RMD Open

Rheumatic & Musculoskeletal Diseases

ORIGINAL RESEARCH

Identification of PsA phenotypes with machine learning analytics using data from two phase III clinical trials of guselkumab in a bio-naïve population of patients with PsA

Pascal Richette [1,2] Marijn Vis,[3] Sarah Ohrndorf [4] William Tillett [5,6]
Julio Ramírez,[7] Marlies Neuhold,[8] Michel van Speybroeck,[9] Elke Theander,[10]
Wim Noel,[11] Miriam Zimmermann [8] May Shawi,[12] Alexa Kollmeier,[13]
Alen Zabotti [14]

ABSTRACT

Objectives Psoriatic arthritis (PsA) phenotypes are typically defined by their clinical components, which may not reflect patients' overlapping symptoms. This post hoc analysis aimed to identify hypothesis-free PsA phenotype clusters using machine learning to analyse data from the phase III DISCOVER-1/DISCOVER-2 clinical trials.

Methods Pooled data from bio-naïve patients with active PsA receiving guselkumab 100 mg every 8/4 weeks were retrospectively analysed. Non-negative matrix factorisation was applied as an unsupervised machine learning technique to identify PsA phenotype clusters; baseline patient characteristics and clinical observations were input features. Minimal disease activity (MDA), disease activity index for psoriatic arthritis (DAPSA) low disease activity (LDA) and DAPSA remission at weeks 24 and 52 were evaluated.

Results Eight clusters (n=661) were identified: cluster 1 (feet dominant), cluster 2 (male, overweight, psoriasis dominant), cluster 3 (hand dominant), cluster 4 (dactylitis dominant), cluster 5 (enthesitis, large joints), cluster 6 (enthesitis, small joints), cluster 7 (axial dominant) and cluster 8 (female, obese, large joints). At week 24, MDA response was highest in cluster 2 and lowest in clusters 3, 5 and 6; at week 52, it was highest in cluster 2 and lowest in cluster 5. At weeks 24 and 52, DAPSA LDA and remission were highest in cluster 2 and lowest in clusters 4 and 6, respectively. All clusters improved with guselkumab treatment over 52 weeks.

Conclusions Unsupervised machine learning identified eight PsA phenotype clusters with significant differences in demographics, clinical features and treatment responses. In the future, such data could help support individualised treatment decisions.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Patterns of joint involvement in psoriatic arthritis (PsA) are difficult to define, largely due to the complex and heterogeneous nature of the disease.
⇒ Treatment selection is currently based on the individual domains or clinical components of the disease, as well as the related conditions and comorbidities, but this is of limited use in individualised medicine.

WHAT THIS STUDY ADDS

⇒ This study demonstrates that it is possible to use machine learning analytics to identify PsA phenotype clusters that differ in their demographic and clinical characteristics.
⇒ Using data from a phase III clinical trial programme evaluating guselkumab in a bio-naïve population of patients with PsA, eight phenotype clusters were identified.
⇒ These could be clearly distinguished by patterns of involvement of joints, skin/nails, dactylitis, enthesitis and axial manifestations.
⇒ Response to therapy also differed across clusters.
⇒ These clusters also demonstrated some alignment with patients enrolled in real-world evidence studies.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our findings support the use of unsupervised machine learning for the identification of PsA phenotype clusters, which may help inform treatment choices.
⇒ In this analysis, guselkumab, an interleukin-23 inhibitor, demonstrated efficacy across all clusters.

INTRODUCTION

Psoriatic arthritis (PsA) is a heterogeneous, chronic, progressive, immune-mediated inflammatory joint disease[1 2] with an average global prevalence of approximately 0.13%, although this could be as high as 0.42%.[3 4] Currently, PsA is considered based on the domains or clinical components of the

BMJ

disease, which is useful for classification and diagnosis, but may not reflect the whole spectrum and overlap of disease phenotypes.[5] At disease onset, oligoarthritis is considered the most frequent joint pattern, although several other patterns have been identified.[1] Musculoskeletal involvement includes arthritis, enthesitis, dactylitis and axial involvement.[1] Extra-articular manifestations are common, particularly skin (psoriasis) and nail (eg, pitting, onycholysis) changes.[1] Furthermore, comorbidities (eg, obesity, cardiovascular diseases, depression, anxiety) are frequent in patients with PsA.[1 6–8]

Several PsA therapies are available, including nonsteroidal anti-inflammatory drugs, conventional disease-modifying anti-rheumatic drugs (DMARDs), biologic DMARDs and targeted synthetic DMARDs.[1 9] However, PsA management is frequently associated with challenges, including lack of response to therapy or loss of response over time.[2 10] The large spectrum of PsA phenotypes contributes to these challenges.[10] Therefore, optimal clinical management of PsA requires a personalised approach to treatment, based on the specific clinical manifestations in each individual patient, as well as related conditions and comorbidities.[5]

Guselkumab has high specificity and affinity for the p19 subunit of interleukin (IL)-23, the overexpression of which is a key driver of clinical features in PsA.[2] The efficacy of guselkumab in improving symptoms of joint and skin disease and preventing structural damage progression and its safety versus placebo have been demonstrated at week 24 of the phase III COSMOS (NCT03796858),[11] DISCOVER-1 (NCT03162796)[12] and DISCOVER-2 (NCT03158285)[13] clinical trials. Data are available through 1 year for COSMOS and DISCOVER-1 and through 2 years for DISCOVER-2.[2 11 12]

We performed a post hoc analysis using unsupervised machine learning analytics[14] to identify clusters of PsA phenotypes, which were defined according to patients' clinical features and characteristics at baseline, using data from bio-naïve patients with PsA enrolled in DISCOVER-1 and DISCOVER-2. For each phenotype cluster, the clinical response to guselkumab at week 24 and week 52 was assessed.

METHODS
Data sources
The design and methods of DISCOVER-1 and DISCOVER-2 are described in detail elsewhere.[12 13] Briefly, DISCOVER-1 and DISCOVER-2 were randomised controlled trials in adults with active PsA despite standard treatment. Patients from DISCOVER-1 who had prior biological treatment with one or two tumour necrosis factor (TNF) inhibitors (approximately 30% of study participants)[12] were excluded from this analysis. The combined inclusion criteria were: ≥3 swollen and ≥3 tender joints, and C reactive protein (CRP) ≥0.3 mg/dL in DISCOVER-1; and ≥5 swollen and ≥5 tender joints, and CRP ≥0.6 mg/dL in DISCOVER-2.[12 13] Presence of spondylitis was assessed at enrolment, and spondylitis with peripheral disease was the primary presentation in ~35% of patients. This study retrospectively analysed pooled data from bio-naïve patients with active PsA treated with guselkumab 100 mg every 8 or 4 weeks.

Machine learning analytics
The main objective of this analysis was to use unsupervised machine learning to identify PsA phenotype clusters that could be differentiated in terms of demographic and clinical features. Non-negative matrix factorisation (NMF) was applied to identify PsA phenotype clusters, with baseline characteristics and baseline clinical observations as input features. NMF is a frequently used dimensionality reduction technique,[15] whereby a matrix (M) comprising m rows and n columns decomposes into two non-negative matrices (U and N) of the original n columns by u clusters and those same u clusters by the m original rows (figure 1).[15 16] Mathematically, the following distance is minimised: $\|M - UN\|_F^2 = \sum_{i,j} (M - UN)_{ij}^2$, whereby the F subscript denotes the Frobenius norm.

NMF was performed in Python, using the Frobenius distance as the beta divergence cost function. All available baseline variables were converted to a set of 115 non-negative features in the range 0–1. Variables used to identify PsA phenotypes were: involved joint location, including the hand, wrist, elbow, shoulder, temporomandibular joint, hip, knee, ankle and foot; number
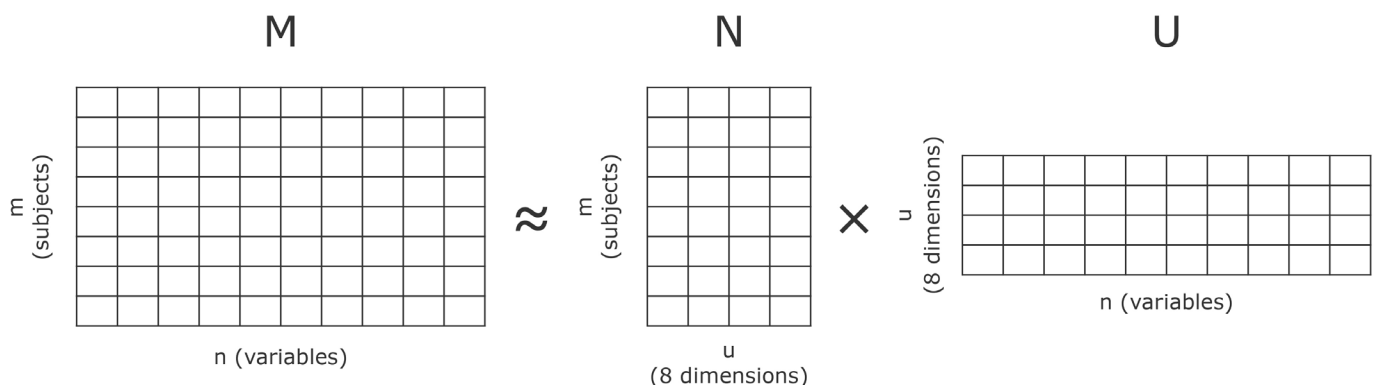

Figure 1 Non-negative matrix factorisation.

of involved joints assessed using 68/66 tender joint count (TJC)/swollen joint count (SJC); skin involvement assessed using Psoriasis Area and Severity Index (PASI) and psoriasis location (nail, scalp and hand and/or foot); dactylitis assessed using digit count and total dactylitis score; and enthesitis of Achilles tendon insertion (left/right), lateral epicondyle humerus (left/right) and medial femoral condyle (left/right). Baseline characteristics were selected for feasibility of clustering and included: age group (<45, 45–65, >65 years), sex, body mass index (BMI) (<25, 25 to <30, ≥30 kg/m^2), CRP (continuous), PsA duration (continuous), PsA subtype (arthritis mutilans, asymmetric peripheral arthritis, distal interphalangeal (DIP) joint involvement, polyarticular arthritis, spondylitis with peripheral arthritis), body surface area (BSA) (<3%, ≥3% to <10%, ≥10%), PASI and smoking status (past and current history). These variables were agreed prior to analysis.

Different numbers of PsA phenotype clusters were tested (four to eight). The optimal number of clusters, eight, was determined by assessing the clinical relevance and discriminatory ability for different cluster numbers; U and N matrices were calculated for these eight clusters. As shown in figure 1, the N matrix is an m×8 matrix and gives the relative weight for each subject to the eight dimensions, while the U matrix gives the weight of each variable to the eight dimensions. Following the NMF clustering method, patients were assigned automatically to the dominant cluster (ie, the column in matrix N where, for a given patient, the maximum value was calculated). Only baseline characteristics and disease features, not treatment response, were included in the clustering. Cluster homogeneity and heterogeneity were expressed as average distances.

Identified clusters were described according to their characteristics and clinical features, as well as achievement of stringent composite disease activity scores—minimal disease activity (MDA), disease activity index for psoriatic arthritis (DAPSA) low disease activity (LDA) and DAPSA remission—and lack of radiographic progression at week 24 and week 52. Radiographic progression was assessed in DISCOVER-2 only, resulting in a smaller sample size.

Descriptive statistics were used to describe the clusters when evaluating MDA/DAPSA response.

Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

RESULTS
Phenotype clusters

The original dataset included 669 bio-naïve patients. After removing patients for whom one or more observed variables were missing, data were pooled from 661 patients, and 8 distinct PsA phenotype clusters were identified (figure 2). Baseline characteristics of the eight PsA phenotype clusters are presented in table 1 and for the overall patient population in online supplemental table 1.

Cluster 1

Cluster 1 was characterised by a high frequency of lower limb involvement (predominantly impacting the metatarsophalangeal joints, ankles and knees) including enthesitis, particularly in the Achilles tendon. The proportion of patients with severe psoriasis was low (only 36.7% had BSA >10%).

Cluster 2

Cluster 2 was characterised by high psoriasis skin involvement, the highest proportion of overweight (not obese) patients (70.4% with BMI 25 to <30 kg/m^2), the highest proportion of male patients (80.0%), high scalp (88.8%) and nail (74.4%) psoriasis involvement and an asymmetric phenotype. The transverse tarsal joint was involved in 70% of patients.
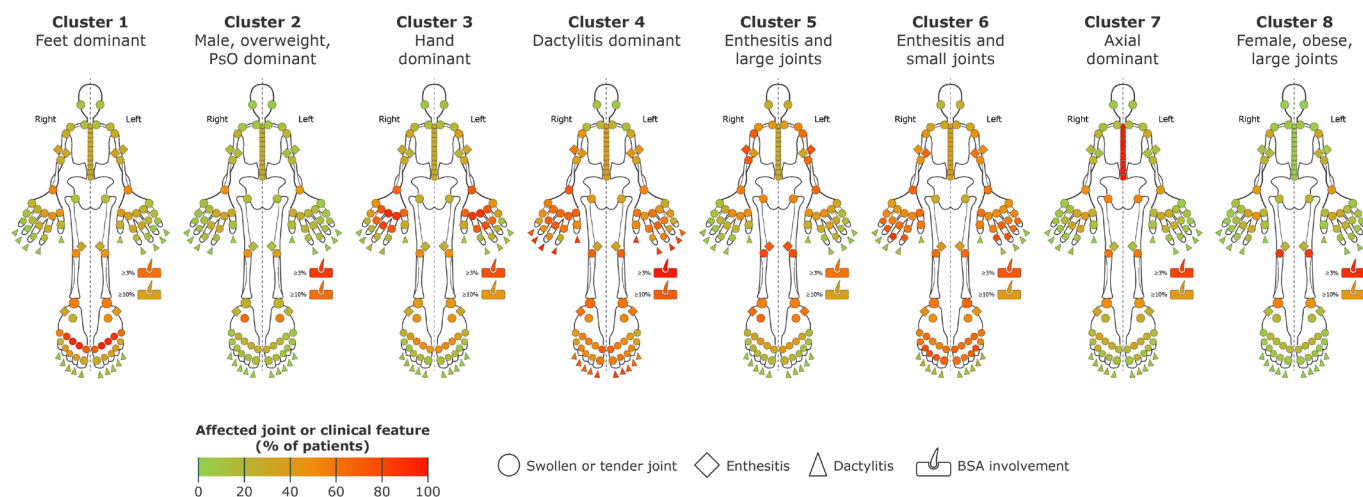

Figure 2 Identified PsA phenotype clusters based on frequency of joint involvement and clinical features. BSA, body surface area; PsA, psoriatic arthritis; PsO, psoriasis.

Table 1 Baseline characteristics of identified PsA phenotype clusters

	Cluster 1 Feet dominant	Cluster 2 Male, overweight, PsO dominant	Cluster 3 Hand dominant	Cluster 4 Dactylitis dominant	Cluster 5 Enthesitis and large joints	Cluster 6 Enthesitis and small joints	Cluster 7 Axial dominant	Cluster 8 Female, obese, large joints
Number of randomised and treated patients	79	125	95	38	57	60	83	124
Age, years	45.8 (10.4)	45.8 (13.2)	48.9 (11.8)	44.8 (12.3)	43.6 (13.3)	45.8 (11.6)	43.8 (10.4)	47.5 (11.1)
Female	54.4	20.0	60.0	26.3	64.9	41.7	36.1	61.3
BMI, kg/m^2	29.3 (5.4)	27.4 (3.6)	29.3 (6.1)	26.9 (4.6)	28.6 (7.5)	29.0 (6.2)	28.4 (6.9)	32.1 (6.5)
Normal <25	21.5	16.8	28.4	34.2	35.1	35.0	41.0	20.2
Overweight 25 to <30	35.4	70.4	31.6	39.5	29.8	20.0	19.3	12.9
Obese ≥30	43.0	12.8	40.0	26.3	35.1	45.0	39.8	66.9
CRP, mg/dL	1.7 (1.8)	1.7 (2.0)	1.4 (2.3)	2.5 (2.8)	1.7 (1.9)	1.7 (1.8)	2.2 (3.1)	1.5 (1.8)
Disease duration, years	5.0 (5.3)	5.1 (4.8)	5.8 (6.5)	6.1 (5.5)	6.8 (7.6)	6.1 (5.6)	4.7 (4.6)	5.2 (6.7)
HLA-B27 positive (n)*	19.0 (42)	29.3 (75)	30.5 (59)	20.0 (20)	33.3 (36)	25.0 (36)	34.8 (46)	26.3 (76)
SJC, 0–66	13.2 (7.1)	8.2 (3.8)	15.0 (8.0)	18.0 (10.1)	10.1 (5.1)	17.5 (11.8)	9.0 (4.2)	8.8 (3.9)
TJC, 0–68	23.4 (10.1)	12.8 (5.9)	26.0 (12.0)	30.6 (15.3)	23.2 (12.6)	37.5 (18.6)	14.6 (6.5)	12.7 (5.2)
BSA, %†	12.6 (19.4)	20.7 (19.8)	14.8 (19.5)	29.7 (26.4)	14.6 (21.6)	14.5 (18.4)	15.2 (19.4)	14.4 (15.7)
<3	36.7	13.6	23.4	0.0	38.6	27.1	18.1	9.7
≥3 to <10	26.6	23.2	33.0	28.9	24.6	30.5	40.9	46.0
≥10	36.7	63.2	43.6	71.1	36.8	42.4	41.0	44.4
PASI	7.8 (11.5)	12.1 (12.0)	7.7 (8.9)	16.5 (13.4)	8.9 (12.6)	11.4 (12.8)	9.1 (10.0)	8.5 (8.5)
Hand/foot PsO	45.6	44.8	35.8	57.9	35.1	48.3	51.8	57.3
Nail PsO	57.0	74.4	57.9	63.2	50.9	78.3	49.4	58.1
Scalp PsO	77.2	88.8	85.3	94.7	75.4	83.3	75.9	75.0
Dactylitis	48.1	30.4	42.1	100.0	36.8	53.3	49.4	31.5
Dactylitis score, 0–60‡	3.0 (4.8)	1.4 (3.2)	2.7 (4.8)	27.5 (12.3)	2.3 (5.2)	3.9 (6.7)	2.2 (2.9)	1.3 (2.6)
Enthesitis	70.9	49.6	58.9	81.6	96.5	73.3	69.9	45.2
LEI score, 0–6§	2.0 (1.7)	1.0 (1.3)	1.7 (1.8)	2.9 (1.9)	4.2 (1.6)	2.7 (2.3)	1.3 (1.2)	1.0 (1.3)
Current smoker	16.5	23.2	13.7	18.4	12.3	23.3	13.3	21.8
Current alcohol user	29.1	52.0	32.6	28.9	38.6	38.3	38.6	37.9

Data shown are mean (SD) or %.
*n denotes the number of patients for which HLA-B27 status was available in each cluster.
†Data were missing for two patients.
‡Dactylitis score is a measure of the degree of dactylitis using a scoring system (0, none; 1, mild; 2, moderate; 3, severe) at each of the 20 digits, giving an individual score range of 0–60.
§The LEI is a measure of the presence or absence of enthesitis at six sites (bilateral lateral epicondyles, medical femoral condyles and Achilles tendon insertions), giving an individual score range of 0–6.[18]
BMI, body mass index; BSA, body surface area; CRP, C reactive protein; LEI, Leeds Enthesitis Index; PASI, Psoriasis Area and Severity Index; PsA, psoriatic arthritis; PsO, psoriasis; SJC, swollen joint count; TJC, tender joint count.

Cluster 3

Cluster 3 was characterised by higher disease burden in hands and wrists, particularly the metacarpophalangeal and proximal interphalangeal joints, but very low DIP joint involvement compared with most other clusters. The total joint counts were high because of the high rate of small joint involvement. There were low rates of enthesitis and moderate rates of dactylitis and more female than male patients in this cluster.

Cluster 4

Cluster 4, the smallest cluster, scored highly in all categories, indicating severe disease and high inflammatory burden; it was distinguished by the highest degree of dactylitis involvement, the second highest proportion of enthesitis involvement, the highest skin involvement, high joint counts and the highest inflammatory burden, along with the second highest proportion of male patients.

Cluster 5

Cluster 5 had the highest proportion of patients with baseline enthesitis, particularly in the epicondyle humerus and femoral condyle, the highest proportion of female patients (64.9%), the lowest mean age (43.6 years) and the highest rate of large joint involvement. This cluster also had a high proportion of patients with no/mild psoriasis (BSA <3%), along with the longest disease duration of any cluster.

Cluster 6

Cluster 6 was characterised by a high level of small joint involvement in hands and feet, especially the DIP and interphalangeal joints, and a high rate of axial involvement and baseline enthesitis (mostly medial epicondyle humerus) but low mean dactylitis score; nail involvement was also present (78.3%), and this cluster had the highest TJC-68.

Cluster 7

Cluster 7 was characterised by axial involvement confirmed by imaging at baseline (100% of patients had imaging-confirmed sacroiliitis), dactylitis in nearly half of patients and enthesitis in more than two-thirds of patients. This cluster contained the highest proportion of HLA-B27-positive patients (34.8%). Additionally, most patients had BSA ≥3% at baseline, which included 41.5% of patients with BSA ≥3% to <10%, indicating moderate disease, and 40.2% with BSA ≥10%, indicating severe disease. Compared with most other clusters, a larger proportion of patients were male. There were also low rates of small joint involvement, resulting in a relatively low TJC-68/SJC-66.

Cluster 8

Cluster 8 was characterised by high rates of extensive skin involvement, the highest proportion of patients with obesity (66.9% with BMI >30 kg/m²), low rates of small joint involvement and higher rates of large joint involvement than other clusters (approximately 70% and 50%

with knee and ankle involvement, respectively), resulting in low TJC-68/SJC-66. Additionally, there was a high proportion of female patients (61.3%), and most patients had BSA ≥3% at baseline (BSA ≥3% to <10%, 46.0%; BSA ≥10%, 44.4%).

Treatment response

Minimal disease activity

Among all 661 bio-naïve patients treated with guselkumab, 24.7% and 34.5% of patients achieved an MDA response at week 24 and week 52, respectively, with increases between the two timepoints in all clusters. This composite measure includes assessment of multiple PsA domains, including skin and enthesitis, and patient-reported outcomes. At week 24, MDA response rates were highest in cluster 2 and lowest in clusters 3, 5 and 6. At week 52, MDA response rates were highest in cluster 2 and lowest in cluster 5; clusters 7 and 8 had consistently high response rates at week 24 and week 52 (figure 3). Clusters 3 and 4 were characterised by low MDA response rates at week 24, which increased at week 52 (figure 3).

Responses for individual MDA domains were also assessed (online supplemental figure 1). In general, MDA thresholds for SJC, PASI and Leeds Enthesitis Index were achieved more frequently than those for TJC, Health Assessment Questionnaire Disability Index (HAQ-DI), pain visual analogue scale (VAS) and patient global VAS. Increases in the proportion of patients achieving MDA thresholds from week 24 to week 52 differed by cluster: cluster 4 had the largest increases and cluster 7 had the smallest increases, across most domains. Larger increases in HAQ-DI, pain VAS and patient global VAS from week 24 to week 52 were seen in clusters 3 and 4 than in other clusters.

Disease activity index for psoriatic arthritis

Overall, DAPSA LDA response was achieved by 41.6% of patients at week 24 and 58.2% at week 52; 9.9% and 19.9% achieved DAPSA remission at week 24 and week 52, respectively. Increased DAPSA response rates were observed across all clusters from week 24 to week 52. DAPSA responses followed a similar pattern to MDA, although some of the factors assessed in this composite score differ, as it focusses on joint counts, CRP and patient-reported outcomes. DAPSA LDA response and DAPSA remission rates at week 24 and week 52 were highest in cluster 2 and lowest in clusters 4 and 6, respectively (figure 4). Clusters 3, 4 and 6 showed low DAPSA LDA response rates at week 24, which increased by 26.5%, 17.3% and 16.1%, respectively, at week 52. The clusters characterised by small joint involvement, clusters 1, 3, 4 and 6, showed low DAPSA remission rates at week 24; these increased by 14.0%, 8.7%, 11.3% and 8.9%, respectively, at week 52.

Radiographic progression

Overall radiographic progression in DISCOVER-2 was low, although clusters with more small joint involvement appeared to have greater progression
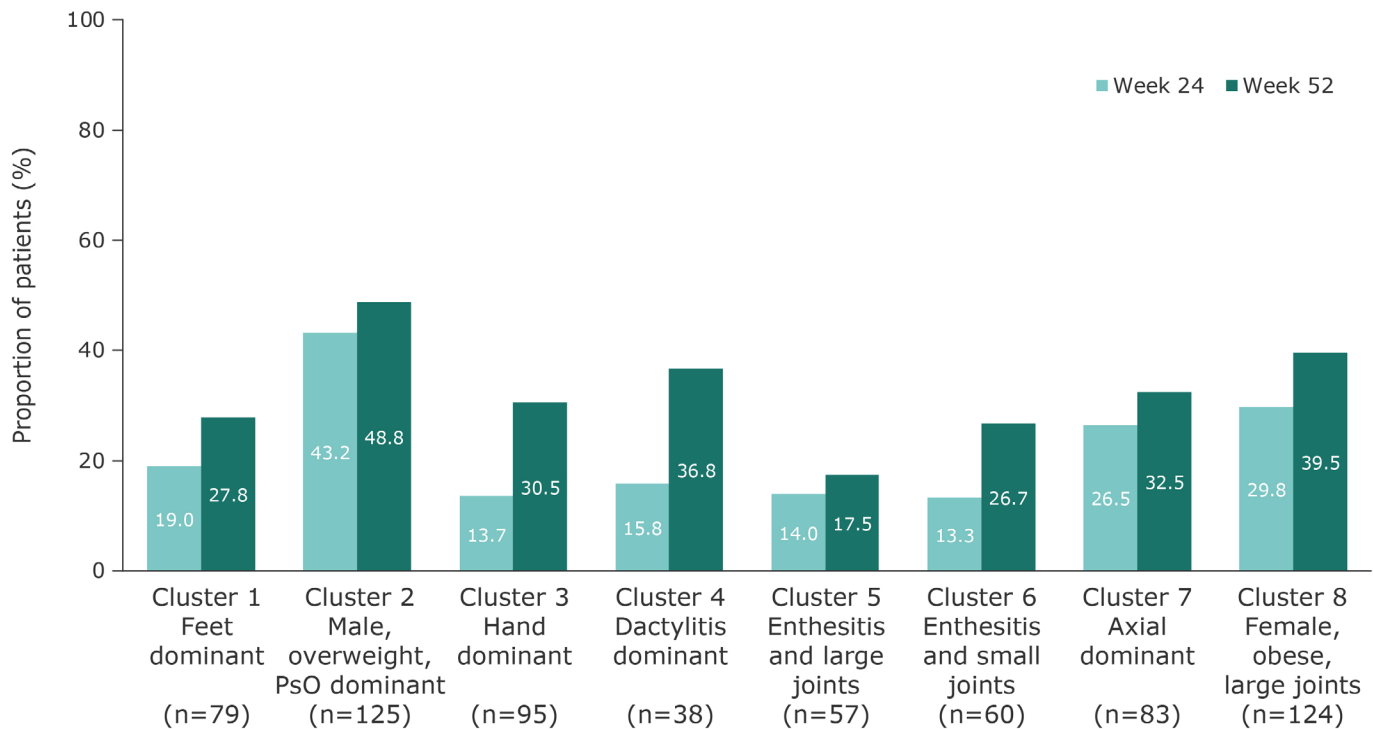
Figure 3 MDA response to guselkumab 100 mg (every 8 weeks and every 4 weeks pooled) at week 24 and increased at week 52 across PsA phenotype clusters. MDA, minimal disease activity; PsA, psoriatic arthritis; PsO, psoriasis.

than those with low small joint involvement (online supplemental figure 2).

DISCUSSION

Unsupervised machine learning[14] identified eight distinct PsA phenotype clusters showing significant differences in demographic and clinical features, including patterns of joint involvement, dactylitis, enthesitis, skin/nail manifestations and MDA and DAPSA response rates. MDA is a multidomain composite score that includes joint counts, patient pain and patient global assessment, and other domains such as skin and enthesitis,[17] thereby providing a holistic overview of patient improvement; whereas DAPSA is a composite score focussing on joint counts, patient pain, patient global assessment and CRP.[18] MDA and DAPSA are tools designed specifically for the assessment of PsA; despite its focus on joints, DAPSA correlates well with structural damage, physical function and patient perception.[17 19 20] Evaluating both scores together provides a comprehensive overview of patient responses, and the similarity of MDA and DAPSA response rates across clusters indicates the consistency of this analysis method.

Despite different MDA and DAPSA response rates, guselkumab was effective across all clusters, with continuous improvement through week 52. Clusters characterised by high skin involvement (clusters 2 and 8) showed strong responses by week 24, increasing through week 52. Importantly, these clusters had the highest proportions of overweight patients or patients with obesity, who typically have poor responses to TNF inhibitors.[21] This finding is

consistent with the results of the PsABIO study, a real-world cohort including many overweight (BMI >27 kg/m^2) patients with PsA, in which patients with similar clinical characteristics to the patients in this study demonstrated improved joint-related measures with IL-12/23 and TNF inhibitors, although responses to the latter were negatively influenced by higher BMI.[21] Cluster 5, characterised by enthesitis and large joints, had a lower rate of clinical response than clusters 2 and 8, and the baseline characteristics may indicate a low inflammatory burden, which is supported by the CRP level, rates of dactylitis, skin burden and SJC. This cluster also had the lowest rate of radiographic progression through week 24 and was among those with the lowest rates through week 52 (online supplemental figure 2).

Clusters appeared to differ in their initial versus late responses to guselkumab, and greater differences were seen for clusters with small joint involvement (clusters 3, 4 and 6), in which increasing therapy responses were seen from week 24 to week 52, particularly in achieving the MDA thresholds for SJC, PASI and patient-reported outcome domains. Furthermore, the increase in response rates over time for cluster 4, which had a high burden of inflammation and high rates of dactylitis, was consistent with temporal patterns of improvement in dactylitis reported with other biologics.[22–26] Taken together, these data may have relevance for clinical decision-making and require further analysis to clarify domains involved in response kinetics.

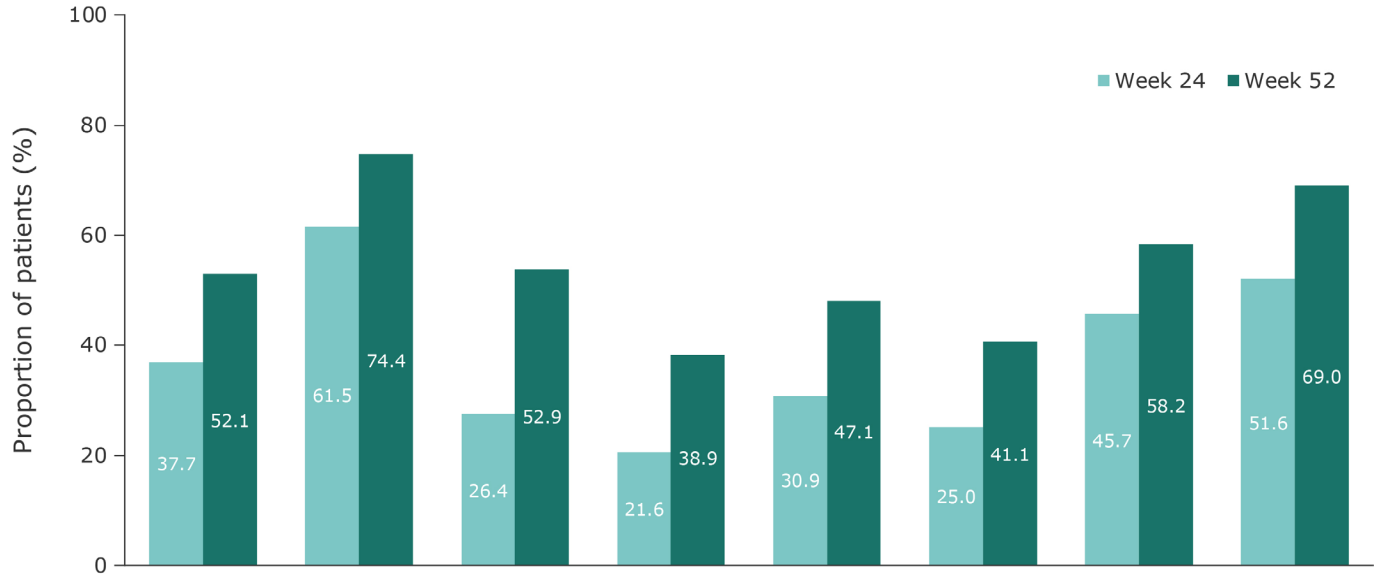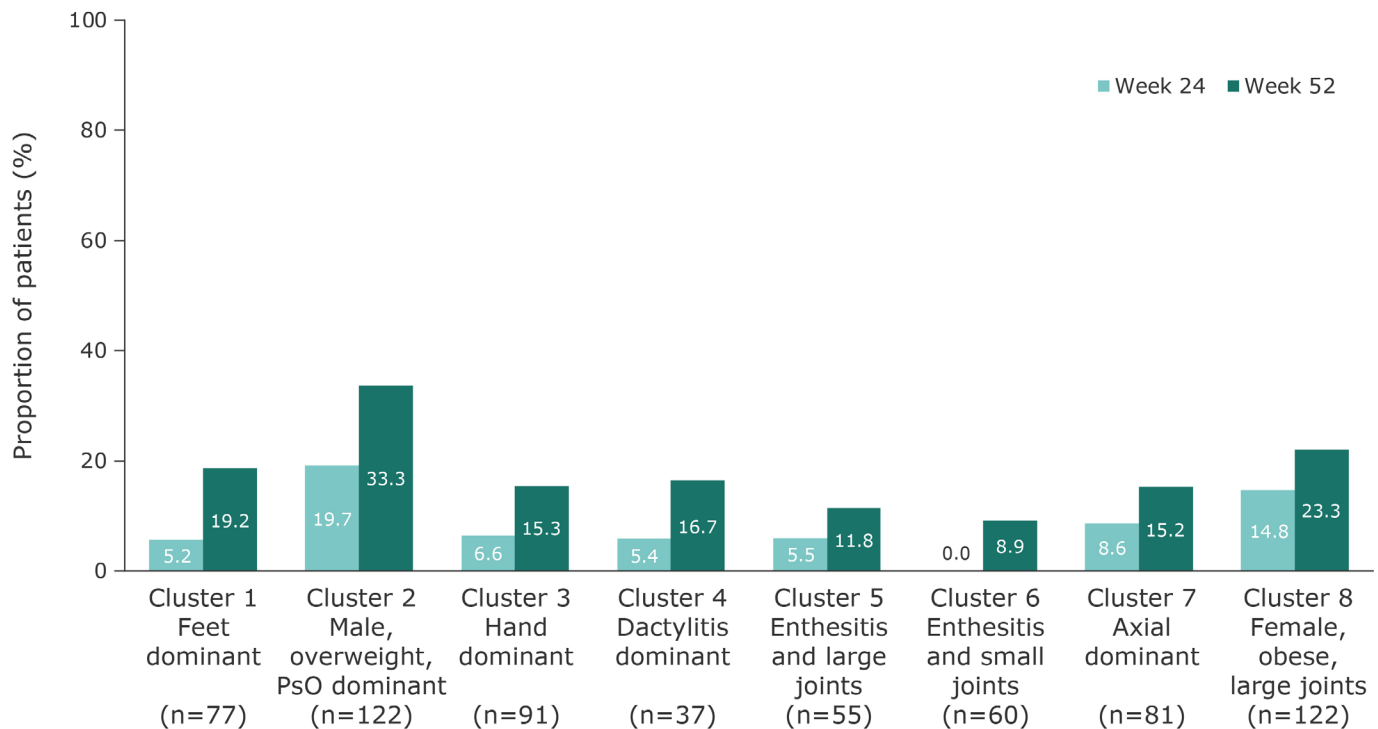The response pattern in cluster 7 (axial dominant), confirmed by baseline imaging (previous radiograph or

A DAPSA LDA

B DAPSA remission

Figure 4 DAPSA LDA response (A) and DAPSA remission (B) with guselkumab 100 mg (every 8 weeks and every 4 weeks pooled) at week 24 and increased at week 52 across PsA phenotype clusters.* *Patients with missing data at week 52: cluster 1, n=4; cluster 2, n=5; cluster 3, n=6; cluster 4, n=1; cluster 5, n=4; cluster 6, n=4; cluster 7, n=2; cluster 8, n=6. DAPSA, disease activity index for psoriatic arthritis; LDA, low disease activity; PsA, psoriatic arthritis; PsO, psoriasis.

MRI of sacroiliac joints, or pelvic radiograph at screening in DISCOVER-2), suggests that patients with axial involvement seem to respond strongly when using composite endpoints that include overall patient pain and patient global assessment. Axial disease management is typically challenging, and an increased understanding of the pathophysiology and differentiation from ankylosing spondylitis is still needed. Prospective data on axial disease management are rare, currently limited to the MAXIMISE study (a limitation of which is use of ankylosing spondylitis disease measures as

endpoints) and post hoc analyses, including those suggesting guselkumab may be effective at alleviating axial symptoms of PsA.[27 28] The efficacy of guselkumab in the treatment of axial symptoms and inflammation in active PsA axial disease is currently under investigation in the STAR study.[29]

The use of machine learning to classify phenotypes is an expanding area of research, and differences in clustering between studies are to be expected due to differences in methodology and patient population, such as inclusion criteria and treatment prior to clinical trial enrolment.

Importantly, characteristics of clusters from this analysis are consistent with real-world data, current literature and PsA pathophysiology. For example, the largest clusters, clusters 2 and 8, appear to reflect patient characteristics seen in clinical practice, as observed through similarities in joint counts, BMI and skin involvement, and similarities with patient populations in the PsABIO study.[21] There could be a potential link between cluster 3, which showed index finger predominance, and the research conducted by Tillett *et al* on frequent erosion in dominant joints[30] and by Helliwell *et al* in feet.[31] Finally, the characteristics of cluster 6 (enthesitis, small joint involvement and nail psoriasis involvement in almost 80% of patients) can be linked back to PsA pathophysiology, where it has been hypothesised that nail psoriasis is related to enthesitis.[32]

Further confirmation of this cluster analysis in real-life situations (eg, application to data from real-world evidence studies, such as PsABIO),[21] could support with defining treatment strategies, such as intensified and comprehensive early treatment approaches.

One strength of this analysis is use of robust data from two large-scale clinical trials with well-documented baseline characteristics. The analysis included only bio-naïve patients, resulting in a homogenous cohort, and response rates by cluster were consistent across multiple endpoints. Use of unsupervised machine learning provides an objective hypothesis-free method of classification without the bias of human interference.

However, this unsupervised technique and lack of defined number of clusters are also limitations of this study. We recognise this is a novel trend in research requiring further validation. We acknowledge the importance of physicians and patients being able to understand the applicability of these artificial intelligence methods and the caution needed in interpretation, particularly in clusters with low patient numbers.[33] Additionally, it is widely accepted that patients in clinical trials are not necessarily representative of patients in clinical practice, and potentially relevant variables may not have been captured at enrolment. The inclusion criteria that required patients to have ≥3 swollen and ≥3 tender joints (DISCOVER-1) and ≥5 swollen and ≥5 tender joints (DISCOVER-2) results in these clusters excluding patients with PsA and low joint involvement or isolated axial disease.

Our study adds valuable information to the current understanding of patterns of joint involvement and therapy response in patients with PsA. It demonstrates that unsupervised machine learning may support the identification of PsA phenotype clusters differentiated by demographic and clinical features as well as MDA and DAPSA responses. Using this method, the IL-23 inhibitor guselkumab was shown to be effective across clusters, with continuous improvement through week 52. Given that optimal clinical management of PsA requires a personalised treatment approach based on each patient's specific clinical manifestations, larger future studies could apply phenotype clusters to real-world cohorts to progress our understanding of disease heterogeneity in PsA and optimise individualised treatment selection.

Author affiliations
[1]Rheumatology Department, AP-HP, Lariboisière Hospital, INSERM U1132, Paris, France
[2]Université Paris Cité, Paris, France
[3]Rheumatology, Erasmus MC Universitair Medisch Centrum, Rotterdam, The Netherlands
[4]Department of Rheumatology and Clinical Immunology, Charité – Universitätsmedizin Berlin, Berlin, Germany
[5]Department of Rheumatology, Royal National Hospital for Rheumatic Diseases, Bath, UK
[6]Department of Life Sciences, Centre for Therapeutic Innovation, University of Bath, Bath, UK
[7]Rheumatology Department, Hospital Universitari Clínic and IDIBAPS, Universidad de Barcelona, Barcelona, Spain
[8]EMEA Medical Affairs Immunology, Janssen-Cilag, Zug, Switzerland
[9]Janssen Pharmaceutica, Beerse, Belgium
[10]EMEA Medical Affairs Immunology, Janssen-Cilag, Solna, Sweden
[11]EMEA Medical Affairs Immunology, Janssen-Cilag BV, Breda, The Netherlands
[12]Medical Affairs, Janssen Global Services LLC, Horsham, Pennsylvania, USA
[13]Research and Development, Immunology, Janssen Research and Development LLC, San Diego, California, USA
[14]Department of Medical and Biological Sciences, Rheumatology Institute, University of Udine, Udine, Italy

ORCID iDs
Pascal Richette http://orcid.org/0000-0003-2132-4074
Sarah Ohrndorf http://orcid.org/0000-0001-5943-4688
William Tillett http://orcid.org/0000-0001-7531-4125
Miriam Zimmermann http://orcid.org/0000-0001-6063-3000
Alen Zabotti http://orcid.org/0000-0002-0573-464X

REFERENCES

1 Coates LC, Helliwell PS. Psoriatic arthritis: state of the art review. *Clin Med (Lond)* 2017;17:65–70.
2 McInnes IB, Rahman P, Gottlieb AB, *et al*. Efficacy and safety of guselkumab, an interleukin-23p19-specific monoclonal antibody, through one year in biologic-naive patients with psoriatic arthritis. *Arthritis Rheumatol* 2021;73:604–16.
3 Cimmino MA. Epidemiology of psoriasis and psoriatic arthritis. *Reumatismo* 2007;59 Suppl 1:19–24.
4 Scotti L, Franchi M, Marchesoni A, *et al*. Prevalence and incidence of psoriatic arthritis: a systematic review and meta-analysis. *Semin Arthritis Rheum* 2018;48:28–34.
5 Coates LC, Soriano ER, Corp N, *et al*. Group for Research and Assessment of Psoriasis and Psoriatic Arthritis (GRAPPA): updated treatment recommendations for psoriatic arthritis 2021. *Nat Rev Rheumatol* 2022;18:465–79.
6 Perez-Chada LM, Merola JF. Comorbidities associated with psoriatic arthritis: review and update. *Clin Immunol* 2020;214:108397.
7 Kamalaraj N, El-Haddad C, Hay P, *et al*. Systematic review of depression and anxiety in psoriatic arthritis. *Int J Rheum Dis* 2019;22:967–73.
8 Queiro R, Lorenzo A, Tejón P, *et al*. Obesity in psoriatic arthritis: comparative prevalence and associated factors. *Medicine (Baltimore)* 2019;98:e16400.
9 Gossec L, Baraliakos X, Kerschbaumer A, *et al*. EULAR recommendations for the management of psoriatic arthritis with pharmacological therapies: 2019 update. *Ann Rheum Dis* 2020;79:700–12.
10 Gratacós J, Behrens F, Coates LC, *et al*. A 12-point recommendation framework to support advancement of the multidisciplinary care of psoriatic arthritis: a call to action. *Joint Bone Spine* 2021;88:105175.
11 Coates LC, Gossec L, Theander E, *et al*. Efficacy and safety of guselkumab in patients with active psoriatic arthritis who are inadequate responders to tumour necrosis factor inhibitors: results through one year of a phase IIIB, randomised, controlled study (COSMOS). *Ann Rheum Dis* 2022;81:359–69.
12 Deodhar A, Helliwell PS, Boehncke W-H, *et al*. Guselkumab in patients with active psoriatic arthritis who were biologic-naive or had previously received TNFα inhibitor treatment (DISCOVER-1): a double-blind, randomised, placebo-controlled phase 3 trial. *Lancet* 2020;395:1115–25.
13 Mease PJ, Rahman P, Gottlieb AB, *et al*. Guselkumab in biologic-naive patients with active psoriatic arthritis (DISCOVER-2): a double-blind, randomised, placebo-controlled phase 3 trial. *Lancet* 2020;395:1126–36.
14 Shoop-Worrall SJW, Cresswell K, Bolger I, *et al*. Nothing about us without us: involving patient collaborators for machine learning applications in rheumatology. *Ann Rheum Dis* 2021;80:1505–10.
15 Eng SWM, Aeschlimann FA, van Veenendaal M, *et al*. Patterns of joint involvement in juvenile idiopathic arthritis and prediction of disease course: a prospective study with multilayer non-negative matrix factorization. *PLoS Med* 2019;16:e1002750.
16 Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature* 1999;401:788–91.
17 Coates LC, Fransen J, Helliwell PS. Defining minimal disease activity in psoriatic arthritis: a proposed objective target for treatment. *Ann Rheum Dis* 2010;69:48–53.
18 Mease PJ. Measures of psoriatic arthritis: tender and swollen joint assessment, Psoriasis Area and Severity Index (PASI), Nail Psoriasis Severity Index (NAPSI), modified Nail Psoriasis Severity Index (mNAPSI), Mander/Newcastle Enthesitis Index (MEI), Leeds Enthesitis Index (LEI), Spondyloarthritis Research Consortium of Canada (SPARCC), Maastricht Ankylosing Spondylitis Enthesis Score (MASES), Leeds Dactylitis Index (LDI), Patient Global for Psoriatic Arthritis, Dermatology Life Quality Index (DLQI), Psoriatic Arthritis Quality of Life (PsAQOL), Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F), Psoriatic Arthritis Response Criteria (PsARC), Psoriatic Arthritis Joint Activity Index (PsAJAI), Disease Activity in Psoriatic Arthritis (DAPSA), and Composite Psoriatic Disease Activity Index (CPDAI). *Arthritis Care Res (Hoboken)* 2011;63 Suppl 11:S64–85.
19 Aletaha D, Alasti F, Smolen JS. Disease activity states of the DAPSA, a psoriatic arthritis specific instrument, are valid against functional status and structural progression. *Ann Rheum Dis* 2017;76:418–21.
20 Gorlier C, Orbai A-M, Puyraimond-Zemmour D, *et al*. Comparing patient-perceived and physician-perceived remission and low disease activity in psoriatic arthritis: an analysis of 410 patients from 14 countries. *Ann Rheum Dis* 2019;78:201–8.
21 Smolen JS, Siebert S, Korotaeva TV, *et al*. Effectiveness of IL-12/23 inhibition (ustekinumab) versus tumour necrosis factor inhibition in psoriatic arthritis: observational PsABIO study results. *Ann Rheum Dis* 2021;80:1419–28.
22 McGonagle D, Tan AL, Watad A, *et al*. Pathophysiology, assessment and treatment of psoriatic dactylitis. *Nat Rev Rheumatol* 2019;15:113–22.
23 Mease PJ, Gottlieb AB, van der Heijde D, *et al*. Efficacy and safety of abatacept, a T-cell modulator, in a randomised, double-blind, placebo-controlled, phase III study in psoriatic arthritis. *Ann Rheum Dis* 2017;76:1550–8.
24 Mease P, Hall S, FitzGerald O, *et al*. Tofacitinib or adalimumab versus placebo for psoriatic arthritis. *N Engl J Med* 2017;377:1537–50.
25 Wells AF, Edwards CJ, Kivitz AJ, *et al*. Apremilast monotherapy in DMARD-naive psoriatic arthritis patients: results of the randomized, placebo-controlled PALACE 4 trial. *Rheumatology (Oxford)* 2018;57:1253–63.
26 Nash P, Mease PJ, McInnes IB, *et al*. Efficacy and safety of secukinumab administration by autoinjector in patients with psoriatic arthritis: results from a randomized, placebo-controlled trial (FUTURE 3). *Arthritis Res Ther* 2018;20:47.
27 Mease PJ, Helliwell PS, Gladman DD, *et al*. Efficacy of guselkumab on axial involvement in patients with active psoriatic arthritis and sacroiliitis: a post-hoc analysis of the phase 3 DISCOVER-1 and DISCOVER-2 studies. *Lancet Rheumatology* 2021;3:e715–23.
28 Baraliakos X, Gossec L, Pournara E, *et al*. Secukinumab in patients with psoriatic arthritis and axial manifestations: results from the double-blind, randomised, phase 3 MAXIMISE trial. *Ann Rheum Dis* 2021;80:582–90.
29 ClinicalTrials.gov. A study of guselkumab administered subcutaneously in bio-naive participants with active psoriatic arthritis axial disease (STAR). Available: https://clinicaltrials.gov/ct2/show/NCT04929210 [Accessed Dec 2022].
30 Tillett W, Shaddick G, Jadon D, *et al*. Novel composite radiographic score for longitudinal observational studies of psoriatic arthritis: a proof-of-concept study. *J Rheumatol* 2016;43:367–70.
31 Helliwell PS, Hetthen J, Sokoll K, *et al*. Joint symmetry in early and late rheumatoid and psoriatic arthritis: comparison with a mathematical model. *Arthritis Rheum* 2000;43:865–71.
32 McGonagle D. Enthesis: an autoinflammatory lesion linking nail and joint involvement in psoriatic disease. *J Eur Acad Dermatol Venereol* 2009;23 Suppl 1:9–13.
33 Watson DS, Krutzinna J, Bruce IN, *et al*. Clinical applications of machine learning algorithms: beyond the black box. *BMJ* 2019;364:l886.