3%/79.8%) and specificity (70.6%/74.7%) for 23 radiologists and 100 rheumatologists, respectively. Training did not improve overall performance.⁴⁶ In the current study, pelvic films were assessed up to nine times by up to four readers. The mean score was calculated for each SI joint. According to the central limit theory, the mean represents the true value of a certain measure better than single observations. Our quantitative scoring allowed assessing how 'free of sacroiliitis' patients

with nr-axSpA really are (table 1). That means, the nr-axSpA group comprises truly patients with nr-axSpA (they do not have borderline sacroiliitis). Consensus judgement would not have enabled quantitative assessment. This finding supports good performance regarding sensitivity. Furthermore, we used the absence of sacroiliitis among all HLA-B27(-) FDR (of HLA-B27(+) probands) as the gold standard for sufficient specificity. Therefore, this post-hoc evaluation is in our view reassuring. However, for future studies we strongly suggest standardised blinded central reading by qualified readers with known sensitivity and specificity in assessment of sacroiliitis. Evaluation should include randomly inserted control radiographs of patients with AS and unaffected persons and assessment of observer reliability.

In conclusion, given the demonstrated heterogeneity of axSpA, clinical studies such as clinical trials should take genetics, sex and structural damage (radiographic sacroiliitis) into consideration, either in study design or retrospectively, for example by applying stratified or otherwise adjusted analysis. This heterogeneity also needs to be considered in clinical assessment and management of patients with axSpA. *One size does not fit all.*

Author affiliations

¹Faculty of Health, School of Biomedical Sciences, Queensland University of Technology (QUT), Brisbane, Queensland, Australia

²Centre for Genomics and Personalised Health, Queensland University of Technology (QUT), Brisbane, Queensland, Australia

³University of Bern, Bern, Switzerland

⁴Department of Internal Medicine, University of Maastricht, Maastricht, The Netherlands

⁵Case Western Reserve University School of Medicine, Cleveland, Ohio, USA

⁶SVMB, Flims, Switzerland

⁷Rheumatology Research, Mortroux, Belgium

⁸Department of Mathematical Sciences, Kent State University, Kent, Ohio, USA

⁹Medical Center Monbijou, Meikirch, Switzerland

¹⁰Rheumatology and Clinical Immunology, Medical Center Monbijou, Meikirch, Switzerland

¹¹Department of Medical and Molecular Genetics, King's College London, London, UK

¹²Genomics England, London, UK

Acknowledgements We like to thank all patients, spouses and relatives for their kind cooperation and are grateful to the Swiss city and village administrations for retrieving current addresses. We also thank Hans-Ueli Rentsch* MD, Hans Valkenburg* MD, Arnold Cats* MD, Herman Kroon MD and Niklaus Gerber MD for their contributions to performing the study. Caroline Kaegi provided helpful secretarial assistance. *deceased.

Contributors All authors had full access to all of the data in the study and take responsibility for the data and the accuracy of the data analysis. Concept and design: SMvdL, MAK, HB, HvZ and MAB conceived the study. Designing the questionnaire: SMvdL, HB and HvZ. Drafting of the manuscript: SMvdL, MAK and MAB. Genetic analysis: ZL and MAB. Statistical analysis: SMvdL, ZL, MKK and MAB. Obtained funding: SMvdL and MAB. Administrative, technical or material support: HvZ and PMV. Acquisition of data: SMvdL, HvZ and PMV. Critical revision of the manuscript for important intellectual content: all authors. MAB acts in the role of guarantor.

Funding The 1985 baseline study was funded by the Swiss National Fund, Schweizer Rück Insurance and Ciba-Geigy, Switzerland. The 2018 follow-up study was funded/supported by the National Institute for Health and Care Research (NIHR) Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London and/or the NIHR Clinical Research Facility.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study involves human participants and was approved by the University Hospital of Bern and the Ethical Committee of the Kanton Bern, Switzerland (#2017-00536). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Zhixiu Li <http://orcid.org/0000-0002-2924-9120>

Sjef M van der Linden <http://orcid.org/0000-0002-0087-0351>

Muhammad Asim Khan <http://orcid.org/0000-0003-4704-8311>

Mohammad Kazim Khan <http://orcid.org/0000-0001-9857-0578>

Peter M Villiger <http://orcid.org/0000-0002-8859-9964>

Matthew A Brown <http://orcid.org/0000-0003-0538-821>

REFERENCES

- de Blecourt J, Polman A, de Blecourt BLECOURT-Meindersma MEINDERSMA T. Hereditary factors in rheumatoid arthritis and ankylosing spondylitis. *Ann Rheum Dis* 1961;20:215–23.
- Bakland G, Nossent HC, Gran JT. Incidence and prevalence of ankylosing spondylitis in northern Norway. *Arthritis Rheum* 2005;53:850–5.
- De Angelis R, Salaffi F, Grassi W. Prevalence of spondyloarthropathies in an Italian population sample: a regional community-based study. *Scand J Rheumatol* 2007;36:14–21.
- Exarchou S, Lindström U, Askling J, et al. The prevalence of clinically diagnosed ankylosing spondylitis and its clinical manifestations: a nationwide register study. *Arthritis Res Ther* 2015;17:118.
- Haglund E, Bremander AB, Petersson IF, et al. Prevalence of spondyloarthritis and its subtypes in southern Sweden. *Ann Rheum Dis* 2011;70:943–8.
- Trontzas P, Andrianakos A, Miyakis S, et al. Seronegative spondyloarthropathies in Greece: a population-based study of prevalence, clinical pattern, and management. The ESORDIG study. *Clin Rheumatol* 2005;24:583–9.
- Stolwijk C, van Onna M, Boonen A, et al. Global prevalence of spondyloarthritis: a systematic review and meta-regression analysis. *Arthritis Care Res* 2016;68:1320–31.
- van der Linden SM, Valkenburg HA, de Jongh BM, et al. The risk of developing ankylosing spondylitis in HLA-B27 positive individuals. A comparison of relatives of spondylitis patients with the general population. *Arthritis Rheum* 1984;27:241–9.
- Khan MA, van der Linden SM, Kushner I, et al. Spondylitic disease without radiologic evidence of sacroiliitis in relatives of HLA-B27 positive ankylosing spondylitis patients. *Arthritis Rheum* 1985;28:40–3.
- Rudwaleit M, Khan MA, Sieper J. The challenge of diagnosis and classification in early ankylosing spondylitis: do we need new criteria? *Arthritis Rheum* 2005;52:1000–8.
- Rudwaleit M, van der Heijde D, Landewé R, et al. The development of assessment of spondyloarthritis International Society classification criteria for axial spondyloarthritis (Part II): validation and final selection. *Ann Rheum Dis* 2009;68:777–83.
- Rudwaleit M, Landewé R, van der Heijde D, et al. The development of assessment of spondyloarthritis International Society classification criteria for axial spondyloarthritis (Part I): classification of paper patients by expert opinion including uncertainty appraisal. *Ann Rheum Dis* 2009;68:770–6.

- 13 Khan MA, van der Linden S. Axial spondyloarthritis: a better name for an old disease: a step toward uniform reporting. *ACR Open Rheumatol* 2019;1:336–9.
- 14 Sieper J, van der Heijde D. Review: Nonradiographic axial spondyloarthritis: new definition of an old disease? *Arthritis Rheum* 2013;65:543–51.
- 15 Robinson PC, Wordsworth BP, Reveille JD, et al. Axial spondyloarthritis: a new disease entity, not necessarily early ankylosing spondylitis. *Ann Rheum Dis* 2013;72:162–4.
- 16 van der Linden S, Akkoc N, Brown MA, et al. The ASAS criteria for axial spondyloarthritis: strengths, weaknesses, and proposals for a way forward. *Curr Rheumatol Rep* 2015;17:62.
- 17 van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the new York criteria. *Arthritis Rheum* 1984;27:361–8.
- 18 Cronbach L, Gleser G, Nanda H. *The dependability of behavioural measurements: theory of generalisability for scores and profiles*. New York: Wiley, 1972.
- 19 Huang X-F, Li Z, De Guzman E, et al. Genomewide association study of acute anterior uveitis identifies new susceptibility loci. *Invest Ophthalmol Vis Sci* 2020;61:3.
- 20 Li Z, Wu X, Leo PJ. Polygenic risk scores have high diagnostic capacity in ankylosing spondylitis. *Ann Rheum Dis* 2021.
- 21 van der Linden SM, Khan MA, Li Z, et al. Factors predicting axial spondyloarthritis among first-degree relatives of probands with ankylosing spondylitis: a family study spanning 35 years. *Ann Rheum Dis* 2022. doi:10.1136/annrheumdis-2021-222083. [Epub ahead of print: 11 Mar 2022].
- 22 Calin A, Porta J, Fries JF, et al. Clinical history as a screening test for ankylosing spondylitis. *JAMA* 1977;237:2613.
- 23 Kiltz U, Baraliakos X, Karakostas P, et al. Do patients with non-radiographic axial spondylarthritis differ from patients with ankylosing spondylitis? *Arthritis Care Res* 2012;64:1415–22.
- 24 Kellgren J, Jeffrey M, Ball J. *The epidemiology of chronic rheumatism*. Blackwell: Oxford, 1963.
- 25 Calin A, Brophy S, Blake D. Impact of sex on inheritance of ankylosing spondylitis: a cohort study. *Lancet* 1999;354:1687–90.
- 26 Fisher RA. XV.—The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh* 1919;52:399–433.
- 27 Cortes A, Maksymowych WP, Wordsworth BP, et al. Association study of genes related to bone formation and resorption and the extent of radiographic change in ankylosing spondylitis. *Ann Rheum Dis* 2015;74:1387–93.
- 28 Ward MM, Hendrey MR, Malley JD, et al. Clinical and immunogenetic prognostic factors for radiographic severity in ankylosing spondylitis. *Arthritis Rheum* 2009;61:859–66.
- 29 Lee W, Reveille JD, Davis JC, et al. Are there gender differences in severity of ankylosing spondylitis? results from the psoas cohort. *Ann Rheum Dis* 2007;66:633–8.
- 30 Rusman T, van Vollenhoven RF, van der Horst-Bruinsma IE. Gender differences in axial spondyloarthritis: women are not so lucky. *Curr Rheumatol Rep* 2018;20:35.
- 31 Rusman T, John M-LB, van der Weijden MAC, et al. Presence of active MRI lesions in patients suspected of non-radiographic axial spondyloarthritis with high disease activity and chance at conversion after a 6-month follow-up period. *Clin Rheumatol* 2020;39:1521–9.
- 32 Akkoç N, Yarkan H, Kenar G, et al. Ankylosing spondylitis: HLA-B*27-Positive versus HLA-B*27-Negative disease. *Curr Rheumatol Rep* 2017;19:26.
- 33 Brown MA, Li Z, Cao K-AL. Biomarker development for axial spondyloarthritis. *Nat Rev Rheumatol* 2020;16:448–63.
- 34 Neuenschwander R, Hebeisen M, Micheroli R, et al. Differences between men and women with nonradiographic axial spondyloarthritis: clinical characteristics and treatment effectiveness in a real-life prospective cohort. *Arthritis Res Ther* 2020;22:233.
- 35 Glintborg B, Sørensen IJ, Østergaard M, et al. Ankylosing spondylitis versus Nonradiographic axial spondyloarthritis: comparison of tumor necrosis factor inhibitor effectiveness and effect of HLA-B27 status. An observational cohort study from the nationwide DANBIO registry. *J Rheumatol* 2017;44:59–69.
- 36 Gulfe A, Kapetanovic MC, Kristensen LE. Efficacy and drug survival of anti-tumour necrosis factor-alpha therapies in patients with non-radiographic axial spondyloarthritis: an observational cohort study from southern Sweden. *Scand J Rheumatol* 2014;43:493–7.
- 37 Ørnbjerg LM, Brahe CH, Askling J, et al. Treatment response and drug retention rates in 24 195 biologic-naïve patients with axial spondyloarthritis initiating TNFi treatment: routine care data from 12 registries in the EuroSpA collaboration. *Ann Rheum Dis* 2019;78:1536–44.
- 38 Sieper J, van der Heijde D, Dougados M, et al. Efficacy and safety of adalimumab in patients with non-radiographic axial spondyloarthritis: results of a randomised placebo-controlled trial (ABILITY-1). *Ann Rheum Dis* 2013;72:815–22.
- 39 Sieper J, van der Heijde D, Dougados M, et al. A randomized, double-blind, placebo-controlled, sixteen-week study of subcutaneous golimumab in patients with active nonradiographic axial spondyloarthritis. *Arthritis Rheumatol* 2015;67:2702–12.
- 40 Landewé R, Braun J, Deodhar A, et al. Efficacy of certolizumab pegol on signs and symptoms of axial spondyloarthritis including ankylosing spondylitis: 24-week results of a double-blind randomised placebo-controlled phase 3 study. *Ann Rheum Dis* 2014;73:39–47.
- 41 Rudwaleit M, Claudepierre P, Wordsworth P, et al. Effectiveness, safety, and predictors of good clinical response in 1250 patients treated with adalimumab for active ankylosing spondylitis. *J Rheumatol* 2009;36:801–8.
- 42 Weber U, Lambert RGW, Østergaard M, et al. The diagnostic utility of magnetic resonance imaging in spondylarthritis: an international multicenter evaluation of one hundred eighty-seven subjects. *Arthritis & Rheumatism* 2019;62:3048–58.
- 43 Ye L, Liu Y, Xiao Q, et al. Mri compared with low-dose CT scanning in the diagnosis of axial spondyloarthritis. *Clin Rheumatol* 2020;39:1295–303.
- 44 de Winter J, de Hooge M, van de Sande M, et al. Magnetic resonance imaging of the Sacroiliac joints indicating sacroiliitis according to the assessment of spondyloarthritis International Society definition in healthy individuals, runners, and women with postpartum back pain. *Arthritis Rheumatol* 2018;70:1042–8.
- 45 Weber U, Jurik AG, Zejden A, et al. Frequency and anatomic distribution of magnetic resonance imaging features in the sacroiliac joints of young athletes: exploring "background noise" toward a data-driven definition of sacroiliitis in early spondyloarthritis. *Arthritis Rheumatol* 2018;70:736–45.
- 46 van Tubergen A, Heuft-Dorenbosch L, Schulp G, et al. Radiographic assessment of sacroiliitis by radiologists and rheumatologists: does training improve quality? *Ann Rheum Dis* 2003;62:519–25.