Combining patient-reported outcome measures to screen for active disease in rheumatoid arthritis and psoriatic arthritis

Published in:
RMD Open

Publication status and date:
Published: 18/10/2024

DOI (link to publisher):
10.1136/rmdopen-2024-004687

Document Version
Publisher's PDF, also known as Version of record

RMD Open
Rheumatic & Musculoskeletal Diseases

ORIGINAL RESEARCH

Combining patient-reported outcome measures to screen for active disease in rheumatoid arthritis and psoriatic arthritis

Agnes E M Looijen ,[1] Selinde V J Snoeck Henkemans ,[1] Annette H M van der Helm-van Mil ,[1,2] Paco M J Welsing,[3] Gonul Hazal Koc ,[1] Jolanda J Luime ,[1] Marc R Kok ,[4] Ilja Tchetverikov,[5] Lindy-Anne Korswagen,[6] Paul Baudoin,[7] Marijn Vis,[1] Pascal H P de Jong [1]

To cite: Looijen AEM, Snoeck Henkemans SVJ, van der Helm-van Mil AHM, *et al*. Combining patient-reported outcome measures to screen for active disease in rheumatoid arthritis and psoriatic arthritis. *RMD Open* 2024;**10**:e004687. doi:10.1136/rmdopen-2024-004687

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/rmdopen-2024-004687).

AEML and SVJSH contributed equally.

Received 27 June 2024
Accepted 29 September 2024

For numbered affiliations see end of article.

Correspondence to
Selinde V J Snoeck Henkemans; s.snoeckhenkemans@erasmusmc.nl

ABSTRACT

Objectives To investigate whether a combination of general health (Visual Analogue Scale (VAS)), Health Assessment Questionnaire-Disability Index (HAQ-DI), pain (VAS/Numerical Rating Scale (NRS)), quality of life (EQ-5D), fatigue (VAS/NRS) and presenteeism (0%–100% productivity loss) could aid as a screening tool to detect active disease in patients with rheumatoid arthritis (RA) and psoriatic arthritis (PsA).

Methods RA patients from the tREACH trial and TARA trial (n=683) and PsA patients from the DEPAR cohort (n=525) were included. The association of a deterioration in the aforementioned patient-reported outcome measure (PROM) scores between two consecutive visits and having active disease was assessed. Active disease was defined as a change from disease activity score (DAS) ≤2.4 to DAS >2.4 in RA or Disease Activity Index in Psoriatic Arthritis (DAPSA) ≤14 to DAPSA >14 in PsA. The area under the curve (AUC) of the sum score of deteriorated PROMs was evaluated.

Results 4594 RA and 1154 PsA visits were evaluated and active disease occurred in 358 (8%) RA and 177 (15%) PsA visits. In both RA and PsA, a deterioration in general health (VAS), HAQ-DI, EQ-5D and pain (VAS/NRS) was significantly associated with active disease. The combination of these PROMs showed acceptable to excellent discriminative ability (RA AUC=0.76, PsA AUC=0.85). If a cut-point of ≥1 deteriorated PROMs is used, 40% of the visits in which RA patients have remission or low disease activity are correctly specified (specificity of 40%), while 10% of visits with active disease are overlooked (sensitivity of 90%). In PsA, these percentages are 41% and 4%, respectively.

Conclusion A combination of general health, HAQ-DI, EQ-5D and pain could aid as a screening tool for active disease in patients with RA and PsA. These data could help facilitate remote monitoring of RA and PsA patients in the future.

Trial registration numbers ISRCTN26791028, NTR2754.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ To improve healthcare efficiency for patients with inflammatory arthritis while maintaining a high quality of care, remote monitoring using patient-reported outcome measures (PROMs) may add value. However, one PROM is not responsive enough to detect changes in disease activity status.

WHAT THIS STUDY ADDS

⇒ A combination of Visual Analogue Scale (VAS) general health/patient global assessment, Health Assessment Questionnaire-Disability Index, EQ-5D and VAS pain, was able to recognise patients with active rheumatoid arthritis (RA) and psoriatic arthritis (PsA) with acceptable diagnostic accuracy.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The combination of aforementioned PROMs could be used as a screening tool for active disease in RA and PsA patients.

INTRODUCTION

Current management guidelines for rheumatoid arthritis (RA) and psoriatic arthritis (PsA) recommend a treat-to-target (T2T) approach.[1 2] In a T2T approach, treatment is adjusted if the target of low disease activity (LDA) is not achieved.[1–4] Within this approach, disease activity is regularly assessed with a disease activity index for which a physical examination is essential. However, with the increasing demands on our healthcare system, due to the ageing population and the impending shortage of healthcare providers, delivery of care should become more efficient.[5 6] One solution could be to screen for patients who need to be assessed in the outpatient clinic through remote monitoring. Patient-reported outcome measures (PROMs) might be of added value for this purpose.[7] An advantage of using PROMs for remote monitoring is the ability to capture

them with digital applications that can automatically calculate PROM scores.[8] A challenge, on the other hand, is that patients' and physicians' perspectives on disease activity do not always align.[9]

PROMs give healthcare providers insight into the impact of the disease on patients' lives, which is one of the reasons why they are increasingly being implemented in routine clinical care. The International Consortium for Health Outcomes Measurement (ICHOM) has recommended regular measurement of disease activity, using a disease activity index (eg, Disease Activity Score (DAS) or Disease Activity Index in Psoriatic Arthritis (DAPSA)), as well as PROMs covering the following outcome domains: activity limitations, health impact, pain, fatigue and work ability. If these PROMs would correlate well with disease activity as measured by the rheumatologist, they could be used for remote monitoring in addition to providing insight into the impact of the disease on patients' lives.[10] This could ultimately lead to a reduction of unnecessary visits, as well as the detection of patients who need to be assessed earlier than their regular scheduled visit for a possible disease flare.

Remote monitoring with PROMs has already been shown to be promising in patients with inflammatory arthritis (IA). A systematic review of the clinical impact of using electronic PROMs to remotely monitor patients with IA has suggested that the use of remote monitoring in addition to outpatient follow-up is not inferior to regular care.[11] However, most of the included studies offered other interventions in combination with the use of electronic PROMs, such as education on the disease, limiting conclusions about the equivalence of remote monitoring with PROMs. Moreover, the majority of the included studies involved RA patients.[11] To the best of our knowledge, no studies on remote monitoring with PROMs have been conducted in PsA patients.

We have previously shown that on a group level general health/patient global assessment (PGA), functional ability (Health Assessment Questionnaire-Disability Index (HAQ-DI)), quality of life (EQ-5D) and pain are responsive to improvement or worsening in disease activity status in early and established RA patients.[12] However, the responsiveness of one single PROM to flare development in individual patients was poor, suggesting that one PROM cannot adequately differentiate between active disease and LDA or remission. However, a combination of PROMs might improve the discriminatory ability on a patient level. Thus far, one study in RA has found an association between a combination of PROMs, that is, the Visual Analogue Scale (VAS) general health, VAS disease activity and VAS pain, and disease activity status according to the rheumatologist.[13] Another study has shown that the combination of a low score on the Routine Assessment of Patient Index Data 3 (RAPID3) and a flare question, that is, 'Are you having a flare of your RA at this time?', were likely to rule out patients in need of treatment intensifications.[14] Although these studies showed promising results, not all ICHOM-recommended outcome domains were taken into account. The inclusion of more outcome domains may improve test characteristics. Furthermore, the selected PROMs should ideally be suitable for all IA patients.

We hypothesise that a combination of PROMs, that show responsiveness on a population level, will have a promising discriminative ability. Therefore, our aim is to examine whether a combination of PROMs from the ICHOM-recommended domains can adequately discriminate between a well-controlled (DAS ≤2.4 or DAPSA ≤14) and active disease (DAS >2.4 or DAPSA >14).[15 16]

METHODS

Patients and study data

For this study, we included patients who participated in the 'treatment in the Rotterdam Early Arthritis Cohort' trial (tREACH, ISRCTN26791028), the 'Tapering strategies in Rheumatoid Arthritis' trial (TARA, NTR2754) or the 'Dutch southwest Early Psoriatic Arthritis cohoRt' (DEPAR).

The tREACH trial was a stratified, single-blinded randomised controlled trial with a follow-up period of 5 years. The trial was carried out in eight rheumatology centres in the southwestern part of the Netherlands. Patient recruitment took place between July 2007 and April 2011. Disease-modifying anti-rheumatic drug (DMARD)-naïve early undifferentiated arthritis (UA) and RA patients with an arthritis in ≥1 joint and symptom duration of <1 year were included. RA diagnosis was based on fulfilment of the 1987 or 2010 classification criteria.[17 18] In the tREACH trial, multiple initial treatment strategies were compared. Patients received either (1) methotrexate, including DMARD combination therapies with or without glucocorticoid bridging therapy; (2) hydroxychloroquine or (3) non-steroid anti-inflammatory drugs/glucocorticoids as initial treatment. The tREACH trial had a T2T management approach that was aimed at reaching LDA, defined as a DAS with 44 joints ≤2.4. Treatment was intensified until LDA was achieved. Medication was tapered if patients were in sustained remission, defined as DAS ≤1.6 at two consecutive visits. If a flare occurred, defined as a DAS >2.4, full treatment was restarted according to the stage of the protocol. The full treatment protocol is described elsewhere.[19]

The TARA trial was a single-blinded randomised controlled trial with a follow-up period of 2 years. The trial was carried out in 12 rheumatology centres in the southwestern part of the Netherlands. The patient recruitment took place between September 2011 and July 2016. RA patients with well-controlled disease, defined as a DAS ≤2.4 and swollen joint count (SJC) ≤1 at two consecutive visits within 3 months, who used both ≥1 conventional synthetic (cs) DMARDs and a TNF-inhibitor, were included. RA diagnosis was based on fulfilment of the 1987 or 2010 classification criteria.[17 18] The median symptom duration (SD) of the included RA patients was 6.2 (4.1–8.9) years and 98% of the patients

had a disease duration ≥1 year, which is defined as established RA.[20] The TARA trial compared the effectiveness of different tapering strategies. Patients were randomised into tapering their csDMARD in the first year followed by tapering their TNF inhibitor in the second year or vice versa. If a disease flare occurred, defined as DAS >2.4 or SJC >1, the last effective treatment was restarted and intensified until well-controlled disease was re-established. Extensive information on the treatment protocol can be found elsewhere.[21]

The DEPAR is an ongoing real-world, prospective observational cohort that started in 2013 and includes DMARD-naïve PsA patients. This study is carried out in 10 rheumatology centres in the Netherlands. PsA diagnosis is based on the expert opinion of the treating rheumatologist, thus no classification criteria were applied. Treatment decisions are based on the insight of the treating rheumatologist and in shared decision with the patient, which is in line with current (inter)national guidelines. DEPAR aims to gain more insight into the disease course and most effective treatment of PsA patients. For this analysis, data up to March 2023 were used. Included patients are followed for 3 years, which could be extended up to 10 years. Further details on the DEPAR study are described in detail elsewhere.[22]

Patients from aforementioned studies were included if they had their disease activity measured at two consecutive visits and had filled out ≥1 PROM score of interest at those two consecutive visits.

Data collection

For these analyses, we used pre-existing data that were collected every 3 months in each aforementioned study. During each visit, disease activity measures were taken by trained research nurses. In addition, a variety of PROMs were collected. For the early UA/RA patients, these 3 monthly visits took place during a follow-up period of 3 years, for established RA patients during 2 years of follow-up and for PsA patients during the first year of the follow-up period.

In all studies, patients completed questionnaires, including the PROMs selected from each ICHOM-recommended domain. These PROMs are general health/PGA (VAS, 0–100 mm), HAQ-DI, EQ-5D, pain (VAS, 0–100 mm/Numeric Rating Scale (NRS), 0–10), fatigue (VAS, 0–100 mm/NRS, 0–10) and presenteeism (0%–100% productivity loss). Table 1 shows the collected PROMs per outcome domain and patient group.[23]

Disease activity and clinical outcomes

Disease activity is usually objectified by a physician in the outpatient clinic using a disease activity index. The composite disease activity measures were chosen because current T2T guidelines recommend the use of the DAS for RA and the DAPSA for PsA to assess disease activity.[1 2]

In both early UA/RA and established RA, disease activity was assessed by the DAS with a 44-SJC, tender joint count (TJC) measured with the Ritchie Articular Index in 53 joints, erythrocyte sedimentation rate (in millimetres (mm)/hour) and general health (VAS, 0–100 mm) or PGA (VAS, 0–100 mm).[24] In addition, anti-citrullinated protein antibody (ACPA) and rheumatoid factor (RF) status were determined as well as the C reactive protein (CRP, in milligrams/decilitres) for the RA patients.

In PsA, disease activity was measured with the DAPSA score, which is the sum of the 66-SJC, 68-TJC, CRP (in milligrams/decilitres), general health (VAS, 0–100 mm divided by 10) and pain (VAS, 0–100 mm divided by 10).[25 26] We chose the DAPSA as this is a continuous disease activity measure comparable to the DAS. In addition, the number of patients with arthritis, psoriasis and enthesitis was determined together with the body surface area (BSA, 0%–100%) and the Leeds Enthesitis Index (LEI, 0–6).

Table 1 Included PROMs per ICHOM-recommended outcome domain and patient group

	Patient groups		
Outcome domains	**Early UA/RA**	**Established RA**	**PsA**
Disease activity component	General health*	PGA*	General health*
Activity limitations	HAQ-DI	HAQ-DI	HAQ-DI
Health impact	EQ-5D-3L	EQ-5D-3L	EQ-5D-5L
Pain	NRS, 0–10	NRS, 0–10	VAS, 0–100 mm
Fatigue	VAS, 0–100 mm	NRS, 0–10	NRS, 0–10
Work ability	Presenteeism†	Presenteeism†	Presenteeism†

*Measured with a Visual Analogue Scale from 0 to 100 mm.
†Measured using the percentage of productivity loss due to working while sick (0%–100%).
EQ-5D-3L, 3-level EQ-5D; EQ-5D-5L, 5-level EQ-5D; HAQ-DI, Health Assessment Questionnaire-Disability Index; ICHOM, International Consortium for Health Outcomes Measurement ; NRS, Numeric Rating Scale; PGA, patient global assessment of disease activity; PROM, patient-reported outcome measures; PsA, psoriatic arthritis; RA, rheumatoid arthritis; UA, undifferentiated arthritis; VAS, Visual Analogue Scale.

Component of disease activity

General health (in early UA/RA and PsA) and PGA (in established RA) were included as a PROM to capture the inflammatory disease activity component of the ICHOM-recommended outcome domain. Both general health and PGA were measured on a 0–100 mm VAS. Higher scores indicate a poorer perceived general health.[27]

Activity limitations and health impact

Activity limitation was measured with the HAQ-DI in all three patient groups. The total score ranges from 0 to 3 and higher scores indicate more functional impairment.[28–30]

In both early UA/RA and established RA, health impact was measured with the 3-level EQ-5D (EQ-5D-3L).[31–33] The EQ-5D-3L measures five health domains on a 3-point Likert scale. In the PsA cohort, the health impact was measured with the 5-level EQ-5D (EQ-5D-5L), which measures five health domains on a 5-point Likert scale. From these answers, utility scores were calculated based on Dutch reference values.[34] Scores range from below 0 to 1; where 0 equals a health status similar to death and 1 equals perfect health.

Pain and fatigue

In early UA/RA and established RA, joint pain was measured with an NRS ranging from 0 to 10. In the PsA cohort, pain was measured on a 0–100 mm VAS. The NRS was transformed to a 0–100 scale to make both scores comparable. Higher scores indicate more experienced pain.

Fatigue was measured on a 0–100 mm VAS in early UA/RA. In established RA and PsA, an NRS from 0 to 10 was used. Again, this NRS was transformed to a 0–100 scale. Higher scores represent more severe fatigue.[35]

Work ability

For this study, we included presenteeism as a measure of productivity loss. Presenteeism is defined as loss of productivity due to working while sick. Presenteeism was assessed with an NRS ranging from 0 to 10.[36 37] This score was converted into a percentage productivity loss, where 0% indicates no productivity loss and 100% indicates the inability to work. The score is only available for patients who had paid work at the time of filling out the questionnaire. Since not all patients have paid work, we assessed the additional value of this outcome measure in a sensitivity analysis.

Statistical analysis

The following PROMs were evaluated for selection in the main prediction model: general health/PGA, HAQ-DI, EQ-5D, pain and fatigue. The change in PROM scores between two consecutive visits was categorised into two groups: 'deterioration' (1) and 'no deterioration' (0). For example, if the HAQ-DI score increased over a 3-month interval, indicating increased functional impairment, it was labelled as 'deterioration'. If the HAQ-DI score remained the same or decreased, it was categorised as 'no deterioration'. For the EQ-5D and HAQ-DI, any worsening in score was considered a deterioration. For the PROMs that use a (transformed) 0–100 scale, a worsening of ≥10 points was defined as deterioration. This threshold was chosen to ensure consistent cut-off points between the VAS and transformed NRS.

All dichotomised PROM changes were then included as potential explanatory variables in a regression model, with active disease as the dependent variable. The aim was to determine whether a deterioration in PROM score between the previous and current visit was associated with having active disease at the current visit. In RA, active disease was defined as a change in disease activity from DAS ≤2.4 to DAS >2.4 over time.[15] In PsA, active disease was defined by a change from DAPSA ≤14 to a DAPSA >14 over time.[16] The analysis included visits where patients either transitioned from well-controlled to active disease or maintained well-controlled disease over both visits. Mixed effects logistic regression models with an unstructured variance-covariance structure of the random effects and a robust sandwich covariate estimator were used. The significant variables were identified using backward elimination. Variables with a p>0.05 were removed from the model. To account for repeated measurements within one patient, a random intercept at the patient level was used.

The selected PROMs from the regression analysis were combined to form a summarised score: $PROM_1 + PROM_2 + PROM_3 + \ldots PROM_n$. This score indicates the number of deteriorated PROMs. For example, a score of 0 means that none of the selected PROMs deteriorated, a score of 2 means that two PROMs deteriorated. The combination of PROMs was tested for its ability to discriminate between active and well-controlled disease. Sensitivity and specificity were calculated for different cut-off values of the summarised score. The overall diagnostic performance was assessed using the area under the receiver-operating curve (AUC-ROC). An AUC of 0.5–0.6 indicates no discriminative ability, 0.6–0.7 is considered poor, 0.7–0.8 is acceptable, 0.8–0.9 is excellent and an AUC greater than 0.9 is considered outstanding.[38]

We did not report the positive and negative predictive values, as the prevalence of disease flare in trial data may not accurately represent the general patient population. Bootstrapping was used for internal validation of the diagnostic accuracy of the main model. All analyses were conducted separately for RA and PsA.

In addition, several sensitivity analyses were performed. First, to assess any circularity of the results due to a PROM (general health/PGA) in the DAS, we also analysed the discriminative ability of the combined PROMs using the three-item DAS44 (ie, the DAS without general health/PGA).[24] Second, to assess the robustness of the PsA model, we also used a disease activity measure that includes enthesitis and skin involvement, namely minimal disease activity (MDA), which was used instead of the DAPSA. Active disease was defined as shifting from MDA to no MDA. A patient is in MDA if ≥5 out of 7 criteria are met: 66-SJC ≤1, 68-TJC ≤1, BSA ≤3%,

VAS pain ≤15 mm, VAS general health ≤20 mm, HAQ ≤0.5 and LEI ≤1. Third, we excluded observations that changed from well-controlled disease to active disease with a non-clinically relevant change in DAS or DAPSA. Thus, only observations with a minimal DAS change of >0.6 for RA, or a minimal DAPSA change of >7.2 for PsA, were included.[24 39] Fourth, we considered a change above the minimal clinically important difference (MCID) for the HAQ-DI (MCID ≥0.22 for RA and ≥0.35 for PsA) and EQ-5D (MCID ≥0.05). Fifth, we assessed if the discriminative ability would improve if we took the prognostic power of the individual PROMs into account. We accounted for this by adding the magnitude of the regression coefficients (β) to the summarised score: $\beta_1 \cdot PROM_1 + \beta_2 \cdot PROM_2 + \beta_3 \cdot PROM_3 + \ldots \beta_n \cdot PROM_n$. Finally, we assessed the additional value of productivity loss for the ability to detect active disease and added presenteeism into the sum score to see whether the model would improve.

If general health/PGA was missing, we used the three-item DAS to determine a patient's disease status. This was the case in 57 out of 4594 RA visits. To determine if there was a difference in response between patients with a well-controlled disease and those with an active disease, we analysed missing PROM scores separately for each group. As shown in online supplemental table S1, the percentage of missing PROM scores varied between PROMs, ranging from 0% to 1% for PGA to 67% for productivity loss. In RA, missingness of PROM scores was similar between visits with a well-controlled and active disease, while in PsA missingness was 6%–7% higher in visits where patients developed a disease flare compared with those who continued having a well-controlled disease. Missing values were treated as missing and were not included in the analysis. All analyses were performed by using Stata V.18 (StataCorp).

RESULTS

Patient characteristics

From the 587 early UA/RA and 189 established RA patients, 495 and 188 had disease activity and PROM data at the same consecutive time points available, respectively. From the PsA cohort, 525 patients were included. Baseline characteristics of the included patients are shown in table 2. The median symptom duration (IQR) was 0.4 (0.3–0.6) years for early UA/RA and 6.2 (4.1–8.9) years for established RA. The mean DAS (SD) at baseline was 3.0 (1.0) in the early UA/RA and 1.0 (0.5) in the established RA group. ACPA positivity and RF positivity were lower in the early UA/RA (41% and 41%, respectively) cohort compared with the established RA cohort (72% and 68%, respectively) because UA patients were also included in the early UA/RA cohort.

In the PsA cohort, the median symptom duration (IQR) at baseline was 0.7 (0.3–2.3) years and the median DAPSA (IQR) was 14.0 (9.8–21.3). Arthritis in one or more joints was present in 77% of the patients at time of diagnosis. Psoriasis occurred in 84% of PsA patients with a median (IQR) BSA of 3% (1%–5%). Enthesitis was present in 35% of patients with a median (IQR) LEI of 1 (1-2).

Selection of PROMs

In total, 4594 UA/RA visits and 1154 PsA visits were included in our analysis. Patients had active disease in 358 (8%) and 177 (15%) of those UA/RA and PsA visits, respectively. In the logistic regression analysis, deterioration in general health/PGA, HAQ-DI, EQ-5D and pain were associated with having active disease in both UA/RA and PsA (table 3). As can be seen in table 3A, the regression coefficients of general health/PGA within the UA/RA group are higher compared with the other PROMs. For the PsA group, both general health and pain have larger coefficients compared with the HAQ-DI and EQ-5D (table 3B).

Test characteristics and discriminative ability of combination of explanatory PROMs

For the PROMs that remained significant in the multivariable analysis, that is, general health/PGA, HAQ-DI, EQ-5D and pain, a sum score was calculated for every visit for UA/RA and PsA separately. This resulted in a sum score ranging from 0 to 4 per visit. In table 4, the number of visits per sum score is shown. In UA/RA, 69% of the visits in which patients had a continuous well-controlled disease and in 28% of the visits in which patients had an active disease the sum score was ≤1. In PsA, 73% and 16% of the visits with a continuous well-controlled and active disease had a sum score ≤1, respectively. The mean difference in disease activity (DAS or DAPSA) increased with the number of deteriorated PROMs, which applied both to visits where patients had a continuous well-controlled disease and to visits where they developed a disease flare (table 4).

The combination of the four selected PROMs (general health/PGA, HAQ-DI, EQ-5D and pain) showed an acceptable discriminative ability for UA/RA (AUC=0.763, figure 1A) and excellent discriminative ability for PsA (AUC=0.856, figure 1B). Table 5 shows the sensitivity, specificity and positive-likelihood and negative-likelihood ratios for different cut-points. For example, if a cut-point of ≥1 deteriorated PROMs is used, 40% of the visits in which UA/RA patients have a continuous well-controlled disease are correctly specified (specificity), while 10% of the visits in which patients have an active disease are overlooked (1-sensitivity). In PsA, these percentages are 41% and 4%, respectively. Internal validation using bootstrapping showed similar results: optimism adjusted AUC (95% CI) for UA/RA is 0.759 (0.725–0.798) and for PsA 0.856 (0.825–0.887).

Sensitivity analyses

In the first sensitivity analysis, the three-item DAS was used instead of the original DAS to define active disease to exclude circularity for general health/PGA. The

Table 2 Baseline characteristics for included early UA/RA, established RA and PsA patients

	Early UA/RA (n=495)	Established RA (n=188)	PsA (n=525)
Demographics			
Sex (female), n (%)	320 (65)	124 (66)	245 (47)
Age (years), mean (SD)	53 (15)	57 (12)	50 (14)
Symptom duration (years)	0.4 (0.3–0.6)	6.2 (4.1–8.9)	0.7 (0.3–2.3)
UA/RA-specific disease activity			
ACPA (positive), n (%)	204 (41)	134 (72)	–
RF (positive), n (%)	204 (41)	127 (68)	–
DAS, mean (SD)	3.0 (1.0)	1.0 (0.5)	–
PsA-specific disease activity			
DAPSA	–	–	14.0 (9.8–21.3)
Arthritis, n (%)	–	–	406 (77)
Psoriasis, n (%)	–	–	403 (84)
BSA in case of psoriasis, n (%)	–	–	3 (1–5)
Enthesitis, n (%)	–	–	185 (35)
LEI in case of enthesitis	–	–	1 (1–2)
Disease activity components			
Swollen joint count*	6 (2–11)	0 (0–0)	2 (1–5)
Tender joint count†	7 (3–13)	0 (0–1)	3 (1–6)
ESR (mm/hour)	19 (11–35)	8 (3–15)	10 (6–24)
CRP (mg/L)	7 (3–18)	2 (1–5)	4 (1–11)
VAS GH/PGA (0–100 mm)	49 (28–66)	14 (4–27)	45 (23–63)
VAS pain (0–100 mm)	–	–	44 (23–63)

All results are shown as median (IQR) unless indicated otherwise.
*Number of swollen joints measured in 44 joints in RA patients and in 66 joints in PsA patients.
†Number of tender joints measured in 53 joints in RA patients and in 68 joints in PsA patients.
ACPA, anti-citrullinated protein antibody; BSA, body surface area; CRP, C reactive protein; DAPSA, Disease Activity Index in Psoriatic Arthritis; DAS, Disease Activity Score; ESR, erythrocyte sedimentation rate; GH, general health; LEI, Leeds Enthesitis Index; PGA, patient global assessment; PsA, psoriatic arthritis; RA, rheumatoid arthritis; RF, rheumatoid factor; UA, undifferentiated arthritis; VAS, Visual Analogue Scale.

discriminative ability decreased if the three-item DAS was used for UA/RA, resulting in an AUC of 0.71 (online supplemental figure S1A and table S2A). If MDA instead of the DAPSA was used in PsA, the discriminative ability decreased but remained excellent (AUC 0.82) (online supplemental figure S1B and table S2B). If observations with only a small change in disease activity were excluded, the AUC improved for UA/RA (AUC=0.78) as well as for PsA (AUC=0.89) (online supplemental figure S2 and table S3). We also considered a change above the MCID for the HAQ-DI and EQ-5D, which showed results similar or better to our main analysis (AUC=0.75 for UA/RA and AUC=0.89 for PsA, data not shown). The application of weight factors, which are based on the coefficients from the multivariable analysis as shown in table 3, improved the discriminative ability for UA/RA, as well as PsA (AUC RA=0.79, AUC PsA=0.87, online supplemental figure S3 and table S4). Lastly, if productivity loss was added to the sum score for the patients with paid work, the discriminative ability decreased, resulting in an AUC of 0.73 for UA/RA and 0.86 for PsA (online supplemental figure S4 and table S5).

DISCUSSION

This study investigated whether a combination of the ICHOM-recommended PROMs can adequately discriminate between well-controlled and active disease in RA and PsA patients. If the combined PROMs have a good discriminative ability, this could be used to screen for active disease in patients with inflammatory arthritis in the future. We found that a combination of general health/PGA, HAQ-DI, EQ-5D and pain has acceptable discriminative ability in RA and excellent discriminative ability in PsA. Depending on the cut-point chosen, the combined PROMs can either adequately rule out active disease (when a low cut-point is used) or adequately detect active disease (when a high cut-point is used). If a low cut-point of ≥1 deteriorated PROMs was used, the sensitivity was 90% and 96%, and the specificity was 40% and 41% for

Table 3 Multivariable analysis on the association between deteriorated PROMs between the prior and evaluated visit and having active disease, stratified for UA/RA and PsA

ΔPROM	β	OR (95% CI)	P value	Weight*
A	Undifferentiated arthritis/rheumatoid arthritis (n=4594)			
General health/PGA†	1.95	7.02 (4.95-9.95)	<0.001	2
HAQ-DI	0.69	2.00 (1.44-2.77)	<0.001	0.5
EQ-5D-3L	0.39	1.47 (1.06-2.06)	0.023	0.5
Pain†	0.44	1.56 (1.10-2.20)	0.012	0.5
B	Psoriatic arthritis (n=1154)			
General health†	1.87	6.50 (3.67-11.49)	<0.001	2
HAQ-DI	0.86	2.37 (1.45-3.88)	0.001	1
EQ-5D-5L	0.71	2.03 (1.24-3.32)	0.005	0.5
Pain†	1.63	5.09 (2.86-9.05)	<0.001	1.5

*Weight was applied according to the rounded coefficients (β) of the multivariable analysis.
†A minimal change of ≥10 was necessary to be classified as a deteriorated PROM score between 2 consecutive visits.
EQ-5D-3L, 3-level EQ-5D; EQ-5D-5L, 5-level EQ-5D; HAQ-DI, Health Assessment Questionnaire-Disability Index; PGA, patient global assessment; PROM, patient-reported outcome measure; PsA, psoriatic arthritis; RA, rheumatoid arthritis; UA, undifferentiated arthritis.

detecting active disease in RA and PsA, respectively. Sensitivity analyses showed that after applying weight factors to the sum score, the discriminative ability increased. If only observations with a clinically relevant change in DAS or DAPSA were included the AUC increased as well. Adding productivity loss to our working population was not of added value. If the three-item DAS or MDA is used to differentiate between active and well-controlled disease, the AUC decreased in both RA and PsA. Thus, if a more strict or objective disease activity score is used, the model performance decreases.

The prerequisite for remote monitoring in inflammatory arthritis is a good balance between not missing active disease and reducing 'unnecessary' outpatient clinic visits. The ability to detect active disease with remote monitoring is important since, according to the T2T recommendations, this requires a treatment change.[1–4] As an example, if our model with a cut-point of ≥1

Table 4 Number of visits with a well-controlled and active disease per total number of deteriorated PROMs for UA/RA and PsA

Deteriorated PROMs (n)	Visits with well-controlled disease, n (%)	ΔDAS/ΔDAPSA between visits*	Visits with active disease, n (%)	ΔDAS/ΔDAPSA between visits*
A	Undifferentiated arthritis/rheumatoid arthritis			
0	1257 (40)	−0.19 (0.51)	25 (9)	0.76 (0.52)
1	918 (29)	−0.07 (0.52)	49 (19)	0.77 (0.48)
2	541 (17)	0.06 (0.54)	63 (24)	1.01 (0.50)
3	265 (9)	0.28 (0.59)	69 (26)	0.93 (0.56)
4	132 (4)	0.63 (0.55)	58 (22)	1.28 (0.75)
Total, n (%)	3113 (100)		264 (100)	
B	Psoriatic arthritis			
0	386 (41)	−2.59 (3.31)	6 (4)	6.83 (3.80)
1	300 (32)	−1.11 (3.29)	19 (12)	7.26 (5.75)
2	147 (16)	0.82 (3.17)	32 (20)	7.03 (4.47)
3	75 (8)	2.52 (2.93)	47 (28)	8.09 (4.82)
4	26 (3)	5.45 (3.37)	54 (34)	12.11 (8.21)
Total, n (%)	934 (100)		158 (100)	

*Shown as mean (SD). For UA/RA the ΔDAS was shown, while for PsA the ΔDAPSA was shown.
Δ, delta; DAPSA, Disease Activity Index in Psoriatic Arthritis; DAS, Disease Activity Score; PROMs, patient-reported outcome measures; PsA, psoriatic arthritis; RA, rheumatoid arthritis; UA, undifferentiated arthritis.

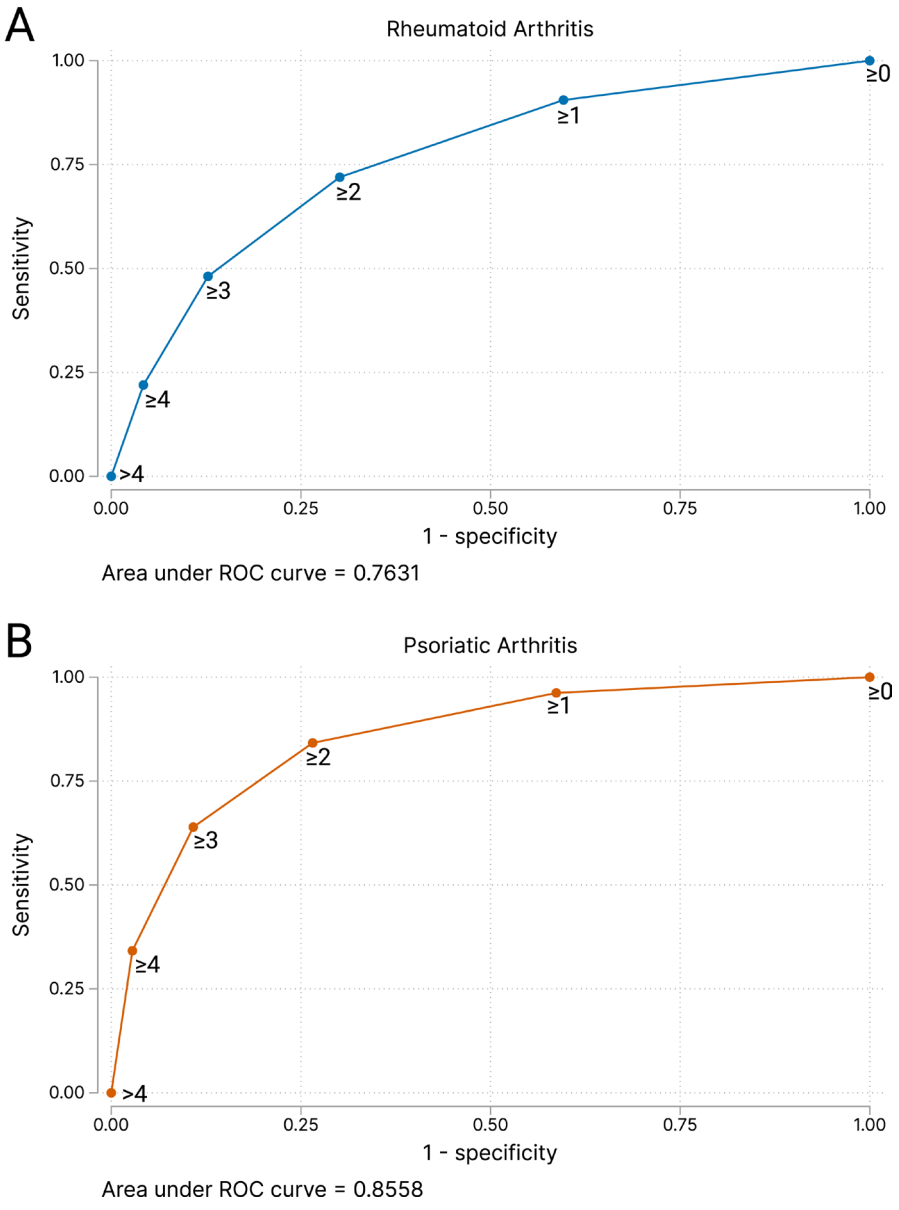

Figure 1 Test characteristics per cut-point for (A) UA/RA and (B) PsA. Each cut-point represents the number of PROMs that deteriorated between the prior and evaluated visit. PROM, patient-reported outcome measure; PsA, psoriatic arthritis; RA, rheumatoid arthritis; ROC, receiver operating characteristics; UA, undifferentiated arthritis.

deteriorated PROMs would be used, in 38% of RA visits (n=1282) the test would indicate well-controlled disease. In 25 of these 1282 visits (2%), the patient would actually have active disease which would be overlooked. In PsA in 36% of visits (n=392), the test would indicate well-controlled disease. In 6 of these 392 visits (2%), active disease would be missed. Still, RA and PsA patients with an active disease and 0 deteriorated PROMs only have a mean increase in their DAS and DAPSA of 0.8 and 6.8, respectively. For RA, this is just above the MCID of the DAS (>0.6) and for PsA this is below the MCID of the DAPSA (7.2). Moreover, in a real-world setting, patients should always have the option of scheduling a visit if they think it is necessary.

Previous studies on remote monitoring with a combination of PROMs were less sensitive in detecting active disease, but more patients with a well-controlled disease were recognised. Hendrikx *et al*, for example, reported that with a combination of VAS general health, VAS disease activity and VAS pain 56% of RA patients with an active disease and 70% with a well-controlled disease were correctly specified, respectively.[13] Seppen *et al*, on the other hand, used a combination of the RAPID3 and a flare question and they were able to correctly specify 39% of the patients who would need a treatment intensification over 3 months, and 96% of the patients who would not need an intensification.[14] To improve the sensitivity of the combined PROMs, they could be used in combination with more objective measurements such as digital biomarkers, including wearables and passive smartphone behaviour.[40 41] Alternatively, it might be possible to select the most sensitive items from each PROM and combine those into one new PROM.

Table 5 Sensitivity, specificity, LR+ and LR– for different cut-points of the sum score for (A) UA/RA and (B) PsA

Cut-point	Sensitivity (%)	Specificity (%)	LR+	LR–
A	Undifferentiated arthritis/rheumatoid arthritis			
≥0	100	0	1.0	
≥1	90	40	1.5	0.2
≥2	72	70	2.4	0.4
≥3	48	87	3.7	0.6
≥4	21	96	5.2	0.8
>4	0	100		1.0
B	Psoriatic arthritis			
≥0	100	0	1.0	
≥1	96	41	1.6	0.1
≥2	84	73	3.2	0.2
≥3	64	89	5.6	0.4
≥4	34	97	12.3	0.7
>4	0	100		1.0

LR–, negative likelihood ratio; LR+, positive likelihood ratio; PROM, patient-reported outcome measure; PsA, psoriatic arthritis; RA, rheumatoid arthritis; UA, undifferentiated arthritis.

Apart from a high discriminative ability, other factors are also important to consider before implementing remote monitoring in routine care. For example, barriers are the time burden of filling out questionnaires and difficulties of completing PROM questionnaires for patients with low language or computer literacy.[42] Thus, the use of electronic PROMs should not be too time-consuming, and they should also be easy to use. In addition, rheumatologists should be aware of the influence of joint damage and comorbidities on PROM scores, for example, functional limitations and pain, especially in established disease.[43–45] This is one of the reasons why we examined a change in PROM score instead of a fixed threshold. Furthermore, validation in other datasets is necessary to assess the generalisability of our results. However, if these barriers are taken into account and external validation shows similar results, the use of electronic PROMs might be a valuable addition to screen for active disease, which is very important in the management of chronic conditions such as RA and PsA.

One should also keep in mind that in our study some of the RA and PsA patients with a well-controlled disease had several deteriorated PROMs, indicating a significant increase in disease burden, without having active disease. Therefore, one may question whether the outpatient clinic visits of aforementioned patients are truly 'unnecessary', just because they did not have active disease according to our definition. Especially since nowadays, a dual T2T approach is recommended, in which the targets are (1) control of inflammation and (2) control of disease impact.[10] One option could be to schedule a telephone consultation with patients with significantly deteriorated PROMs to inquire about their health status and to schedule an outpatient clinic visit if necessary. This could also help ensure that inflammatory arthritis healthcare remains accessible in the future.[5 6]

The strengths of our study are the inclusion of early UA/RA and established RA patients, and early PsA patients. Furthermore, the RA data came from two randomised controlled trials with a T2T approach, in which the medication protocol was fixed. In addition, PROMs out of all ICHOM-recommended outcome domains were assessed for inclusion in the combined PROM.

Our study also has limitations. The RA data originated from randomised controlled trials with protocolised treatment strategies, while the PsA data originated from a cohort study. This difference could explain the higher number of PsA visits in which patients had an active disease. In addition, only 1-year follow-up data with 3 monthly visits were available in PsA, whereas 3 and 2 years of follow-up data were available in early UA/RA and established RA, respectively. A challenge in using PROMs is the test–retest reliability. We decided to consider any deterioration in the HAQ-DI and the EQ-5D score as valuable. The test–retest reliability of the HAQ-DI and EQ-5D has shown to be good in both RA and PsA.[31 46–48] However, even if the disease activity is stable, patients will fill in the questionnaires slightly different each time, leading to a small change in PROM scores. However, the same is true for disease activity as measured by a physician. To account for these measurement errors, we also considered a change above the MCID for the DAS/DAPSA and the HAQ-DI and EQ-5D, which showed similar results. Furthermore, we used general health/PGA in the combined PROM, but this PROM is also necessary to calculate the DAS, which might lead to circularity. The sensitivity analysis in which we used the three-item DAS instead of the original DAS indeed showed a slightly lower discriminative ability than our main analysis. However, the test characteristics remained acceptable.

In conclusion, a combination of general health/PGA, HAQ-DI, EQ-5D and pain is able to differentiate between well-controlled and active disease in RA and PsA with acceptable to excellent diagnostic accuracy. The combination of these PROMs could possibly contribute to remote monitoring of disease status in RA and PsA patients in the future.

Author affiliations
[1]Department of Rheumatology, Erasmus MC, Rotterdam, The Netherlands
[2]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[3]Department of Rheumatology & Clinical Immunology and Julius Center, University Medical Center Utrecht, Utrecht, The Netherlands
[4]Department of Rheumatology and Immunology, Maasstad Hospital, Rotterdam, The Netherlands
[5]Department of Rheumatology, Albert Schweitzer Hospital, Dordrecht, The Netherlands
[6]Department of Rheumatology, Franciscus Gasthuis & Vlietland, Rotterdam, The Netherlands
[7]Department of Rheumatology, Reumazorg Zuid West Nederland, Roosendaal, The Netherlands

Acknowledgements We gratefully thank all participating patients, rheumatologists and research nurses.

Contributors AEML performed the statistical analysis. SVJSH and AEML drafted the manuscript. PHPdJ contributed to the analysis. All authors contributed to the design, revised the manuscript and read and approved the final manuscript. PHPdJ is the guarantor.

Funding This study was supported by an unrestricted grant from ZonMW (grant number 171102014), Pfizer (grant number WI229707) and Abbvie B.V. (grant number SA-003854).

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Written informed consent was obtained from all participants according to the Declaration of Helsinki. The studies were approved by the local medical research ethics committee of Erasmus Medical Center Rotterdam, the Netherlands (MEC-2006-252 for tREACH, MEC-2011-141 for TARA and MEC-2012-549 for DEPAR).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

ORCID iDs
Agnes E M Looijen http://orcid.org/0000-0002-5280-2182
Selinde V J Snoeck Henkemans http://orcid.org/0000-0003-2319-7777
Annette H M van der Helm-van Mil http://orcid.org/0000-0001-8572-1437
Gonul Hazal Koc http://orcid.org/0009-0007-3787-4552
Jolanda J Luime http://orcid.org/0000-0002-1071-7932
Marc R Kok http://orcid.org/0000-0003-2394-6926
Pascal H P de Jong http://orcid.org/0000-0001-6628-6222

REFERENCES

1 Smolen JS, Schöls M, Braun J, *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.
2 Smolen JS, Breedveld FC, Burmester GR, *et al*. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. *Ann Rheum Dis* 2016;75:3–15.
3 Smolen JS, Landewé RBM, Bergstra SA, *et al*. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2022 update. *Ann Rheum Dis* 2022.
4 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.
5 Hootman JM, Helmick CG, Barbour KE, *et al*. Updated Projected Prevalence of Self-Reported Doctor-Diagnosed Arthritis and Arthritis-Attributable Activity Limitation Among US Adults, 2015-2040. *Arthritis Rheumatol* 2016;68:1582–7.
6 Battafarano DF, Ditmyer M, Bolster MB, *et al*. American College of Rheumatology Workforce Study: Supply and Demand Projections of Adult Rheumatology Workforce, 2015–2030. *Arthritis Care Res* 2015;70:617–26.
7 Miloslavsky EM, Bolster MB. Addressing the rheumatology workforce shortage: A multifaceted approach. *Semin Arthritis Rheum* 2020;50:791–6.
8 van Riel P, Alten R, Combe B, *et al*. Improving inflammatory arthritis management through tighter monitoring of patients and the use of innovative electronic tools. *RMD Open* 2016;2:e000302.
9 van Tuyl LHD, Sadlonova M, Hewlett S, *et al*. The patient perspective on absence of disease activity in rheumatoid arthritis: a survey to identify key domains of patient-perceived remission. *Ann Rheum Dis* 2017;76:855–61.
10 Ferreira RJO, Ndosi M, de Wit M, *et al*. Dual target strategy: a proposal to mitigate the risk of overtreatment and enhance patient satisfaction in rheumatoid arthritis. *Ann Rheum Dis* 2019;78:e109.
11 Arumalla N, Chan CKD, Gibson M, *et al*. The Clinical Impact of Electronic Patient-Reported Outcome Measures in the Remote Monitoring of Inflammatory Arthritis: A Systematic Review and Meta-analysis. *Arthritis Rheumatol* 2023;75:1892–903.
12 Looijen AEM, van Mulligen E, Vonkeman HE, *et al*. Responsiveness and sensitivity of PROMs to change in disease activity status in early and established rheumatoid arthritis. *Rheumatol (Oxford)* 2024.:keae213.
13 Hendrikx J, Fransen J, van Riel PLCM. Monitoring rheumatoid arthritis using an algorithm based on patient-reported outcome measures: a first step towards personalised healthcare. *RMD Open* 2015;1:e000114.
14 Seppen BF, Verkleij SJ, Wiegel J, *et al*. Probability of Medication Intensifications in Rheumatoid Arthritis Patients With Low Disease Activity Scores on Their Patient-Reported Outcomes: A Medical-Records Review Study. *J Clin Rheumatol* 2022;28:397–401.
15 Dougados M, Aletaha D, van Riel P. Disease activity measures for rheumatoid arthritis. *Clin Exp Rheumatol* 2007;25:S22–9.
16 Schoels MM, Aletaha D, Alasti F, *et al*. Disease activity in psoriatic arthritis (PsA): defining remission and treatment success using the DAPSA score. *Ann Rheum Dis* 2016;75:811–8.
17 Arnett FC, Edworthy SM, Bloch DA, *et al*. The american rheumatism association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis & Rheum* 1988;31:315–24.
18 Smolen JS, Aletaha D, Bijlsma JWJ, *et al*. Treating rheumatoid arthritis to target: recommendations of an international task force. *Ann Rheum Dis* 2010;69:631–7.
19 Claessen SJJ, Hazes JMW, Huisman MAM, *et al*. Use of risk stratification to target therapies in patients with recent onset arthritis; design of a prospective randomized multicenter controlled trial. *BMC Musculoskelet Disord* 2009;10:71.
20 Chatzidionysiou K, Fragoulis GE. Established rheumatoid arthritis - Redefining the concept. *Best Pract Res Clin Rheumatol* 2019;33:S1521-6942(19)30172-X.
21 van Mulligen E, de Jong PHP, Kuijper TM, *et al*. Gradual tapering TNF inhibitors versus conventional synthetic DMARDs after achieving controlled disease in patients with rheumatoid arthritis: first-year results of the randomised controlled TARA study. *Ann Rheum Dis* 2019;78:746–53.
22 Kasiem FR, Luime JJ, Vis M, *et al*. Lessons learned from clinical phenotypes in early psoriatic arthritis: the real-world Dutch south west Early Psoriatic ARthritis study. *Scand J Rheumatol* 2021;50:124–31.
23 Oude Voshaar MAH, Das Gupta Z, Bijlsma JWJ, *et al*. International Consortium for Health Outcome Measurement Set of Outcomes That Matter to People Living With Inflammatory Arthritis: Consensus From an International Working Group. *Arthritis Care Res (Hoboken)* 2019;71:1556–65.
24 Fransen J, van Riel P. The Disease Activity Score and the EULAR response criteria. *Rheum Dis Clin North Am* 2009;35:745–57.
25 Nell-Duxneuner VP, Stamm TA, Machold KP, *et al*. Evaluation of the appropriateness of composite disease activity measures for assessment of psoriatic arthritis. *Ann Rheum Dis* 2010;69:546–9.
26 Schoels M, Aletaha D, Funovits J, *et al*. Application of the DAREA/DAPSA score for assessment of disease activity in psoriatic arthritis. *Ann Rheum Dis* 2010;69:1441–7.
27 Cauli A, Gladman DD, Mathieu A, *et al*. Patient global assessment in psoriatic arthritis: a multicenter GRAPPA and OMERACT study. *J Rheumatol* 2011;38:898–903.
28 Bruce B, Fries JF. The Health Assessment Questionnaire (HAQ). *Clin Exp Rheumatol* 2005;23:S14–8.
29 Boers M, Jacobs JWG, van Vliet Vlieland TPM, *et al*. Consensus Dutch Health Assessment Questionnaire. *Ann Rheum Dis* 2007;66:132–3.
30 Fries JF, Spitz P, Kraines RG, *et al*. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23:137–45.
31 Hurst NP, Kind P, Ruta D, *et al*. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). *Br J Rheumatol* 1997;36:551–9.
32 EuroQol G. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199–208.
33 Brooks R. EuroQol: the current state of play. *Health Policy* 1996;37:53–72.

34 Lamers LM, Stalmeier PFM, McDonnell J, *et al*. Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff. *Ned Tijdschr Geneeskd* 2005;149:1574–8.
35 Strand V, Boers M, Idzerda L, *et al*. It's good to feel better but it's better to feel good and even better to feel good as soon as possible for as long as possible. Response criteria and the importance of change at OMERACT 10. *J Rheumatol* 2011;38:1720–7.
36 van Roijen L, Essink-Bot ML, Koopmanschap MA, *et al*. Labor and health status in economic evaluation of health care. The Health and Labor Questionnaire. *Int J Technol Assess Health Care* 1996;12:405–15.
37 Bouwmans C, Krol M, Severens H, *et al*. The iMTA Productivity Cost Questionnaire: A Standardized Instrument for Measuring and Valuing Health-Related Productivity Losses. *V H* 2015;18:753–8.
38 Hosmer DW, Lemeshow S, Sturdivant RX. *Applied logistic regression*. John Wiley & Sons, 2013.
39 Tillett W, FitzGerald O, Coates LC, *et al*. Composite Measures for Clinical Trials in Psoriatic Arthritis: Testing Pain and Fatigue Modifications in a UK Multicenter Study. *J Rheumatol* 2021.:jrheum.201674.
40 Mollard E, Pedro S, Schumacher R, *et al*. Smartphone-based behavioral monitoring and patient-reported outcomes in adults with rheumatic and musculoskeletal disease. *BMC Musculoskelet Disord* 2022;23:566.
41 Huhn S, Axt M, Gunga H-C, *et al*. The Impact of Wearable Technologies in Health Research: Scoping Review. *JMIR Mhealth Uhealth* 2022;10:e34384.
42 Glenwright BG, Simmich J, Cottrell M, *et al*. Facilitators and barriers to implementing electronic patient-reported outcome and experience measures in a health care setting: a systematic review. *J Patient Rep Outcomes* 2023;7:13.
43 Bombardier C, Barbieri M, Parthan A, *et al*. The relationship between joint damage and functional disability in rheumatoid arthritis: a systematic review. *Ann Rheum Dis* 2012;71:836–44.
44 Walsh DA, McWilliams DF. Pain in rheumatoid arthritis. *Curr Pain Headache Rep* 2012;16:509–17.
45 Cañete JD, Tasende JAP, Laserna FJR, *et al*. The Impact of Comorbidity on Patient-Reported Outcomes in Psoriatic Arthritis: A Systematic Literature Review. *Rheumatol Ther* 2020;7:237–57.
46 Leung YY, Tillett W, Hojgaard P, *et al*. Test-retest Reliability for HAQ-DI and SF-36 PF for the Measurement of Physical Function in Psoriatic Arthritis. *J Rheumatol* 2021;48:1547–51.
47 Maska L, Anderson J, Michaud K. Measures of functional status and quality of life in rheumatoid arthritis: Health Assessment Questionnaire Disability Index (HAQ), Modified Health Assessment Questionnaire (MHAQ), Multidimensional Health Assessment Questionnaire (MDHAQ), Health Assessment Questionnaire II (HAQ-II), Improved Health Assessment Questionnaire (Improved HAQ), and Rheumatoid Arthritis Quality of Life (RAQoL). *Arthritis Care Res (Hoboken)* 2011;63:S4–13.
48 Tsang HHL, Wong CKH, Cheung PWH, *et al*. Responsiveness of the EuroQoL 5-Dimension (EQ-5D) questionnaire in patients with spondyloarthritis. *BMC Musculoskelet Disord* 2021;22:439.