RMD Open
Rheumatic & Musculoskeletal Diseases

ORIGINAL RESEARCH

Change in different classes of chronic back pain suspicious of axial spondyloarthritis: a latent transition analysis of the SPACE cohort

Philipp Bosch ,[1,2,3] Alexandre Sepriano ,[2,4] Mary-Lucy Marques ,[2,5] Désirée van der Heijde ,[2] Robert Landewé ,[6,7] Miranda van Lunteren ,[2] Liese de Bruin,[2] Manouk de Hooge ,[8] Caroline Bastiaenen,[3] Sofia Exarchou,[9] Roberta Ramonda ,[10] Karen Minde Fagerli ,[11] Floris A van Gaalen ,[2] Sofia Ramiro [2,7]

To cite: Bosch P, Sepriano A, Marques M-L, *et al*. Change in different classes of chronic back pain suspicious of axial spondyloarthritis: a latent transition analysis of the SPACE cohort. *RMD Open* 2024;**10**:e004584. doi:10.1136/rmdopen-2024-004584

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/rmdopen-2024-004584).

Received 29 May 2024
Accepted 9 September 2024

For numbered affiliations see end of article.

Correspondence to
Dr Sofia Ramiro;
sofiaramiro@gmail.com

ABSTRACT

Objectives To follow up four previously identified classes 'pure axial spondyloarthritis' (axSpA) ('axial'), 'axSpA with peripheral signs' ('inflammatory back pain+peripheral'), 'axSpA at risk' and 'no spondyloarthritis' ('no SpA'). They reflect the expert-opinion-free construct or 'Gestalt' of chronic back pain suspicious of axSpA. The aim was to assess participants' transitions between these classes over time.

Methods Participants with chronic back pain of ≤2 years duration, suspicious of axSpA from the SPondyloArthritis Caught Early cohort were analysed. Latent class (LCA) and latent transition analysis (LTA) using clinical, laboratory and imaging data at baseline and 2 years were calculated. Conditional and marginal probabilities were obtained, reflecting the probability of a spondyloarthritis feature in a class and the probability of the participant's class membership, respectively. Transitional probabilities were extracted revealing potential switches across classes. The analyses were performed in all participants using imputations for missing data and in participants with full data at baseline and 2 years.

Results Baseline and 2 years LCA models were constructed for 702 participants, resulting in the same four-class model as previously described. LTA revealed only a 3% transition from the 'no SpA' to the 'at-risk' class from baseline to 2 years with all other participants remaining in their initially assigned class. Sensitivity analysis on 384 participants with complete data at both baseline and 2 years showed similar results, underlining the model's robustness.

Conclusions Transitions between the four classes over 2 years were basically inexistent, highlighting the unlikelihood of developing new class-defining features of axSpA after an initial clinical workup.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Four classes comprising patients with axial involvement, peripheral involvement, risk factors for axSpA and chronic back pain without additional features, reflect the 'Gestalt' of chronic back pain suspicious of axSpA.

WHAT THIS STUDY ADDS

⇒ Patients categorised to one of these classes are unlikely to switch to a different class over a 2-year period.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ New features altering the 'Gestalt' of chronic back pain, which is suspicious of axSpA, are unlikely to develop after an initial clinical workup, reassuring rheumatologists about the reliability of their initial findings and conclusions.

INTRODUCTION

The clinical construct or 'Gestalt' of axial spondyloarthritis (axSpA) has been shaped over the last decades by a better understanding of its clinical presentation, imaging findings at the sacroiliac joints (SIJ) and spine, inflammatory markers in the blood, as well as its association with human leucocyte antigen-B27 (HLA-B27) and familial predisposition.[1] This construct is reflected in the Assessment of SpondyloArthritis international Society (ASAS) classification criteria developed to homogenise axSpA populations for clinical trials.[2–4] A limitation of these (and other) classification criteria is that, in the absence of a definite gold standard, expert opinion has been used both for the development and validation of the criteria, which introduces the potential for circular reasoning.[5 6]

To tackle this issue, our group has previously analysed participants of two inception cohorts of axSpA, namely the SPondyloArthritis Caught Early (SPACE) and DEvenir des Spondyloarthropathies Indifférenciées Récentes (DESIR) cohorts, to find underlying, hidden classes of participants.[7] Using a data-driven approach the role of expert opinion in the categorisation process of the cohorts' participants was minimised, leading to the mostly circularity-free identification of classes of participants.

Three classes were identified in both cohorts and labelled as follows: 'pure axSpA' ('axial') representing individuals with positive axial imaging and a familial/genetic predisposition for spondyloarthritis (SpA); 'axSpA with peripheral signs' ('inflammatory back pain (IBP)+peripheral') with participants characterised by peripheral involvement and IBP and 'axSpA at risk', a class with a familial/genetic predisposition to SpA but an otherwise low probability of other features. A fourth class 'no spondyloarthritis' ('no SpA') was found only in SPACE depicting participants with a very low probability for any SpA-associated feature. This is in line with the underlying population of that cohort which includes patients with axSpA and with non-axSpA chronic back pain, while in DESIR only patients with a clinical diagnosis of axSpA were included.

To assess the class membership and potential switch across these classes over time, a latent transition analysis (LTA) was performed in the DESIR cohort revealing only an 11% probability of switching from the 'at risk' to the 'IBP+peripheral' class, with no other switches, over a period of 5 years. This finding was important as it showed the reliability and stability of the classes with participants being unlikely to develop new class-defining features of axSpA after their first clinical assessment.

In SPACE, such an analysis was not possible as follow-up data were still being collected. However, a recent analysis revealed that SPACE participants diagnosed with axSpA during an initial assessment were reliably considered to have the disease after 2 years, while participants without an initial axSpA diagnosis were unlikely to receive one after this follow-up period.[8] Whether this stability of diagnosis also extends to the four classes is uncertain.

Now, with the 2 years data available, this analysis aimed to follow up the previously described classes and assess a potential switch in class membership over time.

METHODS

Participants and study design

This is a longitudinal analysis, for which baseline and 2 years data of participants from the SPACE cohort were used. This multinational cohort has been previously described in detail.[8 9] Briefly, participants aged ≥16 years with chronic back pain lasting ≥3 months and ≤2 years, with an initial onset of pain before 45 years of age, were eligible for inclusion and were assessed at a baseline visit. After the baseline visit, only participants fulfilling at least one of the following criteria were eligible for follow-up: (a) the presence of at least one major SpA feature (positive likelihood ratio (LR) of ≥6.0 for axSpA diagnosis): HLA-B27 positivity, positive family history of SpA, sacroiliitis (MRI or radiographs) or acute anterior uveitis or (b) the presence of two minor SpA features (positive LR of ≥2.5 and <6.0 for axSpA diagnosis): IBP, heel pain, peripheral arthritis, psoriasis, inflammatory bowel disease (IBD), good response to non-steroidal anti-inflammatory drugs (NSAID) or elevated levels of acute phase reactants (C reactive protein (CRP) or erythrocyte sedimentation rate).[10] All patients included in SPACE were included in this analysis and the database was locked in July 2023.

Spondyloarthritis features and data imputations

Imaging data were collected at baseline, 1 and 2 years while clinical data were additionally also collected 3 months after baseline. As previously described,[7] the variables for the analyses to identify the latent classes included imaging features: sacroiliitis on MRI of the SIJ (MRI-SIJ, ASAS definition), bone marrow oedema (BME) on MRI of the spine (MRI-Spine, ≥5 lesions), definitive damage on radiographs of the SIJ (X-SIJ) according to modified New York criteria, ≥1 syndesmophyte on radiographs of the spine (X-Spine)[11–15] and clinical features: IBP (ASAS definition), good response to NSAID, peripheral arthritis, dactylitis, heel pain, family history of SpA, HLA-B27 status, psoriasis, uveitis, IBD and elevated CRP (>5 mg/L).[3 16]

MRI-SIJ, X-SIJ and X-Spine were read by three independent readers. MRI-spine was read by two independent readers and an adjudicator in case of disagreement. All readers were blinded to chronology, clinical data and the results of other modalities.[7 9]

At baseline, all features were either positive if 'ever present' or negative if 'never present', except for dactylitis which was either 'currently present' or 'currently not present'. For the main analysis, if a clinical feature was absent at baseline, it was imputed using data from the 3-month visit if available, except for 'good response to NSAID', for which data could also be imputed from the 1-year visit if not available in a prior visit. Subsequently, if data were still missing, it was imputed with zero, assuming the absence of that feature (this was the case for 5% of all imaging and 0.2% of all clinical features).

For the 2 years follow-up visit, the 'once a feature always a feature' principle was applied: If a feature was positive at baseline, the 3-month or 1-year follow-up visit, it was also considered positive at the 2 years follow-up visit, even if it was negative or missing at that time point. If a feature was negative at baseline and also negative or missing at the follow-up visits, it was considered negative at the 2 years follow-up visit.

Statistical analysis

The statistical analysis including a detailed description of the method has been previously described.[7 17] As a first step, a latent class analysis (LCA) was performed on the

Table 1 Patient characteristics

Characteristic	N=702 Baseline	2-year follow-up
Age, years	30 (8)	n.a.
Male	271 (39%)	n.a.
Body mass index (kg/m2)	25 (5)	n.a.
Duration of back pain (months)	13 (7)	n.a.
Imaging features of SpA		
Sacroiliitis on MRI-SIJ (ASAS)	140 (20%)	179 (25%)
Definitive damage on X-SIJ (mNY)	24 (3%)	32 (5%)
BME on MRI-Spine (≥5 lesions)	15 (2%)	15 (2%)
≥1 syndesmophyte on X-Spine	64 (9%)	83 (12%)
Clinical features of SpA		
Peripheral arthritis ever	104 (15%)	123 (18%)
Dactylitis ever	43 (6%)	70 (10%)
Heel pain ever	147 (21%)	175 (25%)
Psoriasis ever	85 (12%)	100 (14%)
Uveitis ever	57 (8%)	76 (11%)
Inflammatory bowel disease ever	47 (7%)	53 (8%)
Inflammatory back pain ever	490 (70%)	547 (78%)
Elevated CRP (>5 mg/L)	191 (27%)	261 (37%)
HLA-B27 positive	314 (45%)	314 (45%)
Family history of SpA (ASAS)	312 (44%)	333 (47%)
Good response to NSAID ever	250 (36%)	414 (59%)

Mean (SD) or n (%).
Data for BMI available for 673 (96%) participants, for duration of back pain for 696 (99%) participants.
Missing values for imaging and clinical features of SpA have been imputed as described in the 'Methods' section.
ASAS, Assessment of SpondyloArthritis international Society; BME, bone marrow oedema; CRP, C reactive protein; HLA-B27, human leucocyte antigen B27; mNY, modified New York criteria; n.a., not applicable; NSAID, non-steroidal anti-inflammatory drug; SIJ, sacroiliac joints; SpA, spondyloarthritis; X-SIJ, radiograph of the sacroiliac joints; X-Spine, radiograph of the spine.

baseline and, separately, on the 2 years follow-up data of all participants of the SPACE cohort. This included the 465 participants from the previous analysis and new participants who had been included in the cohort since then.[7] LCA was performed, only as an intermediate step, to identify the number of classes at baseline and follow-up that would be used for the final LTA model (eg, a 4–4 class model represents an LTA model with four classes both at baseline and follow-up). Comparisons between LTA and LCA models were not performed as the latter is not suitable to analyse longitudinal data. The number of classes was determined by means of goodness-of-fit parameters including Akaike's information criterion, Bayesian information criterion (BIC), sample size adjusted BIC (for all three lower values mean better model fit), as well as entropy (range 0–1, values close to one reflect a better model accuracy) and a log-likelihood ratio test (comparing the model with the model with n-1 classes), as well as by clinically recognisable patterns within each class (meaning the classes need to make sense clinically). The number of classes was increased one by one until the best-fitting, clinically recognisable model was found. The classes were interpreted and labelled according to their conditional probabilities (ie, the probability of a feature being present in one of the classes) and marginal probabilities (ie, the probability of a participant being in one of the classes) and participants were individually categorised to one of the classes based on their posterior probability of class membership (with the class having the highest probability for each patient determining their assignment).

LCA was performed once with HLA-B27 status as a SpA feature, and once without it, to compare differences in the conditional and marginal probabilities of the two models. This was done because, for the subsequent LTA, HLA-B27 status was omitted as a variable due to its time-invariant nature (ie, its value remains constant over time).

Finally, an LTA was performed by using the number of classes at baseline and 2 years follow-up based on the best fitting, clinically most recognisable LCA models and the class-(in)variance according to clinical reasoning. The latter means that classes at baseline and follow-up were considered to have the same meaning. In line with the LCA, the LTA model reveals underlying latent classes of the patient cohort, but in addition, it takes the presence of different SpA features both at baseline and 2 years into consideration. Model fitness was again compared against models with fewer or more classes at baseline and 2 years using goodness-of-fit parameters as described above (except for a log-likelihood ratio which is not available for LTA) and clinically recognisable patterns. As before, conditional and marginal probabilities were extracted from the model and individual participants were categorised into one of the classes at baseline and then again at follow-up. Transitional probabilities were extracted from the model representing the probability of a participant categorised to a class at baseline to switch to another class at the 2 years follow-up visit. Age and sex were included separately as covariates, to study their individual impact on transitions between classes.

To minimise selection bias, all participants with a baseline visit (with or without a 2-year follow-up) were included in the main analysis and missing data were imputed as described above. Additionally, a sensitivity analysis was performed only including participants with complete data for all measures of interest at both time points.

Demographic data were analysed both for all participants and for the participants in the respective class. All analyses were performed with MPlus V.8.9 and R V.4.3.0, in particular using the package 'MplusAutomation'.[18]

Table 2 Classes of participants with chronic back pain suspicious of axial spondyloarthritis identified in the 2-year latent transition analysis model (n=702)

	Class 1 'axial' (p*=0.18, N†=130)	Class 2 'IBP+peripheral' (p*=0.16, N†=110)	Class 3 'at risk' (p*=0.29, N†=204)	Class 4 'no SpA' (p*=0.37, N†=258)
Sacroiliitis on MRI-SIJ (ASAS)	0.97	0.22	0.03	0.01
BME on MRI-Spine (≥5 lesions)	0.09	0.01	0.00	0.01
Definitive damage on X-SIJ (mNY)	0.14	0.01	0.01	0.02
≥1 syndesmophyte on X-Spine	0.09	0.19	0.09	0.09
Elevated CRP (>5 mg/L)	0.50	0.43	0.28	0.22
Good response to NSAID ever	0.62	0.71	0.43	0.33
Peripheral arthritis ever	0.11	0.65	0.02	0.08
Dactylitis ever	0.05	0.35	0.03	0.01
Heel pain ever	0.15	0.76	0.14	0.10
Family history of SpA (ASAS)	0.45	0.51	1.00	0.00
Psoriasis ever	0.10	0.47	0.08	0.04
Uveitis ever	0.14	0.11	0.11	0.06
Inflammatory bowel disease ever	0.02	0.08	0.06	0.10
Inflammatory back pain ever	0.81	0.96	0.72	0.62

Conditional probabilities (ie, the probability of a feature being present in one of the classes, range: 0–1) were obtained using a latent transition analysis model with full invariance on baseline and 2-year data. Full invariance means that these probabilities are the same at baseline and follow-up.
Cells are coloured in green whenever the conditional probability is ≥0.3. This cut-off was chosen to better visualise differences between the classes.
*Marginal probability of the latent class (ie, a participant's probability of class membership).
† Participants categorised to one of the classes based on their posterior probability of class membership (with the class having the highest probability for each patient determining their assignment).
ASAS, Assessment of SpondyloArthritis international Society; BME, bone marrow oedema; CRP, C reactive protein; IBP, inflammatory back pain; mNY, modified New York criteria; NSAID, non-steroidal anti-inflammatory drug; SIJ, sacroiliac joints; SpA, spondyloarthritis; X-SIJ, radiograph of the sacroiliac joints; X-Spine, radiograph of the spine.

RESULTS

Participant characteristics

702 participants were included. Table 1 shows baseline and 2 years characteristics of all participants. At baseline, the most frequent imaging feature of SpA was sacroiliitis on MRI-SIJ (20%) and the most frequent clinical feature IBP (70%). The frequency of all features increased from baseline to the 2 years follow-up visit, following the described 'once a feature always a feature' principle. BME on MRI-Spine was the only imaging variable that did not increase in frequency from baseline to follow-up. More pronounced increases were seen in the features good response to NSAID (36%–59%), elevated CRP (27%–37%) and IBP (70%–78%), from baseline to the 2 years follow-up. For all other features, the increase was minimal (≤5%).

Latent class analysis

LCA performed on the baseline and 2 years follow-up data resulted in the same four clinically recognisable classes as previously described.[7] This included the 'axial' class with the highest probability of positive axial imaging and a familial/genetic SpA predisposition, the 'IBP+peripheral' class exhibiting a high probability of peripheral involvement and IBP, the 'at-risk' class, demonstrating a high probability of a familial/genetic SpA predisposition exclusively and a 'no SpA' class with a low probability of any SpA-associated features (online supplemental tables S1 and S2).

The four classes were also apparent in the LCA models at baseline and 2 years follow-up when excluding HLA-B27 status. Excluding the variable increased the probability of a positive family history in the 'no SpA' class at baseline (40% vs 22%). Additionally, without HLA-B27 status, fewer participants were classified as 'axial' at follow-up (9% vs 34%), but with a higher probability of sacroiliitis on MRI-SIJ (100% vs 51%) compared with the follow-up model with the genetic marker (online supplemental tables S3 and S4).

On a more technical note, the four-class model at baseline without HLA-B27 status fitted the data best, while for the baseline model including HLA-B27 status and the follow-up models, with and without the genetic marker, model fitness of the four-class model was not superior to the three or five class models. However, the latter two did not reveal clinically recognisable patterns (online supplemental tables S5–S8).

Latent transition analysis

The LTA performed on baseline and follow-up data revealed similar goodness-of-fit parameters (online supplemental table S9) for all models but recognisable clinical patterns only for the 4–4 class model. The

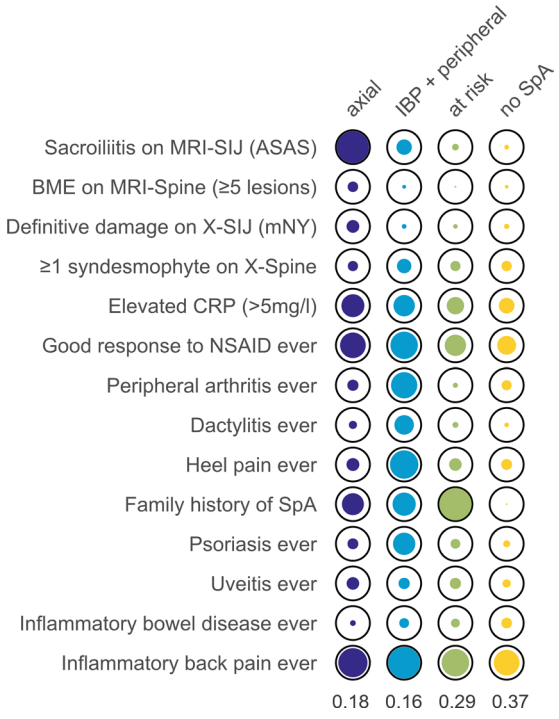

Figure 1 Graphical demonstration of the conditional and marginal probabilities of the 2-year latent transition analysis (LTA) model (n=702). The circles represent the conditional probability for a feature in a respective class, with a higher probability corresponding to a fuller circle. A full circle represents 100% and an empty circle 0% probability. The colours represent the four classes. The numerical values for conditional and marginal probabilities are reported in table 2. The numbers in the last row represent the marginal probabilities, that is, the percentage of participants that according to the LTA belong to one of the classes. ASAS, Assessment of SpondyloArthritis international Society; BME, bone marrow oedema; CRP, C reactive protein; IBP, inflammatory back pain; mNY, modified New York criteria; NSAID, non-steroidal anti-inflammatory drug; SIJ, sacroiliac joints; X-SIJ, radiograph of the sacroiliac joints; SpA, spondyloarthritis; X-Spine, radiograph of the spine.

conditional and marginal probabilities for the final model are shown in table 2 and figure 1. The conditional probabilities of the SpA feature a good response to NSAID, IBP and family history of SpA were high (≥30%) throughout all classes except for family history in the 'no SpA' class which was 0%.

The 'axial' class, had a distinct, high probability of sacroiliitis on MRI-SIJ (97%) and elevated CRP values (50%); the 'IBP+peripheral' class, the highest probability of IBP (96%) and peripheral/cutaneous SpA features (including peripheral arthritis, dactylitis, heel pain and psoriasis); the 'at-risk' class the highest probability of family history of SpA (100%), but otherwise a lack of distinct SpA features; and the 'no SpA' class a negative family history of SpA and overall the lowest probability for any SpA feature.

LTA revealed no switch from the 'axial', 'IBP+peripheral' or 'at-risk' class at baseline to any other class at the 2 years follow-up. The 'no SpA' participants at baseline

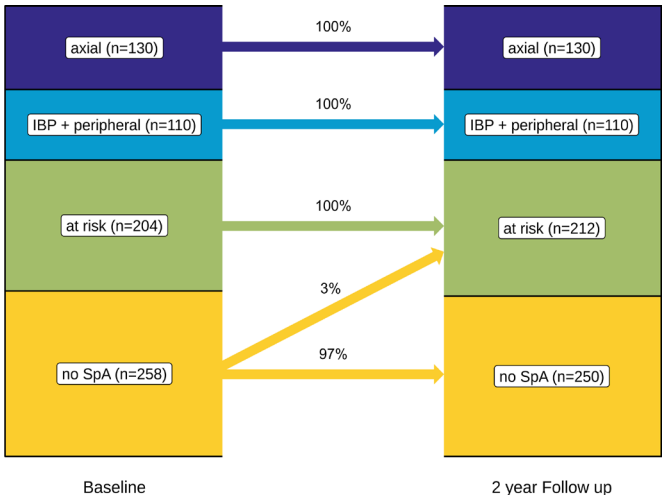

Figure 2 Diagram showing class change over 2 years according to transitional probabilities (LTA analysis). Transitional probabilities were generated using the 4-class latent transition model with 702 patients. IBP, inflammatory back pain; LTA, latent class analysis; SpA, spondyloarthritis.

had a 3% probability to switch to the 'at-risk' class at the 2 years follow-up (figure 2).

The odds of transitioning from the 'no SpA' class to the 'at-risk' class, relative to staying in the 'no SpA' class, were not significantly different for sex (OR (95%CI) for females vs males: 0.28 (0.03 to 2.39)), while age did not have a significant impact on the transition either (OR (95% CI) per 1-year increase: 0.96 (0.88 to 1.05)).

Observed baseline characteristics per class

Observed baseline characteristics per class were similar to the model's conditional probabilities as expected (table 3). Participants in the 'axial' class at baseline showed the lowest mean (SD) age at 28[7] years and the highest percentage of males (69%), while participants in the 'IBP+peripheral', 'at-risk' and 'no SpA' classes were predominantly female with 58%, 72% and 65%, respectively, and slightly older than the participants in the axial class.

Sensitivity LTA

A sensitivity LTA was performed on 384 participants who had full data of all relevant SpA features at baseline and 2 years follow-up. The 4–4 class model exhibited similar goodness-of-fit parameters (online supplemental table S10), the same clinically recognisable classes and similar conditional and marginal probabilities as the main analysis (table 4 and online supplemental figure 1). The sensitivity LTA revealed again no switch from the 'axial', 'IBP+peripheral' or 'at-risk' class at baseline to any other class at follow-up, but a 7% class switch from 'no SpA' at baseline to 'at risk' at the 2 years follow-up (online supplemental figure S2). Observed characteristics per class were also similar to the main analysis (online supplemental table S11).

Table 3 Observed baseline characteristics per latent class (n=702) after data imputation

Characteristic	Class 1 'axial' (n=130, 18%)	Class 2 'IBP+peripheral' (n=110, 16%)	Class 3 'at risk' (n=204, 29%)	Class 4 'no SpA' (n=258, 37%)
Age, years	28 (7)	31 (8)	30 (8)	32 (8)
Male	77 (59%)	46 (42%)	58 (28%)	90 (35%)
Body mass index (kg/m^2)	24 (4)	25 (6)	24 (5)	25 (5)
Duration of back pain (months)	14 (7)	12 (6)	13 (7)	14 (7)
Imaging features of SpA				
Sacroiliitis on MRI-SIJ (ASAS)	125 (96%)	14 (13%)	0 (0%)	1 (0%)
Definitive damage on X-SIJ (mNY)	15 (12%)	1 (1%)	3 (1%)	5 (2%)
BME on MRI-Spine (≥5 lesions)	12 (9%)	1 (1%)	0 (0%)	2 (1%)
≥1 syndesmophyte on X-Spine	9 (7%)	21 (19%)	14 (7%)	20 (8%)
Clinical features of SpA				
Peripheral arthritis ever	13 (10%)	69 (63%)	2 (1%)	20 (8%)
Dactylitis ever	5 (4%)	33 (30%)	2 (1%)	3 (1%)
Heel pain ever	15 (12%)	82 (75%)	28 (14%)	22 (9%)
Psoriasis ever	13 (10%)	48 (44%)	15 (7%)	9 (3%)
Uveitis ever	14 (11%)	10 (9%)	20 (10%)	13 (5%)
Inflammatory bowel disease ever	3 (2%)	8 (7%)	12 (6%)	24 (9%)
Inflammatory back pain ever	97 (75%)	104 (95%)	132 (65%)	157 (61%)
Elevated CRP (>5 mg/L)	55 (42%)	43 (39%)	43 (21%)	50 (19%)
HLA-B27 positive	102 (78%)	54 (49%)	100 (49%)	58 (22%)
Family history of SpA (ASAS)	56 (43%)	52 (47%)	204 (100%)	0 (0%)
Good response to NSAID ever	59 (45%)	57 (52%)	65 (32%)	69 (27%)

Participants were individually categorised to one of the classes (obtained from the latent transition analysis) based on their posterior probability of class membership (with the class having the highest probability for each patient determining their assignment).
Mean (SD) or n (%).
Data for BMI available for 673 (96%) participants, for duration of back pain for 696 (99%) participants.
Missing values for imaging and clinical features of SpA have been imputed as described in the 'Methods' section.
ASAS, Assessment of SpondyloArthritis international Society; BME, bone marrow oedema; BMI, body mass index; CRP, C reactive protein; HLA-B27, human leucocyte antigen B27; IBP, inflammatory back pain; mNY, modified New York criteria; NSAID, non-steroidal anti-inflammatory drug; SIJ, sacroiliac joints; SpA, spondyloarthritis; X-SIJ, radiograph of the sacroiliac joints; X-Spine, radiograph of the spine.

DISCUSSION

The four classes labelled 'pure axSpA' ('axial'), 'axSpA with peripheral signs' ('IBP+peripheral'), 'axSpA at risk' and 'no spondyloarthritis' ('no SpA'), previously described in the SPACE cohort and found through a data-driven approach were followed-up for 2 years. While the first two classes reflect the clinical construct or 'Gestalt' of patients with axial and peripheral manifestations of axSpA, the latter two capture patients with chronic back pain, either without features or only with risk factors for axSpA. The current analysis revealed no relevant class switch over time, more specifically only 3% from 'no axSpA' to 'at risk', underlining the unlikeliness of developing class-defining new features of axSpA over 2 years after an initial clinical workup. The remaining 3% of transitions mainly reflect a positive family history only noted postbaseline, with age and sex showing no significant effect on this transition.

These findings are in line with a similar analysis in the DESIR cohort; the same classes with the exception of the 'no SpA' class were found. The latter is aligned with the different populations included in the two cohorts: In DESIR, only participants with axSpA are included, whereas in SPACE, participants with chronic back pain, with or without an axSpA diagnosis, are included.[9 19] In DESIR, the LTA revealed only an 11% switch of participants from the 'axial SpA at risk' to the 'axial SpA with peripheral signs' class over a time period of 5 years.[7] The development of new peripheral and extramusculoskeletal manifestations in axSpA is well known.[20 21] In SPACE, the follow-up was shorter (2 years) and it is possible that a longer follow-up period would have shown a similar transition, similar to DESIR, also in SPACE. In any case, a similar conclusion is possible in both analyses: there is very little change in the classes over follow-up.

The initial aim for this analytical approach of separating participants in different classes, was to gain an alternative, mostly expert-judgement-free insight into the 'Gestalt' of axSpA. This distinguishes it from diagnosis, where experts assess SpA features based on preconceived

Table 4 Sensitivity analysis: classes of participants with chronic back pain suspicious of axial spondyloarthritis identified in the 2-year latent transition analysis model in the population with complete follow-up (n=384)

	Class 1 'axial' (p*=0.25, N†=96)	Class 2 'IBP+peripheral' (p*=0.15, N†=58)	Class 3 'at risk' (p*=0.31, N†=119)	Class 4 'no SpA' (p*=0.29, N†=111)
Sacroiliitis on MRI-SIJ (ASAS)	1.00	0.26	0.07	0.05
BME on MRI-Spine (≥5 lesions)	0.10	0.02	0.00	0.01
Definitive damage on X-SIJ (mNY)	0.10	0.03	0.02	0.01
≥1 syndesmophyte on X-Spine	0.09	0.23	0.10	0.12
Elevated CRP (>5 mg/L)	0.50	0.44	0.29	0.24
Good response to NSAID ever	0.60	0.73	0.54	0.51
Peripheral arthritis ever	0.14	0.78	0.00	0.13
Dactylitis ever	0.08	0.45	0.03	0.01
Heel pain ever	0.18	0.80	0.14	0.23
Family history of SpA	0.45	0.59	1.00	0.00
Psoriasis ever	0.09	0.50	0.07	0.08
Uveitis ever	0.15	0.08	0.10	0.10
Inflammatory bowel disease ever	0.04	0.08	0.07	0.10
Inflammatory back pain ever	0.80	0.98	0.78	0.86

Conditional probabilities (ie, the probability of a feature being present in one of the classes, range: 0–1) were obtained using a latent transition analysis model with full invariance on baseline and 2-year data. Full invariance means that these probabilities are the same at baseline and follow-up.
Cells are coloured in green whenever the conditional probability is ≥0.3. This cut-off was chosen to better visualise differences between the classes.
*Marginal probability of the latent class (ie, a participant's probability of class membership).
†Participants categorised to one of the classes based on their posterior probability of class membership (with the class having the highest probability for each patient determining their assignment).
ASAS, Assessment of SpondyloArthritis international Society; BME, bone marrow oedema; CRP, C reactive protein; IBP, inflammatory back pain; mNY, modified New York criteria; NSAID, non-steroidal anti-inflammatory drug; SIJ, sacroiliac joints; SpA, spondyloarthritis; X-SIJ, radiograph of the sacroiliac joints; X-Spine, radiograph of the spine.

ideas and experiences and is also different to classification criteria, which involve a similar weighting of features followed by validation against expert opinion.[22–24]

However, all three methods of stratifying patients have their own unique significance or 'truth': (1) the clinical diagnosis represents the rheumatologist's 'truth' which while being influenced by perceptions and prone to circular reasoning is key for treatment and related decisions; (2) classification criteria represent the best possible external 'truth' of the internationally accepted construct of axSpA, which aims to generate homogenous groups of participants to be included in research studies, mostly clinical trials and (3) lastly, LCA and LTA reflect the statistical 'truth' derived from a model that needs human input solely for the selection of patients, variables and detection of meaningful classes, but not for the weighting of the mentioned variables. This means, for instance, that the importance of sacroiliitis is not predefined to be superior to that of heel pain by the researcher, allowing the model itself to determine the significance of each feature, thereby introducing a more objective perspective on the interplay between the SpA features. As a clear hierarchy between these 'truths' does not exist we have on purpose refrained from performing comparisons between them in this study.

Nevertheless, the stability of the classes over time also aligns with the stability of an axSpA diagnosis over time, as described in a recent analysis in SPACE. This analysis revealed that participants diagnosed with axSpA during an initial assessment were reliably considered to have the disease also after 2 years, whereas participants with definitely no initial axSpA diagnosis were unlikely to receive one after this follow-up period. Furthermore, the uncertainty of a diagnosis persisted in up to 30% of the participants.[8] The alignment of these results with the data-driven approach reported here confirms and reassures that our perception of patients with back pain suspicious of axSpA is not biased by placing more value on certain SpA features. It validates the 'Gestalt' of axSpA that we encounter in everyday practice and confirms the unlikelihood of developing new class-defining features over time.

LCA was chosen over traditional clustering methods (eg, k-means, k-medians, hierarchical clustering) as it has been shown to outperform them in correctly identifying classes/clusters in simulation studies, especially when the number of classes/clusters was unknown.[17 25] LCA's 'model-based' approach also provides a more objective way of model comparison through statistical testing.[26] Lastly, LTA, used for the main analysis of obtaining

transitional probabilities for the class switch, builds on LCA and adds the handling of longitudinal data. The main limitation of this study is that several participants with a baseline visit were, per protocol, excluded from follow-up visits in SPACE (100/702). This was due to the very low probability of these individuals getting diagnosed with axSpA, as determined by their clinical, laboratory and imaging data. To minimise selection bias, we still included these participants in the main analysis and imputed their baseline data at 2 years. Similarly, we imputed missing baseline values for clinical or imaging features using data from additional follow-up visits, or if not available, we considered the corresponding SpA features as absent. The latter was, however, done only in a small number of variables (5% of all imaging and 0.2% of all clinical features). The sensitivity analysis, using only participants with complete data (384/702), resulted in very similar results, which adds to the robustness of the findings. The variable HLA-B27 status was removed from the final LTA model due to its time-invariant nature. Furthermore, the value of including both HLA-B27 status and family history as independent variables is questionable, considering their high collinearity.[27] Performing LCA we found meaningful classes with and without HLA-B27 status and a meaningful final LTA model without the genetic marker. Apart from this, the model yielded stable results in line with the previous LCA conducted in SPACE on a subset of participants for whom data was available at that moment.[7] While response to NSAID treatment was included as a latent variable in the LTA model, treatment with a disease-modifying antirheumatic drug (DMARD) was not. The main reason for this was that DMARD treatment would imply an axSpA diagnosis and our aim was to identify and follow-up classes without prior expert knowledge of their diagnosis. A sample size calculation was not performed as there is currently no consensus on a standardised method for LTA. However, our dataset exceeded the minimal sample size of 300 participants that has been recommended and all of the final models converged.[28] Lastly, the 2 years time frame was relatively short for observing changes between the classes over time. Extending the follow-up duration was, however, not possible due to the current availability of scored imaging data only up to the 2 years mark. A future analysis with longer follow-up may reveal more change between the classes, as in DESIR, where only one out of ten participants switched from the 'axSpA at-risk' class to the 'axSpA with peripheral signs' class.

In summary, four distinct latent classes of participants in the SPACE cohort, previously found using a mostly circularity-free data-driven process, were followed up for 2 years. Class switch over time was basically inexistent, highlighting the unlikelihood of participants to develop relevant new features of axSpA over this time period after an initial clinical workup.

Author affiliations
[1]Department of Rheumatology and Immunology, Medical University of Graz, Graz, Austria
[2]Department of Rheumatology, Leiden University Medical Center (LUMC), Leiden, Netherlands
[3]Department of Epidemiology, Maastricht University, Maastricht, Netherlands
[4]Nova Medical School, Universidade Nova de Lisbona, Lisbon, Portugal
[5]Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal
[6]Amsterdam Rheumatology Center, Amsterdam University Medical Center, Amsterdam, Netherlands
[7]Department of Rheumatology, Zuyderland Medical Center, Heerlen, Netherlands
[8]Department of Rheumatology, Ghent University Hospital, Ghent, Belgium
[9]Department of Clinical Sciences Malmö (Rheumatology), Lund University, Malmö, Sweden
[10]Rheumatology Unit, Department of Medicine - DIMED, University Hospital of Padova, Padova, Italy
[11]Center for Treatment of Rheumatic and Musculoskeletal Diseases (REMEDY), Diakonhjemmet Hospital, Oslo, Norway

X Philipp Bosch @philipp_j_bosch, Alexandre Sepriano @AlexSepriano and Manouk de Hooge @MdH42715285

Acknowledgements The authors would like to acknowledge the Assessment of Spondyloarthritis international Society (ASAS) for having supported Philipp Bosch by granting him the ASAS fellowship.

Contributors AS, DvdH, RL, CB, FAvG and SR conceptualised the manuscript's content. DvdH, RL, MvL, LdB, MdH, SE, RR, KMF, FAvG and SR collected patient data. MLM, MvL and LdB performed data management tasks. PB performed the analysis. PB wrote the first draft of the manuscript. All authors made extensive comments and appropriate changes to the manuscript. PB and SR are the guarantors.

Funding PB has received an Assessment of SpondyloArthritis International Society (ASAS) fellowship.

Competing interests PB has received travelling grants and adds board fees from Janssen, speaker fees from Janssen and AbbVie and projects grants from Pfizer; AS has received speaking and/or consulting fees from AbbVie, Novartis, UCB and Lilly. DvdH has received consulting fees from AbbVie, Argenx, BMS, Galapagos, Glaxo-Smith-Kline, Janssen, Lilly, Novartis, Pfizer, Takeda, UCB Pharma and is an associate editor of Annals Rheumatic Diseases, editorial board member of Journal of Rheumatology and RMD Open, Advisor Assessment of Axial Spondyloarthritis international Society and director of Imaging Rheumatology bv. RL has received honoraria for lectures and advisory boards from AbbVie, Galapagos, Janssen, Lilly, Novartis, Pfizer and UCB. RL is owner of Rheumatology Consultancy BV. MdH has received consultancy fees of UCB Pharma. SE has received consultancy fees from AbbVie, Amgen, Janssen, Novartis, UCB Pharma, Eli Lilly and research support from UCB Pharma. RR has received travelling grants, consulting and/or speaking fees, Advisory Boards fees from Abbvie, Novartis, Janssen, Lilly, MSD, Pfizer, and UCB. FvG reports research grants from Stichting vrienden van Sole Mio, Stichting ASAS, research grants and consultancy fees from Novartis, research grants from UCB, fees from MSD, consultancy fees from Abb Vie, fees from Bristol Myers Squibb and Eli Lilly, is a full time employee of the LUMC and member of the ASAS EC and the ASAS treasurer. SR has received grants from AbbVie, Galapagos, MSD, Novartis, Pfizer, UCB and consultancy fees from AbbVie, Eli Lilly, Galapagos, Janssen, MSD, Pfizer, UCB, Sanofi. The other authors have not declared any COI.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by Medical Ethics Committee Leiden The Hague Delft approval (P08.105). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as online supplemental information.

ORCID iDs
Philipp Bosch http://orcid.org/0000-0002-6783-6422
Alexandre Sepriano http://orcid.org/0000-0003-1954-0229
Mary-Lucy Marques http://orcid.org/0000-0002-3071-9425
Désirée van der Heijde http://orcid.org/0000-0002-5781-158X
Robert Landewé http://orcid.org/0000-0002-0577-6620
Miranda van Lunteren http://orcid.org/0000-0002-0601-4863
Manouk de Hooge http://orcid.org/0000-0002-0652-9808
Roberta Ramonda http://orcid.org/0000-0002-9683-8873
Karen Minde Fagerli http://orcid.org/0000-0002-0327-7860
Floris A van Gaalen http://orcid.org/0000-0001-8448-7407
Sofia Ramiro http://orcid.org/0000-0002-8899-9087

REFERENCES

1 Sieper J, Poddubnyy D. Axial spondyloarthritis. *Lancet* 2017;390:73–84.
2 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.
3 Rudwaleit M, van der Heijde D, Landewe R, *et al*. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann Rheum Dis* 2009;68:777–83.
4 Rudwaleit M, van der Heijde D, Landewé R, *et al*. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. *Ann Rheum Dis* 2011;70:25–31.
5 Lukas C, Cyteval C, Dougados M, *et al*. MRI for diagnosis of axial spondyloarthritis: major advance with critical limitations "Not everything that glisters is gold (standard)." *RMD Open* 2018;4:e000586.
6 Sepriano A, Landewé R, van der Heijde D, *et al*. Predictive validity of the ASAS classification criteria for axial and peripheral spondyloarthritis after follow-up in the ASAS cohort: a final analysis. *Ann Rheum Dis* 2016;75:1034–42.
7 Sepriano A, Ramiro S, van der Heijde D, *et al*. What is axial spondyloarthritis? A latent class and transition analysis in the SPACE and DESIR cohorts. *Ann Rheum Dis* 2020;79:324–31.
8 Marques ML, Ramiro S, van Lunteren M, *et al*. Can rheumatologists unequivocally diagnose axial spondyloarthritis in patients with chronic back pain of less than 2 years duration? Primary outcome of the 2-year SPondyloArthritis Caught Early (SPACE) cohort. *Ann Rheum Dis* 2024;83:589–98.
9 van den Berg R, de Hooge M, van Gaalen F, *et al*. Percentage of patients with spondyloarthritis in patients referred because of chronic back pain and performance of classification criteria: experience from the Spondyloarthritis Caught Early (SPACE) cohort. *Rheumatology (Oxford)* 2013;52:1492–9.
10 Rudwaleit M, Heijde D, Khan MA, *et al*. How to diagnose axial spondyloarthritis early. *Ann Rheum Dis* 2004;63:535–43.
11 Creemers MCW, Franssen MJAM, van't Hof MA, *et al*. Assessment of outcome in ankylosing spondylitis: an extended radiographic scoring system. *Ann Rheum Dis* 2005;64:127–9.
12 Lambert RGW, Bakker PAC, van der Heijde D, *et al*. Defining active sacroiliitis on MRI for classification of axial spondyloarthritis: update by the ASAS MRI working group. *Ann Rheum Dis* 2016;75:1958–63.
13 Rudwaleit M, Jurik AG, Hermann K-GA, *et al*. Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. *Ann Rheum Dis* 2009;68:1520–7.
14 de Hooge M, van den Berg R, Navarro-Compán V, *et al*. Patients with chronic back pain of short duration from the SPACE cohort: which MRI structural lesions in the sacroiliac joints and inflammatory and structural lesions in the spine are most specific for axial spondyloarthritis? *Ann Rheum Dis* 2016;75:1308–14.
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;27:361–8.
16 Sieper J, van der Heijde D, Landewé R, *et al*. New criteria for inflammatory back pain in patients with chronic back pain: a real patient exercise by experts from the Assessment of SpondyloArthritis international Society (ASAS). *Ann Rheum Dis* 2009;68:784–8.
17 Brusco MJ, Shireman E, Steinley D. A comparison of latent class, K-means, and K-median methods for clustering dichotomous data. *Psychol Methods* 2017;22:563–80.
18 Hallquist MN, Wiley JF. MplusAutomation: An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. *Struct Equ Model* 2018;25:621–38.
19 Dougados M, d'Agostino M-A, Benessiano J, *et al*. The DESIR cohort: A 10-year follow-up of early inflammatory back pain in France: Study design and baseline characteristics of the 708 recruited patients. *Joint Bone Spine* 2011;78:598–603.
20 Essers I, Ramiro S, Stolwijk C, *et al*. Characteristics associated with the presence and development of extra-articular manifestations in ankylosing spondylitis: 12-year results from OASIS. *Rheumatology (Oxford)* 2015;54:633–40.
21 López-Medina C, Dougados M, Ruyssen-Witrand A, *et al*. Evaluation of concomitant peripheral arthritis in patients with recent onset axial spondyloarthritis: 5-year results from the DESIR cohort. *Arthritis Res Ther* 2019;21:139.
22 Aggarwal R, Ringold S, Khanna D, *et al*. Distinctions between diagnostic and classification criteria? *Arthritis Care Res (Hoboken)* 2015;67:891–7.
23 Poddubnyy D. Classification vs diagnostic criteria: the challenge of diagnosing axial spondyloarthritis. *Rheumatology (Oxford)* 2020;59:iv6–17.
24 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.
25 Magidson J, Vermunt JK. Latent class models for clustering: A comparison with K-means. *Can J of Mark Res* 2002.
26 Schreiber JB. Latent Class Analysis: An example for reporting results. *Res Soc Adm Pharm* 2017;13:1196–201.
27 van Lunteren M, van der Heijde D, Sepriano A, *et al*. Is a positive family history of spondyloarthritis relevant for diagnosing axial spondyloarthritis once HLA-B27 status is known? *Rheumatology (Oxford)* 2019;58:1649–54.
28 Nylund-Gibson K, Garber AC, Carter DB, *et al*. Ten frequently asked questions about latent transition analysis. *Psychol Methods* 2023;28:284–300.