Rheumatology Advances in Practice, 2025, **9(2)**, rkaf031
https://doi.org/10.1093/rap/rkaf031
Original Article

British Society for Rheumatology
RHEUMATOLOGY ADVANCES IN PRACTICE
OXFORD

Clinical science

Disease burden in inflammatory arthritis: an unsupervised machine learning approach of the COVAD-2 e-survey dataset

Vincenzo Venerito[1,‡], Sergio Del Vescovo[1,‡], Sergio Prieto-González[2], Marco Fornaro[1], Lorenzo Cavagna[3,4], Florenzo Iannone[1], Masataka Kuwana[5], Vishwesh Agarwal[6], Jessica Day[7], Mrudula Joshi[8], Sreoshy Saha[9], Kshitij Jagtap[10], Wanruchada Katchamart[11], Phonpen Akarawatcharangura Goo[12], Binit Vaidya[13], Tsvetelina Velikova[14], Parikshit Sen[15], Samuel Katsuyuki Shinjo[16], Ai Lyn Tan[17], Nelly Ziade[18,19], Marcin Milchert[20], Abraham Edgar Gracia-Ramos[21], Carlo V Caballero-Uribe[22], COVAD Study[§], Hector Chinoy[23], Latika Gupta[23,24,25,*], Vikas Agarwal[26]

[1]Department of Precision and Regenerative Medicine and Ionian Area (DiMePre-J), University of Bari Aldo Moro, Bari, Italy
[2]Department of Internal Medicine, Hospital Clinic de Barcelona, Barcelona, Spain
[3]Division of Rheumatology, Faculty of Medicine and Surgery, University of Pavia, Pavia, Italy
[4]Fondazione IRCCS Policlinico San Matteo, Pavia, Italy
[5]Department of Allergy and Rheumatology, Nippon Medical School, Tokyo, Japan
[6]Mahatma Gandhi Mission Medical College, Maharashtra, India
[7]Clinical Translation, Walter and Eliza Hall Institute of Medical Research, Melbourne, VIC, Australia
[8]Byramjee Jeejeebhoy Government Medical College and Sassoon General Hospitals, Pune, India
[9]Mymensingh Medical College, Mymensingh, Bangladesh
[10]Seth Gordhandas Sunderdas Medical College, Mumbai, India
[11]Division of Rheumatology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand
[12]Department of Medicine, Queen Savang Vadhana Memorial Hospital, Chonburi, Thailand
[13]Department of Rheumatology, National Centre for Rheumatic Diseases, Kathmandu, Nepal
[14]Medical Faculty, Sofia University St. Kliment Ohridski, Sofia, Bulgaria
[15]Maulana Azad Medical College, New Delhi, India
[16]Department of Rheumatology, Faculdade de Medicina da, Universidade de São Paulo, São Paulo, Brazil
[17]Leeds Institute of Rheumatic and Musculoskeletal Medicine and Leeds Musculoskeletal Biomedical Research Unit, University of Leeds, Leeds, UK
[18]Rheumatology Department, Saint-Joseph University, Beirut, Lebanon
[19]Department of Rheumatology, Hotel-Dieu De France Hospital, Beirut, Lebanon
[20]Department of Rheumatology, Pomorski Uniwersytet Medyczny w Szczecinie, Szczecin, Poland
[21]Department of Internal Medicine, General Hospital, National Medical Center "La Raza", Instituto Mexicano del Seguro Social Delegacion Distrito Federal Norte, Mexico City, Mexico
[22]Department of Medicine, Hospital Universidad del Norte, Barranquilla, Colombia
[23]Division of Musculoskeletal and Dermatological Sciences, Centre for Musculoskeletal Research, School of Biological Sciences, The University of Manchester, Manchester, UK
[24]School of Infection, Inflammation and Immunology, College of Medicine and Health, University of Birmingham, Birmingham, UK
[25]Department of Rheumatology, Royal Wolverhampton Hospitals NHS Trust, Wolverhampton, UK
[26]Department of Clinical Immunology and Rheumatology, Sanjay Gandhi Post Graduate Institute of Medical Sciences, Lucknow, India

*Correspondence to: Latika Gupta, School of Infection, Inflammation and Immunology, College of Medicine and Health, University of Birmingham, Birmingham, UK. E-mail: drlatikagupta@gmail.com

[‡]V.V. and S.D.V. contributed equally.

[§]The complete list of authors of the COVAD Study Group as well as their affiliations are provided in the supplement.

Received: 4 September 2024. **Accepted:** 15 January 2025

Abstract

Objectives: To comprehensively compare the disease burden among patients with RA, PsA and AS using Patient-Reported Outcome Measurement Information System (PROMIS) scores and to identify distinct patient clusters based on comorbidity profiles and PROMIS outcomes.

Methods: Data from the global COVID-19 Vaccination in Autoimmune Diseases (COVAD) 2 e-survey were analysed. Patients with RA, PsA or AS undergoing treatment with DMARDs were included. PROMIS scores (global physical health, global mental health, fatigue 4a and physical function short form 10a), comorbidities and other variables were compared among the three groups, stratified by disease activity status. Unsupervised hierarchical clustering with eXtreme Gradient Boosting feature importance analysis was performed to identify patient subgroups based on comorbidity profiles and PROMIS outcomes.

Results: The study included 2561 patients (1907 RA, 311 PsA, 343 AS). After adjusting for demographic factors, no significant differences in PROMIS scores were observed among the three groups, regardless of disease activity status. Clustering analysis identified four distinct patient groups: low burden, comorbid PsA/AS, low burden with depression and high-burden RA. Feature importance analysis revealed PROMIS global physical health as the strongest determinant of cluster assignment, followed by depression and diagnosis. The comorbid PsA/AS and high-burden RA clusters showed a higher prevalence of comorbidities (56.47% and 69.7%, respectively) and depression (41.18% and 41.67%, respectively), along with poorer PROMIS outcomes.

Conclusion: Disease burden in inflammatory arthritis is determined by a complex interplay of factors, with physical health status and depression playing crucial roles. The identification of distinct patient clusters suggests the need for a paradigm shift towards more integrated care approaches that equally emphasize physical and mental health, regardless of the underlying diagnosis.

Lay Summary

What does this mean for patients?

This study analysed data from people with rheumatoid arthritis (RA), psoriatic arthritis (PsA) and ankylosing spondylitis (AS) to understand how physical health, mental well-being and other health conditions impact quality of life. Researchers analysed responses from 2561 people worldwide and grouped them based on their health profiles and self-reported symptoms. Key findings showed that physical and mental health struggles, fatigue and daily function were similar across RA, PsA and AS after accounting for age and sex. Four patient clusters were found: those with mild symptoms and few comorbidities, people with PsA/AS with higher comorbidities and depression, people with arthritis and depression but minimal symptoms and comorbidities and RA patients with severe symptoms and multiple comorbidities. These results highlight that arthritis care should focus not just on joints, but also on mental health and comorbid conditions. Doctors are encouraged to work on personalized care plans that treat the person as a whole, not just the arthritis diagnosis.

Keywords: inflammatory arthritis, disease burden, spondyloarthritis, rheumatoid arthritis, patient reported outcome, PROMIS (Patient-Reported Outcome Measurement Information System), mental health, comorbidities, cluster analysis, survey.

Key messages

- Physical and mental comorbidities have a greater influence on disease burden compared with the type of disease.
- PROMIS global physical health and depression are the strongest determinants of patient clustering in inflammatory arthritis.
- Disease burden requires an integrated care approach emphasizing both physical and mental health.

Introduction

RA, PsA and axial SpA (axSpA) are among the most prevalent chronic inflammatory rheumatic diseases, collectively affecting millions of individuals worldwide. RA is characterized by symmetrical polyarthritis primarily affecting the small joints, leading to progressive joint damage and disability [1]. PsA, associated with psoriasis, presents with peripheral arthritis, enthesitis, dactylitis and axial involvement [2]. AS primarily affects the sacroiliac joints and spine, causing inflammatory back pain, stiffness and potential spinal fusion [3]. Despite their distinct features, these debilitating conditions share common clinical manifestations such as joint pain, stiffness, fatigue and reduced physical function, posing a substantial physical and emotional burden on patients' lives [4–7]. A comprehensive evaluation of disease burden from both the physicians' and patients' perspectives is pivotal in guiding therapeutic decisions and optimizing disease management strategies tailored to individual needs.

While previous research has shed light on the comparative disease burden across RA, PsA and AS [8], the existing evidence remains limited, particularly in capturing the patient's voice and experiences through large-scale, multinational studies. The burden of these autoimmune inflammatory diseases, extending beyond objective clinical measures, warrants further exploration to ensure holistic patient care.

Disease burden encompasses not only the physical manifestations of the disease, but also the psychological, social and economic impact on patients' lives. Fatigue, pain, reduced mobility and decreased quality of life are among the many facets of disease burden that can significantly impair patients' daily functioning and well-being [6, 9, 10]. Moreover, the presence of comorbidities, such as cardiovascular diseases, metabolic disorders and mental health issues, can further complicate the disease course and magnify the overall burden experienced by patients [11]. The presence of comorbidities has been associated with increased disease activity, reduced physical function, poorer quality of life, development of difficult-to-treat forms and increased mortality in these patients [12–16]. Comorbidities may also complicate treatment decisions, as certain conditions may contraindicate the use of some therapies or require careful monitoring and management. Moreover, patients with RA, PsA and AS who have multiple comorbidities often have complex healthcare needs that extend beyond the management of their primary rheumatic disease. The cumulative burden of multiple chronic conditions can pose significant challenges for patients and their healthcare providers [17].

Understanding the multifaceted nature of disease burden is crucial for tailoring management strategies that address

patients' diverse needs. By incorporating patient-reported outcomes (PROs) and considering the psychosocial aspects of living with chronic inflammatory conditions, healthcare providers can work towards improving patients' overall health status and minimizing the long-term consequences of these diseases.

Recognizing this need, the COVID-19 Vaccination in Autoimmune Diseases (COVAD) study group, comprising physicians across healthcare centres globally, embarked on an e-survey initiative. The COVAD-1 survey collected data from 94 countries, focusing on domains such as physical function, fatigue, pain, disease activity and steroid/immunosuppressant use [18]. Building upon this foundation, the ongoing COVAD-2 survey expanded its scope, encompassing additional critical aspects like mental and social health, disease flares, long COVID, autoimmune multimorbidity and post-vaccination adverse events. Of note, data have been collected from a diverse pool of 109 countries [19].

In the analyses of rich datasets, unsupervised machine learning (ML) techniques could potentially enhance our understanding of disease burden in these conditions. Such methods may help identify hidden patterns in patient data, revealing subgroups with similar disease profiles or outcomes that are not immediately apparent through traditional statistical approaches [20].

Through the COVAD-2 database, the present study aimed to comprehensively compare the disease burden among patients with RA, AS and PsA by evaluating, in particular, their Patient-Reported Outcome Measurement Information System (PROMIS) scores [21]. Furthermore, an exploratory analysis was conducted to identify distinct patient clusters based on their comorbidity profiles and PROMIS outcomes, potentially highlighting subgroups with varying disease burdens and complex healthcare needs. In addition to PROMIS outcomes, the inclusion of comorbidities in the clustering analysis provides a comprehensive perspective on patients' disease burden, acknowledging that these conditions can significantly impact functional limitations and quality of life regardless of disease activity status. This integrated approach aims to identify distinct patient subgroups and understand how different combinations of disease characteristics, comorbidities and PROs contribute to overall disease burden.

Methods

Data source

The COVAD-2 dataset, generated by the COVAD study group, was employed for this study. The dataset was derived from a global e-survey that was patient self-reported, extensively pilot tested, validated and translated into 18 languages. The detailed methodology of the survey has been previously published. The COVAD study group comprised 156 physicians across healthcare centres in 109 countries, with data collected by contacting patient cohorts seen by collaborators in their clinics or via patient support groups [19].

Ethics approval

The study was approved by the Institutional Ethics Committee of the Sanjay Gandhi Postgraduate Institute of Medical Sciences, Lucknow (IEC code: 2021-143-IP-EXP-39) and informed consent was obtained by all participants.

Study population

The study included patients with a self-reported diagnosis of RA, PsA or AS who were undergoing active treatment with conventional synthetic DMARDs (csDMARDs) and/or biologic DMARDs (bDMARDs).

Data collection and variables

Data were collected between January and August 2022, extracted in August 2022 and the responses were screened. Data were then extracted from the COVAD-2 e-survey, which included information on patient demographics, diagnosis, disease activity, PROMIS scores, comorbidities and therapies. Among comorbidities, the presence of complex multimorbidity with mental health disorders (MHDs) was also extracted. The PROMIS scores assessed were global physical health, global mental health, physical function short form 10a (SF10a) and fatigue 4a [21]. Active disease was defined as the patient's perception of their disease as active in the 4 weeks prior to the survey.

Data cleaning and preprocessing

The survey responses were carefully screened and duplicate or incomplete responses were excluded. Data variables were categorized for assessment. The diagnosis of RA, PsA or AS was verified by rheumatologists. For patients reporting multiple diagnoses, the most likely diagnosis was determined based on the described symptoms to avoid erroneous or incomplete entries.

Statistical analysis

Descriptive statistics were used to summarize patient characteristics, with continuous variables reported as mean and S.D. for normally distributed variables. Categorical variables were expressed as numbers and percentages. Comparisons of PROs, age and sex between the RA, PsA and AS groups were performed separately for patients with active and inactive disease to explore potential differences among the three groups. Analysis of variance (ANOVA) with Bartlett's test and Tukey's test for post hoc pairwise comparisons were used for continuous variables such as age and sex, while the chi-squared test was employed for categorical variables. Moreover, to account for potential confounding by age and sex, we performed multivariate linear regression analysis, allowing us to compare the PROs across the disease groups while adjusting for these demographic factors.

Unsupervised ML techniques

Historically, principal component analysis (PCA) has been used for dimensionality reduction of large datasets [22]. It obtains the components by standardizing the values, building the covariance matrix and, from there, getting the eigenvalues and eigenvectors. PCA arrives at the final components that explain most of the variability in the dataset. Therefore, these new components can be considered linear combinations or composites of the original values. PCA works well with continuous variables, but variability in the dataset might fail to be depicted when categorical data are present, as in the COVAD-2 dataset.

To overcome such an issue, we implemented factor analysis with mixed data (FAMD), a generalization of PCA that can be deployed on large datasets with qualitative and quantitative variables. FAMD was implemented in a Python version 3.9 environment using Prince version 0.7.1, scipy version

1.9.3, scikit-learn version 1.1.3 and numpy version 1.23.2 on a MacBook Pro with Apple Silicon M1MAX 64 GB RAM.

The original COVAD-2 dataset included 2561 observations and 196 features. Briefly, accurate preprocessing with feature standardization and autocorrelation check was carried out for continuous variables. Observations with missing data were ablated. The Prince library automatically one-hot encoded any categorical variable and then divided it by the square root of the probability. The variable was then centred (i.e. subtraction of the mean). A PCA was then run on such adjusted features. One hundred repeats with different seeds were carried out and the 95% CI of the bootstrapped eigenvalues distribution was used to determine the number of FAMD components to retain for clustering (10 components were retained). To define the number of clusters, hierarchical clustering on principal components [23] was performed using Ward's criterion and a Euclidean distance metric on the selected FAMD principal components. The intercluster inertia gain was used to determine the optimal division level.

To assess the relative importance of different features in determining cluster assignments, we employed an eXtreme Gradient Boosting (XGBoost) model. The model was trained to predict cluster assignments using the same set of clinical features used in the initial clustering analysis. Feature importance scores were calculated based on their contribution to the model's prediction accuracy, providing insights into which characteristics were most influential in distinguishing between patient clusters.

After clustering, differences among the obtained clusters were evaluated using the same statistical tests (ANOVA with Tukey's test for continuous variables and chi-squared test for categorical variables) for PROs; the prevalence of RA, PsA and AS; depression; comorbidities; age and sex (Fig. 1).

Results

Descriptive statistics

The study included a total of 2561 patients with inflammatory arthritis, comprising 1907 with RA, 311 with PsA and 343 with AS. The demographic and clinical characteristics of the study population are presented in Table 1.

The majority of patients across all three groups were female, with the highest percentage in the RA group (87.6%), followed by the PsA group (67.2%) and the AS group (36.7%). The mean age of patients was lower in the AS group [43.1 years (s.d. 12.8)] compared with the RA [50.9 years (s.d. 13.7)] and PsA [50.4 years (s.d. 12.7)] groups.

Comorbidities were common across all groups: 42.9% of RA patients, 56.0% of PsA patients and 38.3% of AS patients reported at least one comorbid condition. Depression was the most frequently reported mental health disorder, especially in the PsA group (22.2%), followed by the RA (14.8%) and AS (10.8%) groups. The prevalence of complex multimorbidity, including mental health disorders, was similar across the groups.

In terms of treatment, steroid use was more frequent in the RA group (31.1%) than in the PsA (12.2%) and AS (15.2%) groups. Methotrexate was the most commonly used csDMARD, with the highest use in the RA group (45.6%), followed by the PsA (35.4%) and AS (19.5%) groups. bDMARDs, especially anti-TNF agents, were used more frequently in the PsA and AS groups, while anti-TNF use was lower in the RA group (11.5%) compared with PsA (21.5%) and AS (25.7%) groups. Janus kinase (JAK) inhibitors, a targeted synthetic DMARD class, were used most in the AS group (5.3%) and least in the PsA group (1.6%).

PROs in active and inactive disease

To evaluate the impact of disease activity on PROs, we compared age, sex distribution and the PROMIS scores among patients with RA, PsA and AS, stratified by disease activity status (active or inactive) (Tables 2 and 3).

Active disease

In patients with active disease (Table 2), we observed significant differences in age and sex distribution across the disease groups. Patients with active AS were significantly younger [mean age 42.12 years (s.d. 15.39)] compared with those with active PsA [53.76 years (s.d. 13.71), $P < 0.01$] and active RA [50.28 years (s.d. 13.21), $P < 0.01$]. The proportion of female patients was also significantly higher in the active PsA group (73.9%) compared with the active AS group (62.9%, $P < 0.01$). Female patients consistently reported worse PROMIS scores and higher pain levels compared with male patients across all disease groups ($P < 0.001$ for all comparisons). After adjusting for these demographic differences (sex, age) using a multivariate linear regression model, we found no statistically significant differences in the PROMIS global physical health, global mental health, fatigue 4a and physical function SF10 scores among the three active disease groups (RA, PsA and AS). The mean visual analogue scale (VAS) pain scores were also similar across the active disease groups.

Inactive disease

Among patients with inactive disease (Table 3), age and sex distributions again varied significantly. Patients with inactive PsA [50.81 years (s.d. 12.02)] and RA [52.14 years (s.d. 13.41)] were significantly older than those with inactive AS [43.46 years (s.d. 12.82), $P < 0.001$ for both). The proportion

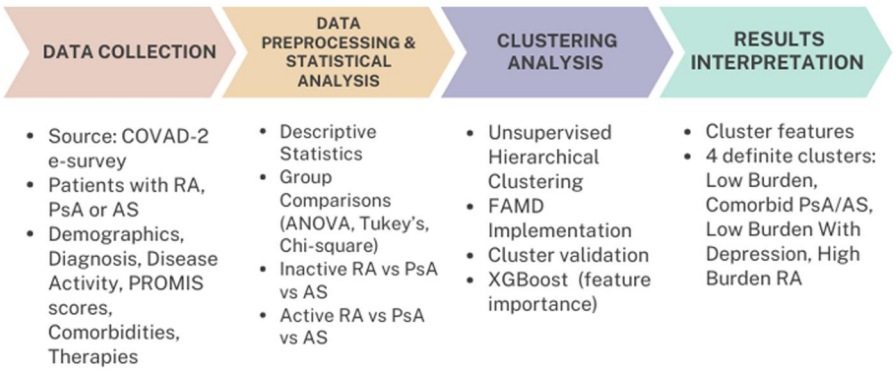

Figure 1. Workflow chart of the study methodology

Table 1. Demographic and clinical characteristics of the study population

Characteristics	RA		PsA		AS	
	Avg. obs.		Avg. obs.		Avg. obs.	
Female, *n* (%)	1907	1671 (87.62)	311	209 (67.20)	343	126 (36.73)
Age, years, mean (s.d.)	1907	50.95 (13.67)	311	50.42 (12.70)	343	43.13 (12.75)
Comorbidities, *n* (%)	1907	821 (42.99)	311	174 (55.96)	343	129 (38.29)
Depression, *n* (%)	1907	282 (14.79)	311	69 (22.19)	343	37 (10.79)
Complex multimorbidity with MHDs, *n* (%)	1907	110 (5.77)	311	24 (7.72)	343	21 (6.12)
Steroids, *n* (%)	1761	547 (31.06)	284	46 (12.20)	302	46 (15.23)
Methotrexate, *n* (%)	1907	870 (45.62)	311	110 (35.37)	343	67 (19.53)
Anti-TNF, *n* (%)	1907	220 (11.54)	311	67 (21.54)	343	88 (25.66)
JAK inhibitor, *n* (%)	1907	68 (3.57)	311	5 (1.61)	343	18 (5.25)

MHDs: mental health disorders; ns: not significant; avg. obs.: average observed.

Table 2. PROs in patients with active RA, PsA and AS

PROs	RA ($n = 189$)		PsA ($n = 38$)		AS ($n = 35$)	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
PROMIS global physical health	11.24	3.41	10.10	2.76	11.05	3.19
PROMIS global mental health	11.89	3.30	10.84	3.63	11.31	3.26
PROMIS fatigue 4a	11.75	4.68	12.84	4.42	12.94	4.87
PROMIS physical function SF10	34.90	9.80	33.52	8.76	35.82	9.62
VAS pain	4.68	2.61	5.0	2.54	4.68	2.77

Table 3. PROs in patients with inactive RA, PsA and AS

PROs	RA ($n = 1167$)		PsA ($n = 179$)		AS ($n = 185$)	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
PROMIS global physical health	12.48	2.90	12.43	3.27	13.13	2.95
PROMIS global mental health	12.84	3.17	12.97	3.33	13.31	3.36
PROMIS fatigue 4a	10.45	4.08	10.58	4.22	9.4	4.13
PROMIS physical function SF10	37.79	8.86	39.27	9.01	41.13	7.39
VAS pain	3.87	2.45	4.04	2.50	3.34	2.39

of female patients was also significantly higher in the inactive RA group (88%) compared with the inactive PsA (63.7%) and AS (38.4%) groups ($P < 0.001$ for both). Similar to the active disease group, female patients consistently demonstrated poorer scores across all PROMIS domains and higher pain levels compared with male patients ($P < 0.001$ for all comparisons). Also in this case, the multivariate regression analysis showed no statistically significant differences in the VAS pain, PROMIS global physical health, global mental health, fatigue 4a and physical function SF10a scores between the inactive disease groups of RA, PsA and AS.

Cluster analysis

Unsupervised hierarchical clustering on principal components was performed to identify distinct patients subgroups based on their comorbidity profiles and PROMIS outcomes. Four clusters were so identified: low burden, comorbid PsA/AS, low burden with depression and high-burden RA (Fig. 2, Table 4).

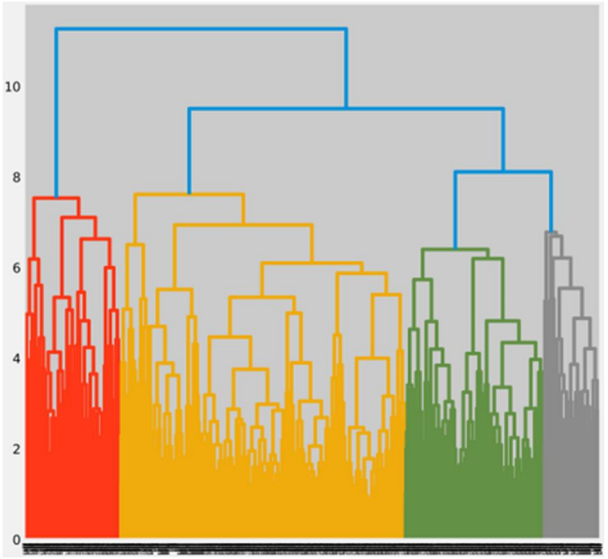

Figure 2. Dendrogram of the unsupervised hierarchical clustering on principal components of patients with RA, PsA and AS based on their comorbidity profiles and PROMIS scores. The coloured boxes indicate the assignment of patients to the four identified clusters

The XGBoost model revealed that PROMIS global physical health was the most influential feature in determining cluster assignments (Supplementary Fig. S1, available at *Rheumatology Advances in Practice* online). Depression emerged as the second most important feature, followed by diagnosis (0.10) and PROMIS physical function SF10a score. PROMIS global mental health, fatigue 4a and comorbidities showed moderate importance, while disease activity status, pain, age and gender demonstrated relatively lower importance in cluster determination.

Significant differences were observed in age, sex and diagnosis distributions among the clusters. Patients in the high-burden RA cluster were older [53.97 years (s.d. 12.72)] compared with those in the comorbid PsA/AS [47.94 years (s.d. 12.23)] and low burden with depression [48.76 years (s.d. 13.35)] clusters. The proportion of female patients was highest in the high-burden RA cluster (91.95%) and lowest in the comorbid PsA/AS cluster (76.19%, $P < 0.001$). RA predominated in the high-burden RA cluster (98.86%), while PsA was most frequent in the comorbid PsA/AS cluster (57.65%).

Table 4. Features of the four patient clusters identified through unsupervised hierarchical clustering on principal components. The clusters are compared in terms of age; sex; prevalence of RA, PsA and AS; PROMIS scores (global physical health, physical function SF10, global mental health, and fatigue 4a), VAS pain, comorbidities and depression.

Features	Cluster 1 (low burden)		Cluster 2 (comorbid PsA/AS)		Cluster 3 (low burden with depression)		Cluster 4 (high-burden RA)		*P*-value
	Obs.		Obs.		Obs.		Obs.		
Age, years, mean (S.D.)	553	52.81 (13.88)	84	47.94 (12.23)	874	48.76 (13.35)	263	53.97 (12.72)	<0.001
Female, *n* (%)	555	449 (80.90)	84	64 (76.19)	881	671 (76.16)	261	240 (91.95)	<0.001
RA, *n* (%)	560	403 (71.96)	85	0 (0)	884	692 (78.28)	264	261 (98.86)	<0.001
PsA, *n* (%)	560	74 (13.21)	85	49 (57.65)	884	92 (10.41)	264	2 (0.76)	<0.001
AS, *n* (%)	560	83 (14.82)	85	36 (42.35)	884	100 (11.31)	264	1 (0.38)	<0.001
PROMIS global physical health, mean (S.D.)	560	11.28 (1.53)	85	8.21 (1.81)	884	14.06 (2.03)	264	8.33 (1.61)	<0.001
PROMIS physical function SF10, mean (S.D.)	560	34.97 (6.59)	85	30.55 (9.12)	884	43.94 (4.73)	264	25.91 (7.43)	<0.001
PROMIS global mental health, mean (S.D.)	560	12.14 (2.24)	84	8.40 (2.07)	884	14.65 (2.50)	264	8.99 (2.56)	<0.001
PROMIS fatigue 4a, mean (S.D.)	560	11.58 (3.07)	84	16.07 (2.74)	884	7.89 (2.87)	264	15.81 (2.01)	<0.001
VAS pain, mean (S.D.)	560	4.91 (1.85)	85	6.26 (2.96)	884	2.37 (1.91)	264	6.51 (1.79)	<0.001
Depression, *n* (%)	560	0 (0)	85	35 (41.18)	884	113 (12.78)	264	110 (41.67)	<0.001
Comorbidity, *n* (%)	560	293 (52.32)	85	48 (56.47)	884	2464 (29.86)	264	184 (69.7)	<0.001
Inactive disease, *n* (%)	560	474 (84.64)	85	54 (63.53)	884	813 (91.97)	264	190 (71.97)	<0.001

Obs.: observed.

AxSpA was distributed across the clusters but was most prevalent in the comorbid PsA/AS (42.35%) cluster.

The low-burden cluster [$n = 560$, 52.81 years (S.D. 13.88), 80.90% female] included patients with fewer comorbidities (52.32%) and no depression. RA was the most common diagnosis (71.96%), followed by PsA (13.21%) and AS (14.82%). This cluster exhibited better PROMIS outcomes, with global physical health [11.28 (S.D. 1.53)], global mental health [12.14 (S.D. 2.24)] and physical function SF10 [34.97 (S.D. 6.59)] scores significantly higher than in the comorbid PsA/AS and high-burden RA clusters ($P < 0.001$). VAS pain was also lower [4.91 (S.D. 1.85), $P < 0.001$].

The comorbid PsA/AS cluster [$n = 85$, 47.94 years (S.D. 12.23), 76.19% female] was characterized by a high prevalence of PsA (57.65%) and AS (42.35%), with no RA patients. This cluster had the highest burden of comorbidities (56.47%) and a high prevalence of depression (41.18%). PROMIS outcomes were poorer, with global physical health [8.21(S.D. 1.81)], global mental health [8.40 (S.D. 2.07)] and physical function SF10 [30.55 (S.D. 9.12)] scores significantly lower than in the low-burden and low burden with depression clusters ($P < 0.001$). VAS pain was higher [6.26 (S.D. 2.96), $P < 0.001$].

The low burden with depression cluster [$n = 884$, 48.76 years (S.D. 13.35), 76.16% female] predominantly included patients with RA (78.28%) but also had smaller proportions of PsA (10.41%) and AS (11.31%). While depression was present in 12.78% of this group, the burden of comorbidities was relatively low (29.86%). PROMIS scores were the best among all clusters, with global physical health [14.06 (S.D. 2.03)], global mental health [14.65 (S.D. 2.50)] and physical function SF10 [43.94 (S.D. 4.73)] significantly higher than in other clusters ($P < 0.001$). PROMIS fatigue 4a [7.89 (S.D. 2.87)] and VAS pain [2.37 (S.D. 1.91)] were the lowest ($P < 0.001$).

The high-burden RA cluster [$n = 264$, 53.97 years (S.D. 12.72), 91.95% female] was almost exclusively composed of RA patients (98.86%), with minimal representation of PsA (0.76%) or AS (0.38%). This cluster had the highest prevalence of comorbidities (69.7%) and depression (41.67%). PROMIS outcomes were the poorest, with global physical health [8.33 (S.D. 1.61)], global mental health [8.99 (S.D. 2.56)] and physical function SF10 [25.91 (S.D. 7.43)] scores significantly lower than in other clusters ($P < 0.001$). VAS pain was also significantly higher [6.51 (S.D. 1.79), $P < 0.001$].

The clusters showed distinct patterns in disease activity status, with the low burden with depression cluster having the highest proportion of inactive disease (92.0%), followed by the low-burden cluster (84.6%), high-burden RA cluster (72.0%) and comorbid PsA/AS cluster (63.5%). Supplementary Tables S1 and S2, available at *Rheumatology Advances in Practice* online show the distribution of inactive disease and active disease patients, respectively, among the four clusters.

Discussion

The present study explored the disease burden in patients with RA, PsA and AS through the cluster analysis of data from the COVAD-2 e-survey, focusing particularly on PROs assessed using the PROMIS. The PROMIS is a well-validated and standardized set of PRO measures that cover various domains of physical, mental and social health. The use of PROMIS, including global physical health, global mental health, fatigue 4a and physical function SF10, allowed for meaningful comparisons across different patient populations and disease states, providing a comprehensive assessment of the patient experience. Notably, female patients consistently reported worse PROMIS scores and higher pain levels across all disease groups and activity states, highlighting the substantial impact of gender on disease burden perception. This finding aligns with previous research demonstrating gender-specific differences in pain perception and functional limitations in rheumatic diseases [24–26]. Moreover, it underscores the critical importance of considering gender when evaluating disease burden and planning therapeutic strategies.

Our analysis revealed no significant differences in PROMIS scores among the three groups, regardless of whether the disease was active or inactive, after adjusting for demographic confounding factors like age and sex. This finding highlights the shared experiences of disease burden among patients with RA, PsA and AS. In active disease, the physical and mental impacts, fatigue levels and physical function limitations were

similar across groups. In inactive disease, while clinical remission was achieved, residual symptoms such as fatigue, pain and reduced physical function persisted at comparable levels across conditions, emphasizing the need for a more holistic approach to patient care.

The unsupervised hierarchical clustering analysis identified four distinct patient clusters based on comorbidity profiles and PROMIS outcomes among patients. The comorbid PsA/AS and high-burden RA clusters were characterized by a higher prevalence of comorbidities and depression, as well as poorer PROs. Comorbid PsA/AS comprised patients with PsA and AS who had depression and comorbidities, resulting in lower PROMIS global physical health and global mental health scores, as well as higher PROMIS fatigue 4a scores. High-burden RA consisted mainly of RA patients with depression and comorbidities, also exhibiting poorer PROs. The low-burden and low burden with depression clusters included patients with RA, PsA and AS who had better PROs. Low burden represented patients with a lower burden of comorbidities and better health status, as indicated by higher PROMIS global physical health and global mental health scores, as well as lower PROMIS fatigue 4a scores. Low burden with depression included patients with a higher prevalence of mental health disorders but relatively better physical health status compared with the comorbid PsA/AS and high-burden RA clusters.

Thus the clustering analysis is consistent with previous studies that have demonstrated the negative impact of comorbidities, particularly mental health disorders, on health outcomes and quality of life in patients with rheumatic diseases [11–13, 27, 28]. Comorbid PsA/AS and high-burden RA are thereby characterized by a higher comorbidity burden and depression prevalence, while demonstrating consistently poorer PROs regardless of the underlying rheumatic diagnosis. This finding suggests that the presence of comorbidities, particularly depression, might be a more important determinant of patient outcomes than the specific type of inflammatory arthritis. The similar PRO patterns observed in these clusters, despite different underlying diagnoses, suggest common pathways of disease impact that might benefit from similar therapeutic approaches.

Moreover, the feature importance analysis using XGBoost revealed a hierarchical structure of factors influencing clusters assignment, with PROMIS global physical health emerging as the strongest determinant, followed by depression and diagnosis. The prominence of depression as the second most important feature is particularly noteworthy and aligns with growing evidence regarding the bidirectional relationship between mental health and inflammatory diseases [29, 30].

These findings have several important implications for clinical practice. First, they support the need for routine screening of depression and other comorbidities in all patients with inflammatory arthritis, regardless of their diagnosis or disease activity status. The high feature importance score for depression suggests that mental health assessment should be an integral part of routine care. Second, the identification of distinct patient clusters suggests that treatment strategies might be more effectively organized around patient profiles rather than diagnostic categories alone. For instance, patients in the comorbid PsA/AS and high-burden RA groups might benefit from more intensive psychological support and comorbidity management alongside their standard rheumatologic care.

The study has several limitations that should be acknowledged. First, the data were collected through a patient self-reported e-survey, which may be subject to recall bias and misclassification. However, the survey was extensively pilot tested, validated and translated to minimize these risks. Second, the number of patients with active disease was relatively small, which may limit the generalizability of the findings in this subgroup. Third, the study relied on patients' self-reported perception of disease activity rather than objective disease activity measures, such as the 28-joint DAS for RA or the Ankylosing Spondylitis Disease Activity Score for AS. While patient-reported disease activity provides valuable insights into the patient experience, future studies should consider incorporating both patient-reported and objective measures for a more comprehensive assessment. Lastly, the analysis did not account for disease duration, which may influence PROs and comorbidity burden. Future research should explore the impact of disease duration on the observed outcomes.

Despite these limitations, the study has several strengths. The COVAD-2 e-survey collected data from a large, diverse and multinational cohort of patients with RA, PsA and AS, enhancing the generalizability of the findings. The use of the PROMIS measures, which are well-validated and standardized PRO assessment tools, allows for meaningful comparisons across different patient populations and disease states. Furthermore, the unsupervised hierarchical clustering analysis provides a novel approach to identifying patient subgroups with distinct comorbidity profiles and PROs, enabling a more personalized understanding of the disease burden.

Conclusion

This study demonstrates that disease burden in inflammatory arthritis is determined by a complex interplay of factors, with physical health status and depression playing particularly crucial roles. The identification of distinct patient clusters, primarily driven by these factors, suggests the need for a paradigm shift in how we approach patient care in rheumatology. The finding that PROMIS global physical health and depression are the strongest determinants of patient clustering calls for a more integrated approach to patient care that places equal emphasis on physical and mental health. Future research should focus on several key areas. First, longitudinal studies are needed to understand how these patterns of disease burden evolve over time and respond to various therapeutic interventions. Second, intervention studies should evaluate the effectiveness of targeted approaches for the specific patient subgroups identified through this analysis. Finally, implementation research is needed to determine how best to integrate routine screening for depression and other key determinants of disease burden into standard rheumatology care.

Supplementary material

Supplementary material is available at *Rheumatology Advances in Practice* online.

Data availability

Data are available from the corresponding author upon reasonable request.

Authors' contributions

Conceptualisation, design, and analysis: V.V., S.D.V., L.G.; Writing – original draft: V.V., L.G., S.D.V.; Review and editing: S.P.-G., M.F., L.C., F.I., M.K., V.A., J.D., M.J., S.S., K.J., W.K., P.A.G., B.V., T.V., P.S., S.K.S., A.L.T., N.Z., M.M., A. E.G.-R., C.V.C.-U., H.C., COVAD Study Group, L.G., V.A.

Funding

No specific funding was received from any funding bodies in the public, commercial or not-for-profit sectors to carry out the work described in this manuscript.

Disclosure statement: A.L.T. has received honoraria for advisory boards and speaking from AbbVie, Gilead, Janssen, Lilly, Novartis, Pfizer and UCB. H.C. has received consulting fees as a speaker from GSK and UCB; is an advisory board member for AstraZeneca, Pfizer, Argenx and Galapagos and is Data and Science Monitoring Board chair for Horizon Therapeutics. J.D. has received research funding from CSL. M.K. has received consulting fees, speaking fees and research grants from AbbVie, Argenx, Asahi Kasei, AstraZeneca, Astellas, Boehringer Ingelheim, Chugai, Corbus, Eisai, GSK, Horizon, Janssen, Kissei, MBL, Mitsubishi Tanabe, Mochida, Nippon Shinyaku, Ono Pharmaceuticals, Pfizer and Taisho. N.Z. has received speaker fees, advisory board fees and research grants from Pfizer, Roche, AbbVie, Lilly, NewBridge, Sanofi-Aventis, Boehringer Ingelheim, Janssen and Pierre Fabre; none related to this manuscript. The remaining authors declare no conflicts of interest relevant to this manuscript.

Acknowledgements

The authors are grateful to all respondents for completing the questionnaire. The authors also thank the Myositis Association, Myositis India, Myositis UK, Myositis Support and Understanding, the Myositis Global Network, Deutsche Gesellschaft für Muskelkranke e.V. (DGM), Dutch and Swedish Myositis patient support groups, Cure JM, Cure IBM, Sjögren's India Foundation, Patients Engage, Scleroderma India, Lupus UK, Lupus Sweden, Emirates Arthritis Foundation, EULAR PARE, ArLAR research group, AAAA patient group, Myositis Association of Australia, APLAR myositis special interest group, Thai Rheumatism sassociation, PANLAR, AFLAR NRAS, Anti-Synthetase Syndrome support group, and various other patient support groups and organizations for their contribution to the dissemination of this survey. Finally, the authors wish to thank all members of the COVAD study group for their invaluable role in the data collection.

COVAD Study Group

Steering Committee: Elena Nikiphorou, Ioannis Parodis, Johannes Knitza, Ashima Makol, Carlos Enrique Toro Gutiérrez, Dzifa Dey, Chris Wincup, Nicola Dalbeth, Gerd-Rüdiger Burmester, Gouchun Wang.

COVAD-2 Study Group members: Arvind Nune, James B. Lilleker, John D. Pauling, Armen Yuri Gasparyan, Naveen R, Bhupen Barman, Yogesh Preet Singh, Prithvi Sanjeevkumar Gaur, Mahabaleshwar Mamadapur, Naitica Darooka, Praggya Yaadav, Babur Salim, Zoha Zahid Fazal, Mahnoor Javaid, Tulika Chatterjee, Aarat Patel, Laura Andreoli, Daniele Lini, Lisa S. Traboco, Syahrul Sazliyana Shaharir, Chou Luan Tan, Suryo Anggoro Kusumo Wibowo, Miguel A. Saavedra, Erick Adrian Zamora Tehozol, Jorge Rojas Serrano, Ignacio García-De La Torre, Iris J. Colunga-Pedraza, Javier Merayo-Chalico, Raquel Aranega, Jesús Loarce-Martos, Leonardo Santos Hoff, Karen Schreiber, Margarita Aleksandrovna Gromova, Or Aharonov, Melinda Nagy-Vincze, Elisa Palalane, Marie Hudson, Océane Landon-Cardinal, Russka Shumnalieva, Olena Zimba, Doskaliuk Bohdana, Ho So, Manuel Francisco Ugarte-Gil, Lyn Chinchay, José Proaño Bernaola, Victorio Pimentel, A.T. M. Tanveer Hasan, Tamer A. Gheita, Hanan Mohamed Fathi, Reem Hamdy A. Mohammed, Yi-Ming Chen, Ghita Harifi, Lina El Kibbi, Hussein Halabi, Yurilís Fuentes-Silva, Lilith Stange Nunez, Cristian Vergara M, Jossiell Then Báez, Rodrigo García Salinas, Radames Sierra-Zorita, Idania Escalante M, Wendy Calapaqui, Ivonne Quezada.

Other COVAD Investigators: Rajiv Ranjan, Avinash Jain, Sapan C. Pandya, Rakesh Kumar Pilania, Aman Sharma, Manesh Manoj M, Vikas Gupta, Chengappa G. Kavadichanda, Pradeepta Sekhar Patro, Sajal Ajmani, Sanat Phatak, Rudra Prosad Goswami, Abhra Chandra Chowdhury, Ashish Jacob Mathew, Padnamabha Shenoy, Ajay Asranna, Keerthi Talari Bommakanti, Anuj Shukla, Arunkumar R. Pande, Kunal Chandwar, Akanksha Ghodke, Sinan Kardeş, Döndü Üsküdar Cansu, Reşit Yıldırım, Margherita Giannini, François Maurier, Julien Campagne, Alain Meyer, Nicoletta Del Papa, Gianluca Sambataro, Atzeni Fabiola, Marcello Govoni, Simone Parisi, Elena Bartoloni Bocci, Gian Domenico Sebastiani, Enrico Fusaro, Marco Sebastiani, Luca Quartuccio, Franco Franceschini, Pier Paolo Sainaghi, Giovanni Orsolini, Rossella De Angelis, Maria Giovanna Danielli, Silvia Grignaschi, Alessandro Giollo, Alessia Alunno, Stylianos Tomaras, Fabian Nikolai Proft, Marie-Therese Holzer, Zoltán Griger, Ihsane Hmamouchi, Imane El Bouchti, Zineb Baba, Uyi Ima-Edomwonyi, Ibukunoluwa Dedeke, Emorinken Airenakho, Nwankwo Henry Madu, Abubakar Yerima, Hakeem Olaosebikan, Celestine Chibuzo Okwara, Becky A, Ouma Devi Koussougbo, Daman Langguth, Vidya Limaye, Merrilee Needham, Nilesh Srivastav, Wilmer Gerardo Rojas Zuleta, Álvaro Arbeláez, Javier Cajas, José António Pereira Silva, João Eurico Fonseca, Karoll Cabriza, Jonathan Losanto, Nelly Colaman, Antonio Cachafeiro-Vilar, Generoso Guerra Bautista, Enrique Julio Giraldo Ho, Hugo Alonzo, Carlos Benito Santiago Pastelin, Alejandro Quiñónez Obiols, Nilmo Chávez, Andrea Bran Ordóñez, Gil Alberto Reyes Llerena, Dina Arrieta, Eduardo Romero Hidalgo, Ricardo Saenz, Gabriela Arredondo, Akira Yoshida, Ran Nakashima, Shinji Sato, Yuko Kaneko, Takahisa Gono, Oliver Distler.

References

1. Di Matteo A, Bathon JM, Emery P. Rheumatoid arthritis. Lancet 2023;402:2019–33.
2. FitzGerald O, Ogdie A, Chandran V *et al.* Psoriatic arthritis. Nat Rev Dis Primer 2021;7:59.
3. Sieper J, Poddubnyy D. Axial spondyloarthritis. Lancet 2017; 390:73–84.
4. Backman CL. Arthritis and pain. Psychosocial aspects in the management of arthritis pain. Arthritis Res Ther 2006;8:221.
5. Cohen SP, Vase L, Hooten WM. Chronic pain: an update on burden, best practices, and new advances. Lancet 2021;397:2082–97.
6. Husni ME, Merola JF, Davin S. The psychosocial burden of psoriatic arthritis. Semin Arthritis Rheum 2017;47:351–60.

7. Cutolo M, Kitas GD, van Riel PLCM. Burden of disease in treated rheumatoid arthritis patients: going beyond the joint. Semin Arthritis Rheum 2014;43:479–88.
8. Michelsen B, Fiane R, Diamantopoulos AP *et al.* A comparison of disease burden in rheumatoid arthritis, psoriatic arthritis and axial spondyloarthritis. PLoS One 2015;10:e0123582.
9. Kilic G, Kilic E, Nas K, Kamanlı A, Tekeoglu İ. Residual symptoms and disease burden among patients with psoriatic arthritis: is a new disease activity index required? Rheumatol Int 2019;39:73–81.
10. Liu V, Fong W, Kwan YH, Leung YY. Residual disease burden in patients with axial spondyloarthritis and psoriatic arthritis despite low disease activity states in a multiethnic Asian population. J Rheumatol 2021;48:677–84.
11. Shah K, Paris M, Mellars L, Changolkar A, Mease PJ. Real-world burden of comorbidities in US patients with psoriatic arthritis. RMD Open 2017;3:e000588.
12. Ballegaard C, Skougaard M, Guldberg-Møller J *et al.* Comorbidities, pain and fatigue in psoriatic arthritis, psoriasis and healthy controls: a clinical cohort study. Rheumatology (Oxford) 2021;60:3289–300.
13. Batko B, Urbański K, Świerkot J *et al.* Comorbidity burden and clinical characteristics of patients with difficult-to-control rheumatoid arthritis. Clin Rheumatol 2019;38:2473–81.
14. Ferguson LD, Siebert S, McInnes IB, Sattar N. Cardiometabolic comorbidities in RA and PsA: lessons learned and future directions. Nat Rev Rheumatol 2019;15:461–74.
15. Scriffignano S, Perrotta FM, De Socio A, Lubrano E. Role of comorbidities in spondyloarthritis including psoriatic arthritis. Clin Rheumatol 2019;38:3–10.
16. Roodenrijs NMT, van der Goes MC, Welsing PMJ *et al.* Difficult-to-treat rheumatoid arthritis: contributing factors and burden of disease. Rheumatology (Oxford) 2021;60:3778–88.
17. Zhao SS, Jones GT, Macfarlane GJ *et al.* Comorbidity and response to TNF inhibitors in axial spondyloarthritis: longitudinal analysis of the BSRBR-AS. Rheumatology (Oxford) 2021;60:4158–65.
18. Sen P, Gupta L, Lilleker JB *et al.* COVID-19 vaccination in autoimmune disease (COVAD) survey protocol. Rheumatol Int 2022;42:23–9.
19. Fazal ZZ, Sen P, Joshi M *et al.* COVAD survey 2 long-term outcomes: unmet need and protocol. Rheumatol Int 2022;42:2151–8.
20. Dalmaijer ES, Nord CL, Astle DE. Statistical power for cluster analysis. BMC Bioinformatics 2022;23:205.
21. Khanna D, Krishnan E, Dewitt EM *et al.* Patient-Reported Outcomes Measurement Information System (PROMIS®)—the future of measuring patient reported outcomes in rheumatology. Arthritis Care Res 2011;63(Suppl 11):S486–90.
22. Lever J, Krzywinski M, Altman N. Principal component analysis. Nat Methods 2017;14:641–2.
23. Husson F, Josse J, Pagès J. Principal component methods-hierarchical clustering-partitional clustering: why would we need to choose for visualizing data ? Technical Report – Agrocampus. Applied Mathematics Department, September 2010. http://www.agrocampus-ouest.fr/math/
24. Bartley EJ, Fillingim RB. Sex differences in pain: a brief review of clinical and experimental findings. Br J Anaesth 2013;111:52–8.
25. Athnaiel O, Cantillo S, Paredes S, Knezevic NN. The role of sex hormones in pain-related conditions. Int J Mol Sci 2023;24:1866.
26. Barcelon E, Chung S, Lee J, Lee SJ. Sexual dimorphism in the mechanism of pain central sensitization. Cells 2023;12:2028.
27. Schmukler J, Li T, Gibson KA *et al.* Patient global assessment is elevated by up to 5 of 10 units in patients with inflammatory arthritis who screen positive for fibromyalgia (by FAST4) and/or depression (by MDS2) on a single MDHAQ. Semin Arthritis Rheum 2023;58:152151.
28. Lai TL, Au CK, Chung HY, Lau CS. Depression in psoriatic arthritis: related to socio-demographics, comorbid loads or disease activity? Int J Rheum Dis 2022;25:474–80.
29. Joshi N, Khanna R, Shah RM. Relationship between depression and physical activity, disability, burden, and health-related quality of life among patients with arthritis. Popul Health Manag 2015;18:104–14.
30. Nerurkar L, Siebert S, McInnes IB, Cavanagh J. Rheumatoid arthritis and depression: an inflammatory perspective. Lancet Psychiatry 2019;6:164–73.

Rheumatology Advances in Practice, 2025, 9, 1–9
https://doi.org/10.1093/rap/rkaf031
Original Article