Seminars in Arthritis and Rheumatism 57 (2022) 152097

Contents lists available at ScienceDirect

Seminars in Arthritis and Rheumatism

journal homepage: www.elsevier.com/locate/semarthrit

Characteristics associated with the perception of high-impact disease (PsAID ≥4) in patients with recent-onset psoriatic arthritis. Machine learning-based model

Rubén Queiro [a,*], Daniel Seoane-Mato [b], Ana Laiz [c], Eva Galíndez Agirregoikoa [d], Carlos Montilla [e], Hye-Sang Park [c], Jose A. Pinto-Tasende [f], Juan J. Bethencourt Baute [g], Beatriz Joven Ibáñez [h], Elide Toniolo [i], Julio Ramírez [j], Ana Serrano García [k], on behalf of Proyecto REAPSER Study Group

[a] Rheumatology Service and the Principality of Asturias Institute for Health Research (ISPA), Faculty of Medicine, Universidad de Oviedo, Avenida de Roma s/n, Asturias, Oviedo 33011, Spain
[b] Research Unit, Spanish Society of Rheumatology, Madrid, Spain
[c] Rheumatology and Autoimmune Disease Department, Hospital Universitari de la Santa Creu i Sant Pau, Barcelona, Spain
[d] Rheumatology Service, Hospital Universitario Basurto, Bilbao, Spain
[e] Rheumatology Service, Hospital Universitario de Salamanca, Salamanca, Spain
[f] Rheumatology Service-INIBIC, Complexo Hospitalario Universitario de A Coruña, A Coruña, Spain
[g] Rheumatology Service, Hospital Universitario de Canarias, Santa, Cruz de Tenerife, Spain
[h] Rheumatology Service, Hospital Universitario 12 de Octubre, Madrid, Spain
[i] Rheumatology Service, Hospital Universitari Son Llàtzer, Palma de Mallorca, Spain
[j] Arthritis Unit, Rheumatology Department, Hospital Clínic Barcelona, Barcelona, Spain
[k] Knowledge Engineering Institute, Universidad Autónoma de Madrid, Madrid, Spain

ARTICLE INFO

Keywords:
Arthritis
Psoriatic
Quality of life
Predictive model
Machine learning

ABSTRACT

Objectives: To evaluate which patient and disease characteristics are associated with the perception of high-impact disease (PsAID ≥4) in recent-onset psoriatic arthritis.
Methods: We performed a multicenter observational prospective study (2-year follow-up, regular annual visits). The study population comprised patients aged ≥18 years who fulfilled the CASPAR criteria and less than 2 years since the onset of symptoms. The dataset was generated using data for each patient at the 3 visits (baseline, first year, and second year of follow-up) matched with the PsAID values at each of the 3 visits. PsAID was categorized into two groups (<4 and ≥4). We trained a logistic regression model and random forest–type and XGBoost machine learning algorithms to analyze the association between the outcome measure and the variables selected in the bivariate analysis. A k-fold cross-validation with $k = 5$ was performed.
Results: The sample comprised 158 patients. Of the patients who attended the clinic, 45.8% scored PsAID ≥4 at baseline; 27.1%, at the first follow-up visit, and in 23.0%, at the second follow-up visit. The variables associated with PsAID ≥4 were, in decreasing order of importance: HAQ, pain, educational level, and physical activity. Higher HAQ (logistic regression coefficient 10.394; IC95% 7.777,13.011), higher pain (5.668; 4.016, 7.320), lower educational level (-2.064; -3.515, -0.613) and high level of physical activity (1.221; 0.158, 2.283) were associated with a higher frequency of PsAID ≥4. The mean values of the measures of validity of the algorithms were all ≥85%.
Conclusions: Despite the higher weight given to pain when scoring PsAID, we observed a greater influence of physical function on disease impact.

* Corresponding author.
E-mail address: rubenque7@yahoo.es (R. Queiro).

https://doi.org/10.1016/j.semarthrit.2022.152097

Available online 9 September 2022

Introduction

Psoriatic arthritis (PsA) affects a high percentage of patients with psoriasis and is one of the most common types of chronic arthritis treated by rheumatologists [1]. The multidomain nature of PsA should be adequately reflected when assessing disease activity and the impact on the patient's daily life [2].

The Psoriatic Arthritis Impact of Disease (PsAID) questionnaire is currently the standard tool for evaluating the impact of PsA on quality of life [3,4]. PsAID covers a series of domains, including pain, skin problems, physical function, sleep, ability to work and engage in leisure activities, as well as different aspects of psychosocial health. A PsAID value <4 is considered acceptable for the patient and, therefore, a therapeutic objective [3,4]. In addition, PsAID correlates well with other outcomes and is highly sensitive to changes in clinical routine [3–5].

A recent multicenter study found that patients with distal interphalangeal joint involvement, a family history of PsA, or high levels of C-reactive protein (CRP) were less likely to achieve low disease impact according to PsAID [6]. In another multicenter study, a PsAID score indicating high disease impact (≥4) was associated with female sex, tender joints, enthesitis, and comorbid conditions [7]. Nevertheless, these studies were carried out in patients with long-standing established disease. The patient and disease characteristics associated with high-impact PsAID when the disease has been present for a shorter period of time remain unknown.

Therefore, it is necessary to analyze the behavior of PsAID in recent-onset PsA. This information is essential if we are to better plan the care provided to affected patients. The objective of the present study was to evaluate which patient and disease characteristics are associated with the perception of high-impact disease (PsAID ≥4) in recent-onset PsA.

Material and methods

This work is part of the REAPSER study. The design of REAPSER has been described in detail elsewhere [8]. It is a multicenter observational prospective study (2-year follow-up, regular annual visits), promoted by the Spanish Society of Rheumatology. The study population comprised patients of both sexes aged ≥18 years who fulfilled the Classification Criteria for Psoriatic Arthritis (CASPAR) [9], with less than 2 years since the onset of symptoms attributable to the disease.

The intention at the baseline visit was to reflect the patient's situation before disease progress was modified by the treatments prescribed in the rheumatology department. In this sense, participants could not have been receiving methotrexate, leflunomide, or apremilast for more than 3 weeks after initiation and could not be receiving biologic disease-modifying antirheumatic drugs (DMARDs). These intervals were fixed considering that the mean time from initiation of treatment until onset of the response to therapy is 4 weeks in the case of synthetic DMARDs and 1 week in the case of biologic DMARDs. In cases where the patient had been receiving synthetic DMARDs for more than 3 weeks, we obtained confirmation from the investigating rheumatologist that the patient had not yet responded to treatment at the baseline visit; this information was sought in only 9 patients, and for all those involved, the time since initiation of synthetic DMARDs was under 2 months.

If patients with psoriasis receiving treatment with synthetic or biologic DMARDs developed PsA and were referred to the rheumatology department for diagnosis and management, then they could be included in the study, since this would not violate the criterion that the baseline visit reflected the situation of the patient before disease progress was modified by the treatment prescribed at the rheumatology clinic.

Patients were invited to participate consecutively at one of their scheduled visits to the rheumatologist. Inclusion began in November 2014 and ended in October 2016. A total of 25 centers from 11 of the 17 Spanish autonomous communities participated in the study.

All patients gave their informed consent to participate. The study centers assigned each participant an identification code in order to ensure data confidentiality in line with current legislation. The study complies with the Declaration of Helsinki and was approved by the Clinical Research Ethics Committees of the Principality of Asturias (study number 14/2014).

Variables and measurement

Variables included in REAPSER have been previously described [8]. For this work, we considered:

- Sociodemographic data: age; sex; educational level (none, primary, secondary, university).
- Family history of PsA, other types of inflammatory arthritis, and psoriasis.
- Personal history and comorbidities (based on a review of medical records): age-adjusted Charlson comorbidity index [10], cardiovascular risk factors (arterial hypertension, hyperlipidemia, diabetes mellitus [differentiating between insulin- and non–insulin-dependent]).
- Anthropometric data: body mass index (BMI).
- Lifestyle: smoking. Alcohol consumption [11]. Physical activity (low, moderate, and high) [12].
- Clinical situation at diagnosis of PsA: year of presentation of symptoms of PsA; clinical form (1. axial, 2. peripheral, 3. mixed); articular pattern (1. oligoarticular, 2. polyarticular, 3. distal, 4. mutilans, 5. spondylitis); presence of dactylitis (yes/no).
- Joint involvement and enthesitis: number of tender joints (NTJ68); number of swollen joints (NSJ66); extended version of the Maastricht Ankylosing Spondylitis Enthesitis Score (MASES) [13]. Polyarthritis was defined as NSJ66 ≥5.
- Pain and global assessment of disease during the previous week: pain, from 0 (no pain) to 10 (very intense); patient global assessment of disease, from 0 (feels very well) to 10 (feels very ill); physician global assessment of disease, from 0 (minimal activity) to 10 (maximum activity).
- Cutaneous and nail involvement (evaluated by a dermatologist): cutaneous psoriasis (yes/no); year of onset of psoriasis; clinical type; specific locations; treatment of psoriasis at PsA diagnosis;Psoriasis Area and Severity Index (PASI) [14]; onychopathy (number of digits affected). For purposes of the analysis, severe psoriasis was defined as PASI >10.
- Functional situation and quality of life: Health Assessment Questionnaire (HAQ) [15], Psoriatic Arthritis Impact of Disease (PsAID) [3]
- Radiographic evaluation at baseline: Bath Ankylosing Spondylitis Radiology Index (BASRI) of the sacroiliac region [16], hand involvement according to the modified Steinbrocker method for PsA [17].
- Laboratory tests: C-reactive protein (CRP), uric acid, total cholesterol, LDL cholesterol, triglycerides. For purposes of the analysis, a series of cut-off points were established to define high values: >0.5 mg/dl for standard CRP; >0.3 mg/dl for high-sensitivity CRP; hyperuricemia if >7 mg/dl in men and >6 mg/dl in women; ≥200 mg/dl for total cholesterol; ≥100 mg/dl for LDL; ≥150 mg/dl for triglycerides.
- Treatment of PsA with DMARDs, date of initiation, date of finalization: synthetic DMARDs (methotrexate, leflunomide, sulfasalazine, cyclosporine), targeted-synthetic DMARDs (apremilast), biologic DMARDs (adalimumab, etanercept, infliximab, golimumab, ustekinumab, certolizumab, secukinumab).

Sample size

As REAPSER study was planned as a registry intended to collect a large number of variables, without prespecified hypothesis, a sample size was not previously calculated for this work.

Statistical analysis

Imputation of missing data:

- The duration of psoriasis was imputed with the median of the remaining patients from the same age range. The age ranges used were as follows: <41 years, 41–60 years, and >60 years.
- Systemic treatment of psoriasis was imputed with 0 (that is, not receiving systemic treatment). The reason for this imputation was that when monitoring we observed that cases in which this data was not available were actually patients with no treatment or topical treatment. There were only two cases with missing information about systemic treatment of psoriasis that could not be compiled after monitoring.
- Radiological involvement of the hands at the baseline visit was not imputed, except for those patients with NTJ28 and NSJ28 value of 0, in which case it was imputed with 0.
- For patients who stopped attending the visits owing to improvement of their condition, the missing values for the variables PsAID and HAQ were imputed with 0.

Generation of the dataset:

We conducted a cross-sectional analysis. The dataset used for bivariate and multivariate analysis included 3 timepoints per patient. It was generated using data for each patient at the 3 visits (baseline, first year, and second year of follow-up) matched with the PsAID values at each of the 3 visits. Atemporal variables such as sex and family history were matched with outcome measures from each visit; therefore, their values are the same for each one. This was also true for variables that were only collected at the baseline visit, such as systemic treatment of psoriasis at PsA diagnosis and clinical form at diagnosis.

PsAID was categorized into two groups, namely, <4 (low-impact disease) and ≥4 (high-impact disease) [3].

Bivariate analysis:

We selected variables whose Spearman correlation was considered significant according to the threshold applied to the ρ correlation coefficient ($(|p| > \frac{2}{\sqrt{N}}$, with N being the number of data items). We also applied methods based on artificial intelligence, specifically the XGBoost algorithm and the SHAP technique, in order to identify informative variables (see Supplementary material for a detailed explanation of both approaches). Finally, of the variables identified in the previous steps, we selected those that were statistically significantly associated with the outcome measure ($p < 0.05$). To do so, we applied the Mann-Whitney test for continuous/discrete variables and the χ^2 test for categorical variables.

Multivariate analysis:

In order to generate models where the independent variables do not share information, and have a significant contribution to the model when adjusting for the rest of the variables included, we selected statistically significant variables (ie, $p < 0.05$) in an iterative fashion using logistic regression models based on artificial intelligence. The steps were performed in the 75% of the sample (training dataset) (see Supplementary material for a detailed explanation).

Next, based on the variables selected, random forest–type and XGBoost machine learning algorithms were trained to analyze the association between the outcome measure and the variables selected (see Supplementary material for more detail). To train the machine learning models the sample is split in two subsets, one to train the model and the other to evaluate its functioning. The division is generated in such a way that the proportion for each class of the outcome measure is the same in both subsets.

When the subsamples generated are imbalanced, the oversampling technique is used to train the models [18]. This is based on duplicating or triplicating those data whose value for the outcome variable is a minority value.

The parameters and thus the predictions of the trained algorithm might depend on the randomness that derives from the training/test split, which means that different splits of the data might result in different models. To reduce this effect, *k*-fold cross-validation was performed. Such method consists in splitting the original dataset into k subsets of the same size, and iteratively training the algorithm with *k*-1 of them while testing the model with the one left. After k iterations, the algorithm will have been trained and evaluated with all the partitions. In this analysis, a *k*-fold cross-validation with $k = 5$ has been used for the random forest and XGBoost (with the same subsets for both).

The contribution of the variables to the prediction of each iteration of the algorithms was calculated by the feature importance of each variable in the training subset. To estimate the performance of the random forest and XGBoost algorithms we calculated the values of accuracy, sensitivity, specificity, positive predictive value and negative predictive value as the mean of the values obtained for each parameter in the five evaluations performed in the cross validation.

Results

The sample eventually comprised 158 patients. The baseline characteristics of the sample have been previously published [19].

Thirty-three patients (20.9%) were lost to follow-up. The investigating rheumatologist at their center confirmed that 10 of these patients had not attended the visit because their PsA had improved.

Of the patients who attended the clinic, 45.8% scored PsAID ≥4 at baseline; 27.1%, at the first follow-up visit, and 23.0%, at the second follow-up visit.

Bivariate analysis

Table 1 shows the variables selected in the bivariate analysis.

Multivariate analysis

The number of observations for the multivariate analysis was 403.

Given that all patients with PASI >10 had PsAID ≥4, it was necessary to apply an L1 regularization in order to assign a coefficient in the logistic regression analysis. The regularization limits the magnitude of the regression coefficients so that the model can generalize for new data. In this case, given that all patients with PASI>10 had PsAID ≥4 in the training data, the model run without regularization assigned coefficients of $+\infty$ to the variable PASI, in such a way that if a patient had PASI>10, he/she would be always classified as PsAID ≥4. The L1 regularization limits the coefficient of the variable PASI >10 so that the model can envisage the case of a patient with PASI >10 having PsAID <4. In mathematical terms, L1 regularization is a technique used during the estimation of the regression coefficients, in which these are limited by adding the sum of the absolute values of the coefficients to the error function, which, in turn, when minimized reveals the coefficients [20].

Table 2 shows the results of the logistic regression analysis. Higher HAQ, higher pain during the previous week and lower educational level were associated with a higher frequency of PsAID ≥4. In the case of physical activity, a moderate level could be associated with a lower frequency of PsAID ≥4 (although this was not statistically significant), while a high level was associated with a higher frequency of PsAID ≥4.

When the random forest–type and XGBoost machine learning algorithms were trained with these 4 variables, the order of importance (from more to less) attributed by most of the models according to the values of feature importances (Table 3) was as follows: HAQ, pain in the previous week, educational level, and physical activity during the previous week.

Table 4 shows the mean of the values of accuracy, sensitivity, specificity, positive predictive value and negative predictive value in the different evaluations performed in the cross validation.

Table 1
Variables associated with PsAID ≥4. Bivariate analysis.

Variable	PsAID <4	PSAID ≥4	*p*-value
Sex (women)	115 (39.9)	74 (55.2)	0.005
Educational level			0.004
None	2 (0.7)	5 (3.7)	
Primary	96 (33.3)	61 (45.5)	
Secondary	124 (43.1)	50 (37.3)	
University	66 (22.9)	18 (13.4)	
Level of physical activity in the previous week			<0.001
Low	38 (16.0)	42 (35.3)	
Medium	124 (52.1)	51 (42.9)	
High	76 (31.9)	26 (21.8)	
Smoking			0.04
Never smoked	115 (42.6)	45 (33.6)	
Exsmoker	73 (27.0)	46 (34.3)	
Occasional smoker	12 (4.4)	1 (0.7)	
Daily smoker	70 (25.9)	42 (31.3)	
Weekly alcohol consumption	0 [0–4]	0 [0–2]	0.02
Body mass index	26.40 [24.22–29.05]	28.36 [24.89–31.94]	0.03
Clinical form at diagnosis			0.02
Axial	25 (8.7)	5 (3.7)	
Peripheral	235 (81.6)	105 (78.4)	
Mixed	28 (9.7)	24 (17.9)	
Joint pattern at diagnosis			0.004
Oligoarticular	152 (52.8)	81 (60.4)	
Polyarticular	80 (27.8)	45 (33.6)	
Distal	23 (8.0)	2 (1.5)	
Spondylitis	33 (11.5)	6 (4.5)	
High C-reactive protein	68 (26.0)	50 (37.9)	0.02
Polyarthritis	11 (4.0)	36 (26.9)	<0.001
Diabetes mellitus			<0.001
No	253 (92.3)	113 (84.3)	
Non–insulin-dependent	21 (7.7)	14 (10.5)	
Insulin-dependent	0	7 (5.2)	
Enthesitis	52 (19.1)	44 (32.8)	0.003
NSJ66	0 [0–1.5]	2 [0–5]	<0.001
NPJ68	1 [0–2]	4 [2–10]	<0.001
Psoriasis affecting the scalp	87 (35.7)	71 (53.8)	0.001
PASI >10	1 (0.4)	7 (5.7)	0.003
Pain in the previous week	2 [1–4]	7 [5–8]	<0.001
Patient global assessment of disease	2 [1–4]	7 [5–8]	<0.001
Physician global assessment of disease	2 [1–3]	4 [3–6]	<0.001
HAQ	0.13 [0–0.38]	1.13 [0.75–1.5]	<0.001

Categorical variables are expressed as n (%). Numerical variables are expressed as median [IQR] if not.

Tabla 2
Variables associated with PsAID ≥4: Logistic regression analysis taking physical activity as a categorical variable

Variable	Regression coefficient*	95% CI	*p* value (Wald test)
HAQ	10.394	(7.777, 13.011)	<0.001
Pain in the previous week	5.668	(4.016, 7.320)	<0.001
Educational level	-2.064	(-3.515, -0.613)	0.005
Low level of physical activity in the previous week (reference)			
Moderate level of physical activity in the previous week	-0.341	(-1.255, 0.573)	0.465
High level of physical activity in the previous week	1.221	(0.158, 2.283)	0.024

* Positive values indicate that the higher the value of the variable, the higher the frequency of PsAID ≥4.

Table 3
Feature importances* of the variables in the different models trained in the cross validation.

Variable	Iteration 1	Iteration 2	Iteration 3	Iteration 4	Iteration 5
Random Forest					
HAQ	0.479	0.513	0.521	0.383	0.537
Pain in the previous week	0.364	0.334	0.330	0.461	0.325
Educational level	0.080	0.081	0.075	0.087	0.077
Level of physical activity	0.078	0.072	0.073	0.069	0.062
XGBoost					
HAQ	0.558	0.482	0.415	0.501	0.568
Pain in the previous week	0.201	0.342	0.442	0.256	0.175
Educational level	0.117	0.094	0.058	0.143	0.132
Level of physical activity	0.124	0.083	0.085	0.100	0.125

* Values from 0 to 1. The higher the value, the greater the importance of the variable in the model. Values are normalized, i.e. in each iteration the sum of the values equals 1.

Table 4
Pooled measures of validity in the different evaluations performed in the cross validation.

Metric	Accuracy	Sensitivity	Specificity	NPV	PPV
Random Forest					
Mean*	90.1	93.3	86.5	92.5	88.5
SD	2.1	4.1	4.4	4.5	3.4
95% CI	84.2, 96.0	82.0, 100	74.2, 98.9	80.0, 100	79.0, 98.0
XGBoost					
Mean&	87.6	90.0	85.0	88.8	86.8
SD	2.5	4.1	4.1	3.9	3.0
95% CI	80.6, 94.5	78.6, 100	73.6, 96.4	77.9, 99.8	78.4, 95.3

SD: standard deviation.
* Mean of the values obtained in the 5 evaluations performed in the cross validation in Random Forest analysis.
& Mean of the values obtained in the 5 evaluations performed in the cross validation in XGBoost analysis.

Discussion

In this multicenter prospective study carried out in patients with recent-onset PsA, assessed at baseline before the potential modification of its natural history because of the treatment prescribed by a rheumatologist, an artificial intelligence–based analysis revealed 4 variables to be associated with high-impact disease according to PsAID. Most of the models stratified these variables from greater to lesser importance as follows: HAQ score, pain in the previous week, educational level, and level of physical activity during the previous week. Higher HAQ, higher pain and lower educational level were associated with a higher frequency of PsAID ≥4. It should be noted that the frecuency of PsAID ≥4 was lower in the group of patients with moderate level of physical activity than in the group with low level (although this association was not statistically significant), whereas it was higher if the level of physical activity was high. The mean values of the measures of validity of the machine learning algorithms were all ≥85%.

HAQ-DI is considered a coreinstrument for assessing physical function in numerous diseases, including PsA [21]. In fact, it is a component of minimal disease activity (MDA), one of the treatment objectives promoted by EULAR as part of their treat-to-target approach in PsA [22].

Moreover, the PsAID itself contains an item aimed at evaluating physical function [3]. Curiously, when the MDA response was analyzed according to whether the patient had a low-impact PsAID score, no differences in inflammatory burden represented by the swollen joint count were seen, whereas the HAQ scores were significantly higher in patients with high-impact PsAID [23]. Agreement between a low-impact PsAID score and preserved physical function (HAQ < 0.5) is quite acceptable [23]. In fact, this agreement is much closer between PsAID and HAQ than between PsAID and MDA or remission measured using DAPSA [24]. A recent study showed that the most severely impaired functions included in the HAQ in patients with PsAID >4 were those depending on upper limb joints [25]. In summary, impaired physical function measured using HAQ is associated with a greater impact on quality of life as measured using PsAID.

Pain was the second most important variable in most of the models. Pain is the item that receives the highest corrective weighting when calculating PsAID score (3-fold) and is usually the variable patients consider most relevant with respect to their quality of life [3,26]. In addition, we must not forget that pain is closely correlated with other important aspects such as fatigue, sleep alterations, poorer physical function, and psychological distress, all of which are included in the PsAID questionnaire [3,27,28].

We are aware that the association of pain and the HAQ with the PsAID is due to the fact that both are aspects directly assessed in the PsAID itself, and that this could be considered circularity and subject to criticism. However, one remarkable finding of our analysis would be that the association of HAQ with PsAID was higher than that of pain, despite the fact that when calculating PsAID score, the item about pain is multiplied x3, while the item about physical function is multiplied x2. This was consistently observed in the cross validation.

The third variable in order of importance was the patient's educational level. Associations between educational level and disease impact and activity have been assessed in several studies. The direction of this association is generally the same (the lower the educational level, the poorer the outcome) [29]. According to the results of our analysis, this association would be independent of possible differences in pain or physical function between subjects with different educational level.

The last variable identified as being associated with high-impact disease according to PsAID was the level of physical activity. Curiously, moderate physical activity could be associated with PsAID <4, whereas a high level of physical activity was associated with PsAID ≥4. When interpreting this result, it must be taken into account that one of the items in PsAID questionnaire refers to the level of fatigue due to the psoriatic arthritis. In our sample, we observed that a high level of physical activity was associated with scores ≥4 in this item after adjusting by pain, HAQ and educational level. A possible explanation for this association could be that subjects with a high level of physical activity score this item higher, not because of the disease, but as a result of the physical activity. A practical corollary of this last PsAID-associated variable, would be that patients could need help when answering this item of the questionnaire in order to avoid misinterpretation.

The main limitation of this study is its sample size and the fact that some data are missing for some variables. This affected the power of the statistical analysis and, therefore, the ability of the study to detect variables associated with the outcome measure. We tried to compensate for this by using models based on artificial intelligence and machine learning. Random forests are "joint" algorithms in which decision trees are trained with different subsets of variables and data. Decision trees are more flexible than many statistical models, since they make it possible to identify many types of association between explanatory variables and the outcome measure. Furthermore, the fact that random forests add variability prevents the model from being overadjusted to the data and can be re-run with new data, thus increasing the robustness of the predictions. On the other hand, XGBoost algorithms use ensembles of decision trees in a sequential manner. In each tree, the observations that were wrongly-classified in the previous one are given a larger weight, thus creating models with very little bias which usually result in very accurate predictions. The counterpart of this phenomenon is a higher risk of the model being overfit to the training dataset. Our analysis showed that the random forest models tended to perform better than XGBoost in terms of all the metrics, which is probably due to the reduced number of observations in the dataset causing the training and test subsets to be quite disparate. Therefore, we could conclude that for such small datasets, an algorithm that overfits less to the training subset such as random forest is more appropriate.

We observed a significant decrease in the percentage of patients with a high-impact PsAID from the baseline visit (more than 40%) to the last visit (around 20%). Of course, some of the variables associated with PsAID ≥4 could vary between the baseline and follow-up visits, but we think that this would mainly affect the drugs received and their effect on PsAID would be reflected by their effect on other variables considered at all visits (such as pain or the HAQ). On the other hand, the reduction in sample size that would entail analyzing baseline and follow-up visits separately would imply a greater probability of statistical errors in the analysis.

The main strength of this study is its ability to record the course of PsA from an early phase before the natural disease evolution is modified by treatment prescribed by the rheumatologist.

Conclusions

Our artificial intelligence–based models showed, with high values of the measures of validity, that higher HAQ, pain, low educational level, and a high level of physical activity are associated with a high-impact PsAID score and, therefore, perception of poorer quality of life among patients with PsA. For this reason, both pain control and control of the disease as a whole, preventing patients from suffering a decrease in their physical function, are first-order treatment objectives. Despite the higher weight given to pain when scoring PsAID, we oberved a greater influence of physical function on disease impact.

CRediT authorship contribution statement

Rubén Queiro: Visualization, Funding acquisition, Data curation, Writing – original draft. **Daniel Seoane-Mato:** Visualization, Formal analysis, Data curation, Writing – original draft. **Ana Laiz:** Validation, Funding acquisition. **Eva Galíndez Agirregoikoa:** Validation, Funding acquisition. **Carlos Montilla:** Validation, Funding acquisition. **Hye-Sang Park:** Validation, Funding acquisition. **Jose A. Pinto-Tasende:** Validation, Funding acquisition. **Juan J. Bethencourt Baute:** Validation, Funding acquisition. **Beatriz Joven Ibáñez:** Validation, Funding acquisition. **Elide Toniolo:** Validation, Funding acquisition. **Julio Ramírez:** Validation, Funding acquisition. **Ana Serrano García:** Formal analysis.

Declaration of Competing Interest

None.

Funding

REAPSER study was supported by AbbVie, that had no role in the design, data collection, data analysis, interpretation, writing or the decision to submit the paper for publication of this manuscript.

Acknowledgments

The authors would like to acknowledge José Luis Fernández Sueiro for the conception of the study; José Miguel Carrasco for his contribution to the design of the study; Nuria Montero and Cristina Oliva for their contribution to data monitoring; Ana González Marcos and Cristina Pruenza for their contribution to data analysis; and Thomas OBoyle for

the translation of the manuscript.

Non-author members of REAPSER Study Group: Juan D. Cañete (Hospital Clinic i Provincial, Barcelona, Spain); Xavier Juanola (Hospital Universitari de Bellvitge, Barcelona, Spain); Jordi Fiter (Hospital Universitario Son Espases, Palma de Mallorca, Spain); Jordi Gratacós (Hospital Universitari Parc Taulí, Barcelona, Spain); Jesús Rodriguez-Moreno, Jaime Notario Rosa (Hospital Universitari de Bellvitge, Barcelona, Spain); Andrés Lorenzo Martín, Anahy Brandy García, Pablo Coto Segura (Hospital Universitario Central de Asturias, Oviedo, Spain); Anna López Ferrer (Hospital de la Santa Creu i Sant Pau, Barcelona, Spain); Silvia Pérez Barrio (Hospital Universitario Basurto, Bilbao, Spain); Andrés J Plata Izquierdo (Hospital Clínico Universitario de Salamanca, Salamanca, Spain); Sagrario Bustabad, Francisco J Guimerá Martín-Neda (Hospital Universitario de Canarias, San Cristóbal de La Laguna, Spain); Eduardo Fonseca Capdevilla (Complexo Hospitalario Universitario A Coruña, A Coruña, Spain); Raquel Rivera Díaz (Hospital Universitario 12 Octubre, Madrid, Spain); Andrea Cuervo, Mercè Alsina Gibert (Hospital Clinic i Provincial, Barcelona, Spain); Pilar Trenor Larraz, Isabel de la Morena Barrio, Laura Puchades Lanza, Diego Bedoya Sanchís (Hospital Clínico Universitario de Valencia, Valencia, Spain); Catalina Meliá Mesquida, Claudia Murillo (Hospital Son Llàtzer, Palma de Mallorca, Spain); Manuel J Moreno Ramos, María D Beteta, Paloma Sánchez-Pedreño Guillén (Hospital Clínico Universitario Virgen de la Arrixaca, Murcia, Spain); Leticia Lojo Oliveira, Teresa Navío Marco, Laura Cebrián, Pablo de la Cueva Dobao (Hospital Universitario Infanta Leonor, Madrid, Spain); Martina Steiner, Santiago Muñoz-Fernández, Ricardo Valverde Garrido (Hospital Universitario Infanta Sofía, Madrid, Spain); Manuel León, Esteban Rubio, Alejandro Muñoz Jiménez, Lourdes Rodríguez Fernández-Freire (Hospital Universitario Virgen del Rocío, Sevilla, Spain); Julio Medina Luezas, María D Sánchez-González, Carolina Sanz Muñoz (Hospital Clínico Universitario de Valladolid, Valladolid, Spain); José M Senabre, José C Rosas, Gregorio Santos Soler, Francisco J Mataix Díaz (Hospital Marina Baixa, Villajoyosa, Spain); Juan C Nieto-González, Carlos González, Juan G Ovalles Bonilla, Ofelia Baniandrés Rodríguez (Hospital Universitario Gregorio Marañón, Madrid, Spain); Fco Javier Nóvoa Medina, Dunia Luján (Complejo Hospitalario Universitario Insular Materno-Infantil, Las Palmas de Gran Canaria, Spain); María D Ruiz Montesino, Ana M Carrizosa Esquivel (Hospital Universitario Virgen Macarena, Sevilla, Spain); Cristina Fernández-Carballido, María P Martínez-Vidal, Laura García Fernández (Hospital General Universitario de Elda, Elda, Spain); Vega Jovani, Rocío Caño Alameda, Silvia Gómez Sabater, Isabel Belinchón Romero (Hospital General Universitario de Alicante, Alicante, Spain); Ana Urruticoechea-Arana, Marta Serra Torres (Hospital Can Misses, Ibiza, Spain); Raquel Almodóvar, José L López Estebaranz (Hospital Universitario Fundación Alcorcón, Alcorcón, Spain); María D López Montilla, Antonio Vélez García-Nieto (Hospital Universitario Reina Sofía, Córdoba, Spain).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.semarthrit.2022.152097.

References

[1] Romero Pérez A, Queiro R, Seoane-Mato D, Graell E, Chamizo E, Chaves Chaparro L, et al. Higher prevalence of psoriatic arthritis in the adult population in Spain? A population-based cross-sectional study. PLoS One 2020;15:e0234556.

[2] Ritchlin CT, Colbert RA, Gladman DD. Psoriatic arthritis. N Engl J Med 2017;376:957–70.

[3] Gossec L, De Wit M, Kiltz U, Braun J, Kalyoncu U, Scrivo R, et al. A patient-derived and patient-reported outcome measure for assessing psoriatic arthritis: elaboration and preliminary validation of the Psoriatic Arthritis Impact of Disease (PsAID) questionnaire, a 13-country EULAR initiative. Ann Rheum Dis 2014;73:1012–9.

[4] Orbai A, Holland R, Leung YY, Tillett W, Goel N, McHugh N, et al. PsAID12 provisionally endorsed at OMERACT 2018 as core outcome measure to assess psoriatic arthritis-specific health- related quality of life in clinical trials. J Rheumatol 2019;46:990–5.

[5] Pardo E, Charca L, Alonso S, Alperi M, Arboleya L, Queiro R. Disease activity in psoriatic arthritis index and psoriatic arthritis impact of disease questionnaire: correlation and sensitivity to change in a real clinical setting. Clin Exp Rheumatol 2020;38:973–7.

[6] Queiro R, Cañete JD, Montoro M, Gómez S, Cábez A, MAAPS study group. Disease features associated with a low disease impact in patients with psoriatic arthritis: results of a cross-sectional multicenter study. Arthritis Res Ther 2020;22:82.

[7] Orbai AM, Perin J, Gorlier C, Coates LC, Kiltz U, Leung YY, et al. Determinants of patient-reported psoriatic arthritis impact of disease: an analysis of the association with gender in 458 patients from 14 countries. Arthritis Care Res (Hoboken) 2020;72:1772–9.

[8] Queiro R, Laiz A, Seoane-Mato D, Galindez Agirregoikoa E, Montilla C, Park HS, et al. Spanish registry of recent-onset psoriatic arthritis (REAPSER study): aims and methodology. Reumatol Clin (Engl Ed) 2019;15:252–7.

[9] Taylor W, Gladman D, Helliwell P, Marchesoni A, Mease P, Mielants H, et al. Classification criteria for psoriatic arthritis: development of new criteria from a large international study. Arthritis Rheum 2006;54:2665–73.

[10] Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J. Chronic Dis 1987;40:373–83.

[11] Gual A, Contel M, Segura L, Ribas A, Colom J. The ISCA (Systematic Interview of Alcohol Consumption), a new instrument to detect risky drinking. Med Clin (Barc) 2001;117:685–9.

[12] The IPAQ Group Guidelines for data processing and analysis of the International Physical Activity Questionnaire (IPAQ)—short and long forms. http://www.ipaq.ki.se/. Accessed 9 Oct 2021.

[13] Heuft-Dorenbosch L, Spoorenberg A, van Tubergen A, Landewe R, van ver Tempel H, Mielants H, et al. Assessment of enthesitis in ankylosing spondylitis. Ann Rheum Dis 2003;62:127–32.

[14] Feldman SR, Krueger GG. Psoriasis assessment tools in clinical trials. Ann Rheum Dis 2005;64(2):ii65–8. Suppldiscussion ii9-73.

[15] Esteve-Vives J, Batlle-Gualda E, Reig A. Spanish version of the Health Assessment Questionnaire: reliability, validity and transcultural equivalency. Grupo para la Adaptacion del HAQ a la Poblacion Espanola. J Rheumatol 1993;20:2116–22.

[16] MacKay K, Mack C, Brophy S, Calin A. The Bath Ankylosing Spondylitis Radiology Index (BASRI): a new, validated approach to disease assessment. Arthritis Rheum 1998;41:2263–70.

[17] Rahman P, Gladman DD, Cook RJ, Zhou Y, Young G, Salonen D. Radiological assessment in psoriatic arthritis. Br J Rheumatol 1998;37:760–5.

[18] Batista G, Prati R, Monard MC. A study of the behavior of several methods for balancing machine learning training data. SIGKDD Explor 2004;6:20–9.

[19] Queiro R, Seoane-Mato D, Laiz A, Galindez Agirregoikoa E, Montilla C, Park HS, et al. Severe disease in patients with recent-onset psoriatic arthritis. prediction model based on machine learning. Front Med (Lausanne) 2022;9:891863.

[20] Vidaurre D, Bielza C, Larranaga P. A survey of L1 regression. Int Stat. Rev. 2013;81:361–87.

[21] Højgaard P, Klokker L, Orbai AM, Holmsted K, Bartels EM, Leung YY, et al. A systematic review of measurement properties of patient reported outcome measures in psoriatic arthritis: a GRAPPA-OMERACT initiative. Semin Arthritis Rheum 2018;47:654–65.

[22] Smolen JS, Schöls M, Braun J, Dougados M, FitzGerald O, Gladman DD, et al. Treating axial spondyloarthritis and peripheral spondyloarthritis, especially psoriatic arthritis, to target: 2017 update of recommendations by an international task force. Ann Rheum Dis 2018;77:3–17.

[23] Queiro R, Pardo E, Charca L. Distribution of the components of the MDA response among patients with psoriatic arthritis with and without an acceptable symptomatic state. Clin Exp Rheumatol 2020;38:575–6.

[24] Queiro R, Cañete JD, Montilla C, Abad MA, Montoro M, Gómez S, et al. Very low disease activity, DAPSA remission, and impact of disease in a Spanish Population with psoriatic arthritis. J Rheumatol 2019;46:710–1.

[25] Queiro R. Analysis of the HAQ-DI components in psoriatic arthritis patients with and without low disease impact. Rheumatology (Oxford) 2020;59:3569–70.

[26] Leung YY, Orbai AM, Hojgaard P, Holland R, Mathew AJ, Goel N, et al. OMERACT Filter 2.1 instrument selection for physical function domain in psoriatic arthritis: provisional endorsement for HAQ-DI and SF-36 PF. Semin Arthritis Rheum 2021;51:1117–24.

[27] Garrido-Cumbrera M, Hillmann O, Mahapatra R, Trigos D, Zajc P, Weiss L, et al. Improving the management of psoriatic arthritis and axial spondyloarthritis: roundtable discussions with healthcare professionals and patients. Rheumatol Ther 2017;4:219–31.

[28] Desthieux C, Granger B, Balanescu AR, Balint P, Braun J, Canete JD, et al. Determinants of patient-physician discordance in global assessment in psoriatic arthritis: a multicenter European study. Arthritis Care Res 2017;69:1606–11.

[29] Sumpton D, Kelly A, Tunnicliffe DJ, Craig JC, Hassett G, Chessman D, et al. Patients' perspectives and experience of psoriasis and psoriatic arthritis: a systematic review and thematic synthesis of qualitative studies. Arthritis Care Res (Hoboken) 2020;72:711–22.