New indicator for discordance between patient-reported and traditional disease activity outcomes in patients with early Rheumatoid Arthritis

Pazmino, Sofia; Lovik, Anikó; Boonen, Annelies; De Cock, Diederik; Stouten, Veerle; Joly, Johan; Doumen, Michaël; Bertrand, Delphine; Westhovens, René; Verschueren, Patrick

Published in:
Rheumatology (Oxford, England)

DOI:
10.1093/rheumatology/keac213

Publication date:
2023

New indicator for discordance between patient-reported and traditional disease activity outcomes in patients with early Rheumatoid Arthritis

Sofia Pazmino [1], Anikó Lovik[2], Annelies Boonen[3,4], Diederik De Cock[1,5], Veerle Stouten[1], Johan Joly[6], Michaël Doumen[1], Delphine Bertrand[1], René Westhovens[1,6], Patrick Verschueren[1,6]

[1]Department of Development and Regeneration, Skeletal Biology and Engineering Research Centre, KU Leuven, Leuven, Belgium

[2]I-BioStat, Department of Public Health and Primary Care, KU Leuven, Leuven, Belgium

[3]Department of Internal Medicine, Division of Rheumatology, Maastricht University Medical Centre, Maastricht, The Netherlands

[4]Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, The Netherlands

[5]Laboratory of Clinical and Experimental Endocrinology, University of Leuven, Leuven, Belgium.

[6]Division of Rheumatology, University Hospitals Leuven, Leuven, Belgium

Corresponding author: Sofia Pazmino. Mailing address: Division of Rheumatology, University Hospitals Leuven, Herestraat 49, B-3000 Leuven, Belgium.

ORCID iD https://orcid.org/0000-0001-8579-6914

E-mail: sofia.pazmino@kuleuven.be; Tel: +32-16-342541; Fax: +32-16-342543

KEY MESSAGES

- Including pain, fatigue and functionality when assessing early RA patients broadens evaluation of disease impact.
- Novel factor-score detects unmet needs to be used for predicting future disease control and quality-of-life.
- The earlier detection of unmet needs could signal the need for timely interdisciplinary interventions.

KEYWORDS:

Early rheumatoid arthritis, patient-reported outcomes, discordance, unmet needs, factor scores, measurement instrument

ABSTRACT

Objective: To unravel disease impact in early Rheumatoid Arthritis (RA) by separately quantifying patient-reported (PRF), clinical (CF) and laboratory (LF) factors. We propose a new indicator, the discordance score (DS), for early identification and prediction of patient's unmet needs and of future achievement of sustained remission (SR) and RA-related quality of life (QoL).

Methods: Factor-scores obtained by factor analysis in the CareRA trial, allowed to compute DS, reflecting the difference between PRF and the mean of CF and LF. Improvement from baseline to week 104 (%) and area-under-the-curve (AUC) across time points per factor-score were calculated and compared between patients achieving/not achieving sustained (week 16-104) remission (DAS28CRP<2.6) with ANOVA. Logistic and linear regressions were used to predict SR based on previous factor and discordance scores, and QoL at year 1 and 2 based on DS at week 16.

Results: PRF, CF and LF scores improved rapidly within 8 weeks. PRF improved 57%, CF 90% and LF 27%, in those achieving SR, compared to 32% (PRF:p=0.13), 77% (CF:p<0.001) and 9% (LF:p=0.36) in patients not achieving SR. Patients achieving SR had an AUC of 15.7, 3.4 and 4.8 for PRF, CF and LF, respectively, compared to 33.2, 10.1, and 7.2 in participants not achieving SR (p<0.001 for all). Early discordance was associated with later factor scores, QoL, and self-efficacy.

Conclusions: All factor scores improved rapidly, especially in patients achieving SR. Patient-reported burden improved less. Discordance scores could help predicting the need for additional non-pharmacological interventions to achieve SR and decrease disease impact.

INTRODUCTION

The primary clinical manifestation of Rheumatoid Arthritis (RA) is inflammation of the peripheral joints resulting in swelling, stiffness and pain.(1) However, more constitutional symptoms such as fatigue, restricted ability to work, and impact on other aspects of health-related quality of life (QoL) can be present.(2) Treating early to a target of remission or at least low disease activity is highly advocated.(3) This treatment strategy aims to maximise the long-term health-related quality of life through control of symptoms, prevention of structural damage, normalisation of function and social participation.(4) However, a clear definition of the treatment target is paramount to this strategy and should be stringent enough to seek "the best possible control of inflammation" while avoiding overtreatment.(5)

Specific instruments are used to measure the target of remission or low disease activity.(3) For example, the ACR/EULAR Boolean remission criterion is stringent, requiring swollen/tender joint counts (SJC/TJC) to be below or equal to (≤) 1, C-reactive protein (CRP) ≤1mg/dL and Patient's global health (PaGH) ≤1 (0-10 scale). When using this criterion for remission, it has been shown that one-third of patients with RA fail to reach remission solely because of PaGH (near-remission).(6) If the current treatment recommendations would be followed,(3) based on the Boolean remission definition a state of near-remission could lead to an adaptation of disease-modifying anti-rheumatic drugs (DMARDs) and potential overtreatment, although isolated PaGH elevation suggests needs that are not necessarily related to inflammation.(7) Hence, in such cases the remaining disease burden might not, or not only be mediated by disease activity. Several studies have shown a statistically significant correlation of PaGH with disease activity;(7,8) however, this correlation is absent in low levels of disease activity where PaGH is related to pain,

function and fatigue.(7–9) Unmet needs incorporated in the PaGH and their relative importance should be uncovered when aiming to reduce the broader impact of RA.

Recent research from our group has shown that adding pain, fatigue and physical function to the components included in the traditional composite disease activity measures may provide valuable additional information in evaluating the disease impact on the patient while keeping the traditional components intact.(10) Overall, three factors or components were identified: a patient-reported (PRF), clinical (CF) and laboratory (LF) factor. The patient-reported factor consists of PaGH, pain, fatigue and physical function (HAQ): four self-reported measures that add a separate patient-experienced aspect of evaluating the disease burden. The clinical factor contains three variables: the physician's global health assessment (PhGH), swollen and tender joint counts, all of which are common clinical measures. The laboratory factor comprises only two laboratory measures: CRP and ESR. CF and LF scores have long been part of measuring disease activity, while patient-reported factor is a novel addition. The three factors have also been replicated in a patient sample with established RA.(11) Based on these results and data from the CareRA trial(12), this paper aims to explore the potential benefit of evaluating these factor scores and their patterns over time to better understand the disease burden, both from the clinical and the patient's perspective. For this reason, we will plot the evolution of the three factors in the CareRA trial and take their differences under the loop with a new indicator called the discordance score, which allows us to study discorance between the traditional disease activity measures and patient-reported outcomes. We will attempt to predict from baseline the patient-reported, clinical and laboratory factors and discordance scores both later scores and QoL outcomes. Moreover, by comparing patients achieving sustained (week 16 to 104) remission

(DAS28CRP<2.6) to those who do not achieve sustained remission, we illustrate how even the patients achieving sustained remission still report unmet needs such as pain and fatigue.(13)

PATIENTS AND METHODS

CareRA was a 2-year open-label investigator-initiated pragmatic superiority trial (EudraCT number: 2008-007225-39, Clinical trials NCT01172639) conducted in 13 Flemish rheumatology centres.

Patients with recently diagnosed RA (≤1 year) were included and stratified into a high- or low-risk group based on classical factors of poor prognosis (erosions, rheumatoid factor (RF) and/or anti-citrullinated cyclic peptide (anti-CCP) positivity and baseline disease activity score in 28 joints with C-reactive protein (DAS28CRP) >3.2) and then randomised into four different treatment arms. High-risk patients were randomised to methotrexate (MTX) 15mg weekly with a step-down glucocorticoid (GC) scheme or to this combination together with either sulphasalazine or leflunomide. Low-risk patients were randomised to a tight step-up treatment starting with MTX monotherapy without GC or to MTX weekly with step-down GCs. Overall, around 70% of the CareRA participants achieved a status of good disease control after 2 years (DAS28CRP <2.6) with a treat-to-target approach.(12)

Clinical outcomes

Patients were assessed at screening, baseline and further at week 8, 16, 28, 40, 52, 65, 78, 91 and 104. Optional visits, if clinically required, could be performed. An electronic case report form was filled out and routinely monitored. Clinical, patient and laboratory parameters were collected at every visit: swollen (SJC28) and tender joint (TJC28) count in 28 joints, patient's global

health assessment (PaGH), physician's global health assessments (PhGH), C-reactive protein (CRP) or erythrocyte sedimentation rate (ESR), health assessment questionnaire (HAQ), pain and fatigue each on a visual analogue scale (VAS) of 0-100. Overall QoL was captured with the Short Form 36 (SF-36) and RA-specific QoL with the RAQoL questionnaire at baseline, week 16, 52 and 104. The SF-36 was standardised in 1990 as a self-report measure of functional health and well-being. The SF-36 consists of eight scales: physical functioning (10 items), role-physical (4 items), bodily pain (2 items), general health (5 items), vitality (4 items), social functioning (2 items), role-emotional (3 items), mental health (5 items), and a final item, termed self-reported health transition. To score the SF-36, scales are standardised with a scoring algorithm to obtain a score ranging from 0 to 100.(14) Higher scores indicate better health status. The RAQoL contains 30 yes/no questions regarding specific activities of daily living and quality of life. Each positive response is one point, a total sum is calculated giving a scale of 0-30.(15) Higher scores indicate worse health status. Self-efficacy was measured by the 2-sub-scale version of the Arthritis Self-Efficacy Scale (ASES) at week 52 and week 104.(16) The subscales for pain (PSE) and other symptoms (OSE) measure the patient's perceived confidence in their ability to control arthritis pain and to control symptoms of arthritis other than pain, respectively, with other means than medication. PSE and OSE are scored separately, each subscale ranges from 1 to 10 ("very uncertain" to "very certain"). The total ASES score (PSE+OSE) thus ranges from 2 to 20, with higher scores indicating stronger self-efficacy beliefs.

Statistical analyses

All randomised patients who had taken at least one medication dose, were considered for the intention to treat (ITT) analysis. Missing data were assumed to be missing at random and were

imputed with multiple imputation (classification and regression trees, n=100) by chained equations.(17) Patient-reported questionnaire data (RAQoL, ASES, and SF-36) were first handled as outlined in the individual questionnaire manuals. Missingness in clinical variables used to estimate the factor scores, disease activity per time point, and questionnaire scores were imputed. Besides the incomplete variables, treatment strategy, centre of recruitment, age, gender, RF, anti-CCP, erosions at baseline, and trial completion were included as predictors in the matrix. Each imputed dataset was analysed separately. Results of the 100 analyses were pooled using Rubin's rules.(18). As described previously by Pazmino et al. (10), three factors were identified using exploratory factor analysis on nine variables: Patient-Reported (PRF), Clinical (CF), and Laboratory (LF) factor. Factor loadings, which represent how strongly a variable relates to its factor, from the exploratory factor analyses (EFAs) were used as weights. The individual variables were normalised to a 0-1 scale considering clinically feasible minimum and maximum. Afterwards, the normalised (n.) variables were weighted -multiplied by the factor loadings- and summed up as follows, to calculate the different factor scores:

$$PRF = (0.90 * n.fatigue) + (0.87 * n.PaGH) + (0.86 * n.pain) + (0.57 * \text{n.HAQ})$$

$$CF = (0.92 * n.SJC28) + (0.89 * n.TJC28) + (0.76 * PhGH)$$

$$LF = (0.87 * n.CRP) + (0.78 * n.ESR)$$

Because the number of variables was different for each factor, the patient-reported, clinical and laboratory factor scores were re-scaled to 0-1 (higher values suggest more health impact) for comparisons. Thus, for each patient three factor scores per visit were obtained.

Furthermore, a discordance score (DS) for the PRF and CF/LF factor scores was calculated by subtracting the mean of the other two factor scores from the PRF score:

$$DS = PRF - \left(\frac{CF + LF}{2}\right)$$

The higher the score, the higher the patient-experienced impact not addressed by traditional measures of disease activity. The DS goes from -1 to 1. Correlations between all three factor scores and discordance scores per timepoint were calculated to assess the strength of the relationship of the factor and discordance scores at different time points. Due to the skewed distribution of the factor scores, Spearman correlations were used.

Furthermore, the percentage (%) of improvement from baseline to week 104 and the area-under-the-curve (AUC) across time points were calculated per score. Differences in % improvement and AUC were compared between patients not achieving and achieving early and sustained (week 16 to week 104) disease activity score remission (DAS28CRP <2.6) with ANOVA. Bonferroni correction was used for multiple testing. We chose to look at patients achieving early and sustained remission as a surrogate for “good responders” in whom we expected disease burden to be less pronounced.(13,19) To find a clinically meaningful cut-off for the discordance score to serve as an alarm system for unmet patient needs, we based ourselves on the SF-36. Receiver Operator Characteristics (ROC) Curves were fitted for the discordance score to be able to classify patients into good/poor health status based on SF-36. Without referring to norms, anytime a scale score is below 50, health status is below average.(14) Logistic regression was fitted to predict sustained remission based on factor and discordance scores at earlier time points with

and without adjustment for age and sex. Linear regressions to predict RA-QoL at week 52 and 104 based on the factor scores was also employed.

All analyses were performed with R V.4.1.2.

Ethics Approval

The study was approved by the leading Ethics Committee of the University Hospitals Leuven after consulting the medical ethics committee of each participating centre (ref s51411), and all study participants gave their written informed consent before inclusion. Patient consent for publication was not required.

RESULTS

Patients with early RA (n=379) were included with a mean (SD) age of 53.9 (13.0), 77% positive to RF or anti-CCP and 69% women (Table 1). In total, 289 were stratified to high risk and 90 to low risk. Good retention rates of 85% were observed overall. Missingness in the clinical variables over two years ranged from 0.3% to 39% per different time point and was 19% overall.

The patient-reported, clinical and laboratory factor scores improved rapidly over the first 8 weeks (Figure 1a). From baseline to week 104 the scores improved 41%, 78% and 10% for the patient-reported, clinical and laboratory factor scores respectively in the entire population (n=379), 57%, 90% and 27% in patients achieving sustained remission (n=122), and 32%, 77% and 9% in patients not achieving sustained remission (n=257). After correction for multiple testing, there was a statistically significant difference in the percentage of improvement from baseline to week 104 of CF ($p<0.001$) between patients in sustained remission or not, but not in the improvement of

PRF (p=0.13) or LF (p=0.36) (Figure 1b). Figure 1 also indicates higher values of PRF scores at all times, while CF and LF scores are rather similar, which illustrates the rationale for using the discordance score. Further detailed information can be found in Supplementary Table S1, available at *Rheumatology* online.

Patients in the CareRA trial had an AUC of 27.4, 7.9 and 6.4 for PRF, CF, LF scores respectively in the overall population. Those who achieved sustained remission had an AUC of 15.6, 3.4 and 4.8 for the patient-reported, clinical and laboratory factor scores respectively, compared to 33.2, 10.1, and 7.2 in participants not achieving sustained remission ($p<0.001$ for all AUCs, between patients in sustained remission or not).

Supplementary Figure S1 (available at *Rheumatology* online) shows the Spearman correlations of all three factor scores and their discordance scores at every time point. The correlations indicate a strong relationship between the discordance scores at an early stage and both factor scores and the discordance score at a later time point.

Furthermore, both the patient-reported factor and the discordance score were predictive of sustained remission from as early as baseline (Table 2) even after correcting for age, gender and treatment group. The odds ratio in the multivariate model for the patient-reported factor score lower than 1, implying that higher values in the PRF score were associated with a lower probability of achieving sustained remission.

Additionally, the discordance score proved to be predictive of RAQoL scores and self-efficacy (ASES) both at week 52 and week 104 (Table 3) indicating the efficacy of the predictive value of the DS for an unrelated outcome measure.

For determining thresholds for the discordance score to classify patients into good/poor health status based on SF-36 (SF-36 score above/below 50), an exploratory ROC curve analysis was performed. The area-under-the-curve for every time point was approximately 0.80. Different cut-offs are needed depending on the time point in the evolution of the disease, one for baseline (right before treatment has started) which was estimated at 0.23 and another for when treatment has taken effect (week 16, 52, 104) of around 0.10 to 0.15 (Figure 2).

DISCUSSION

In this paper, we presented the discordance score as a novel indicator of discrepancy between patient-reported outcomes and more traditional measures of disease activity. We previously identified three factors representing the broader impact of RA using exploratory factor analysis: a patient-reported factor (patient's health assessment, pain, fatigue and HAQ), a clinical factor (physician's health assessment, tender/swollen joint count), and a laboratory factor (ESR and CRP). The PRF, CF and LF scores improved rapidly over time in a treat-to-target setting in patients with early RA participating in the CareRA trial. However, overall, patient-reported impact (PRF) seemed not to improve to the same extent. In fact, PRF remained in all time points higher than either CF or LF scores, despite being normalised to the same scale. Furthermore, the difference seems to be constant over time after week 8. This means that assessment of the discordance between PRF scores and the other factor scores may (from very early in the treatment process) be used as a **warning system** or perhaps even further developed as a **clinical decision support tool** for the clinician, since it predicts future patient-reported impact. Even more importantly, it could detect at an earlier stage, unmet needs even in patients "under control", which is clinically relevant considering that one in five persistent treatment responders from CareRA reported

unmet needs after 1 year, such as remaining pain and fatigue.(13) The fact that two cut-offs are needed for the discordance score seems to indicate that there is a different way of evaluating additional patients' needs not directly explained by disease activity after recently being diagnosed (baseline) versus when treatment has been started. However, replicating these results in independent samples and establishing external validity is needed. At baseline, it is likely that quite an important part of this need is related to unmeasured (CF/LF) but experienced (PRF) disease activity, while this explanation becomes less plausible during further treatment.

Moreover, the a better (lower) PRF score also has a predictive value in forecasting if a patient will be achieving sustained remission. These findings may be somewhat limited since the analyses were exploratory and not powered for assessing prediction models based on the factor and discordance scores. Hence, there could be an effect of the LF which could not be captured due to too low a sample size. The probability of achieving a state of disease control decreases from baseline onwards for every point increase in the PRF score and the negative predictive value of the PRF score becomes even more important by week 16, highlighting again the importance of early disease control, also for global disease impact. Keeping an open mind to the alternative that complementary interventions besides DMARDs could help mitigate the "persisting effect of disease" is crucial. This is an important issue in the current(20,21) and future research. Studies in other patient cohorts than the one the scores were developed in are needed for future validation of these factor and discordance scores and their possible cut-offs for further clinical use. Additional qualitative studies are also needed to understand what is behind this discordance between patient-reported outcomes and clinical variables.

Rapid and persistent disease control as well as baseline psychosocial variables, and not so much treatment choice, have been associated with favourable patient-reported health and illness perceptions after 1 year in CareRA. (19) Moreover, treatment response along with patient-reported outcomes have been proven to be early determinants of long-term self-efficacy in an early RA.(22) The early, intensive, treat-to-target approach(3) definitely has its benefits with snowballing consequences that can go further than controlling disease activity.(23) However, there might be space for broadening the target currently used.(3)

CONCLUSION

Patients' unmet needs in terms of pain, fatigue, functionality and overall well-being should be given more attention during follow-up, even in individuals achieving sustained remission. Looking at the discordance between the patient-reported factor score and the clinical and laboratory factor scores does provide further insights into these needs that cannot be directly explained by standard disease activity parameters. These needs could also represent a target for additional types of care, allowing to broaden the future scope of the treat-to-target principle to multidisciplinary interventions in addition to, and sometimes as an alternative for pharmacological treatment adaptations.

Acknowledgements

We would like to show our gratitude to all participating patients, as well as to the investigators and medical staff at all sites. We appreciate the time invested.

Contributors

PV, RW, AB, SP, and AL made substantial contributions to the conception or design of the study. SP and AL performed the statistical analysis. The manuscript was written by SP, AL, PV, RW, and

AB and subsequently revised critically by all the remaining co-authors. All authors were involved in data interpretation and approved the final version to be submitted for publication.

Patient involvement: The pragmatic CareRA protocol was strongly inspired by daily interactions of the investigators with RA patients in daily clinical practice. Patients were not formally involved in setting the research question or the outcome measures, nor were they invited to comment on study design or the interpretation of the results of this manuscript. However, results of this research will be disseminated to study participants, all stakeholders and the general public in collaboration with patient organisations and the Belgian patient partners program (trained patients who educate physicians, medical students and other health care professionals in collaboration with a rheumatologist).

Funding: The CareRA trial (EudraCT number: 2008-007225-39) was funded by a Flemish governmental grant (Agency for Innovation by Science and Technology [IWT]). Patrick Verschueren holds the Pfizer chair for early rheumatoid arthritis management at the KU Leuven.

Competing interests: The authors have declared no conflicts of interest.

Data availability statement: The authors commit to making the relevant anonymised patient data available for a specified purpose approved by the institution and the principal investigator of the CareRA study and with a signed data access agreement.

REFERENCES

1. Scott DL, Wolfe F, Huizinga TWJ. Rheumatoid arthritis. In: *The Lancet*. 2010. p. 1094–1108. https://doi.org/10.1016/S0140-6736(10)60826-4.

2. Van Riel PLCM. *The development of the disease activity score (DAS) and the disease activity score using 28 joint counts (DAS28)*. Clin Exp Rheumatol. 2014.

3. Smolen JS, Landewé RBM, Bijlsma JWJ, Burmester GR, Dougados M, Kerschbaumer A, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. *Annals of the Rheumatic Diseases*. 2020;79: 685–699. https://doi.org/10.1136/annrheumdis-2019-216655.

4. Smolen JS, Aletaha D, Bijlsma JWJ, Breedveld FC, Boumpas D, Burmester G, et al. Treating rheumatoid arthritis to target: Recommendations of an international task force. *Annals of the Rheumatic Diseases*. 2010;69(4): 631–637. https://doi.org/10.1136/ard.2009.123919.

5. Ferreira RJ, Landewé RB, da Silva JA. Definition of Treatment Targets in Rheumatoid Arthritis: Is It Time for Reappraisal? *The Journal of Rheumatology*. 2021; https://doi.org/10.3899/jrheum.210050.

6. Ferreira RJ de O, Dougados M, Kirwan JR, Duarte C, de Wit M, Soubrier M, et al. Drivers of patient global assessment in patients with rheumatoid arthritis who are close to remission: An analysis of 1588 patients. *Rheumatology (United Kingdom)*. 2017;56(9): 1573–1578. https://doi.org/10.1093/rheumatology/kex211.

7. Ferreira RJO, Carvalho PD, Ndosi M, Duarte C, Chopra A, Murphy E, et al. Impact of Patient's Global Assessment on Achieving Remission in Patients With Rheumatoid Arthritis: A Multinational Study Using the METEOR Database. *Arthritis Care and Research*. 2019;71(10): 1317–1325. https://doi.org/10.1002/acr.23866.

8. Ferreira RJO, Duarte C, Ndosi M, de Wit M, Gossec L, da Silva JAP. Suppressing Inflammation in Rheumatoid Arthritis: Does Patient Global Assessment Blur the Target? A Practice-Based Call for a Paradigm Change. *Arthritis Care and Research*. 2018;70(3): 369–378. https://doi.org/10.1002/acr.23284.

9. Radner H, Yoshida K, Tedeschi S, Studenic P, Frits M, Iannaccone C, et al. Different Rating of Global Rheumatoid Arthritis Disease Activity in Rheumatoid Arthritis Patients With Multiple Morbidities. *Arthritis and Rheumatology*. 2017;69(4): 720–727. https://doi.org/10.1002/art.39988.

10. Pazmino S, Lovik A, Boonen A, de Cock D, Stouten V, Joly J, et al. Does Including Pain, Fatigue, and Physical Function When Assessing Patients with Early Rheumatoid Arthritis Provide a Comprehensive Picture of Disease Burden? *The Journal of Rheumatology*. 2020;49(1). https://doi.org/10.3899/jrheum.200758.

11. Pazmino S, Lovik A, Westhovens R, Verschueren P. Evaluation of Disease Burden by (Separate) Patient-reported, Clinical, and Laboratory Factor Scores in Patients With Established Rheumatoid Arthritis: A Factor Analysis Replication. *The Journal of Rheumatology*. 2022; jrheum.210871. https://doi.org/10.3899/JRHEUM.210871.

12. Stouten V, Westhovens RR, Pazmino S, De Cock D, Van Der Elst K, Joly J, et al. Effectiveness of different combinations of DMARDs and glucocorticoid bridging in early rheumatoid arthritis: two-

year results of CareRA. *Rheumatology (Oxford, England)*. 2019; https://doi.org/10.1093/rheumatology/kez213.

13. Van Der Elst K, Verschueren P, De Cock D, De Groef A, Stouten V, Pazmino S, et al. One in five patients with rapidly and persistently controlled early rheumatoid arthritis report poor well-being after 1 year of treatment. *RMD Open*. 2020;6(1). https://doi.org/10.1136/rmdopen-2019-001146.

14. Ware J. . SF-36 Health Survey Update. *Spine*. 2000;25(24): 3130–3139.

15. Whalley D, Mckenna SP, de Jong Z, van der Heijde D. *Quality of life in Rheumatoid Arthritis*. British Journal of Rheumatology. 1997.

16. Barlow JH, Williams B, Wright CC. The reliability and validity of the arthritis self-efficacy scale in a UK context. *http://dx.doi.org/10.1080/13548509708400556*. 2007;2(1): 3–17. https://doi.org/10.1080/13548509708400556.

17. Burgette LF, Reiter JP. Multiple imputation for missing data via sequential regression trees. *American Journal of Epidemiology*. 2010;172(9): 1070–1076. https://doi.org/10.1093/aje/kwq260.

18. Rubin DB, Schenker N. Multiple imputation for interval estimation from simple random samples with ignorable nonresponse. *Journal of the American Statistical Association*. 1986;81(394): 366–374. https://doi.org/10.1080/01621459.1986.10478280.

19. Van der Elst K, Verschueren P, Stouten V, Pazmino S, De Groef A, De Cock D, et al. Patient reported outcome data from the Care in Early Rheumatoid Arthritis trial: Opportunities for broadening the scope of treating to target. *Arthritis Care & Research*. 2019; https://doi.org/10.1002/acr.23900.

20. Doumen M, Westhovens R, Vandeputte M, Melder R van, Elst K van der, Pazmino S, et al. The perception of stakeholders on the applicability of nurse-led clinics in the management of rheumatoid arthritis. *Rheumatology Advances in Practice*. 2021;5(Suppl 2): ii45. https://doi.org/10.1093/RAP/RKAB052.

21. Doumen M, Westhovens R, Pazmino S, Bertrand D, Stouten V, Neys C, et al. The ideal mHealth-application for rheumatoid arthritis: qualitative findings from stakeholder focus groups. *BMC Musculoskeletal Disorders*. 2021;22(1). https://doi.org/10.1186/S12891-021-04624-8.

22. Doumen M, de Cock D, Pazmino S, Bertrand D, Joly J, Westhovens R, et al. Treatment response and several patient-reported outcomes are early determinants of future self-efficacy in rheumatoid arthritis. *Arthritis research & therapy*. 2021;23(1): 269. https://doi.org/10.1186/S13075-021-02651-3.

23. Schoemaker CG, de Wit MP. Treat to target from the patient perspective is bowling for a perfect strike. *Arthritis & Rheumatology*. 2020; art.41461. https://doi.org/10.1002/art.41461.

Figure 1: Factor-scores over time for a. the entire population and b. per sustained remission.

Figure 2: Patient classification into good/poor health status (based on SF-36) using the discordance score.

AUC = area under the curve

Table 1: Demographic and clinical characteristics at baseline per sustained remission achievement (DAS28CRP <2.6 from week 16 to 104)

	CareRA population n=379	Not sustained remission n=252	Sustained remission n=127
Demographic variables			
Age, years	52 (10)	52 (11)	51 (11)
Women, n (%)	262 (69)	174 (69)	88(69)
Smokers, n smoked ever (%)	209 (55)	143 (57)	66 (52)
Current work n (%)	193 (51)	125 (50)	68 (54)
Clinical variables			
Body mass index, kg/m²	26 (5)	27(4)	26 (4)
Symptom duration, weeks; median (IQR)	20 (12-36)	21 (13-42)	17 (10-29)
RF positive, n (%)	252 (67)	169 (67)	83 (65)
Anti-CCP positive, n (%)	249 (66)	169 (67)	80 (63)
Erosive disease, n (%)	97 (26)	65 (26)	32 (25)
DAS28-CRP	4.50 (1.28)	4.69 (1.25)	4.10 (1.25)
CRP mg/L	9.94 (15.31)	11.56 (26.33)	6.72 (15.07)
SJC28 (0-28)	7 (6)	8 (6)	6 (5)
TJC28 (0-28)	9 (7)	10 (6)	7 (6)
PhGH, mm (0-100)	52 (22)	54 (19)	48 (19)
PaGH, mm (0-100)	55 (27)	58 (23)	49 (24)
Pain, mm (0-100)	56 (24)	60 (24)	49 (24)
Fatigue, mm (0-100)	48 (28)	52 (23)	40 (24)
HAQ score (0-3)	1.03 (0.71)	1.14 (0.71)	0.81 (0.66)
RAQoL (0-30)	11 (8)	13 (7)	9 (7)

Data are presented as mean and standard deviation (SD) unless otherwise specified.
Symptom duration= number of weeks between the onset of symptoms and the start of treatment; RF= Rheumatoid factor; Anti-CCP= Anti cyclic citrullinated protein; DAS28= Disease activity score on 28 joints; CRP= C-reactive protein; SJC28= swollen joint count in 28 joints; TJC28= tender joint count in 28 joints; PhGH= physician's global health assessment; PaGH= patient's global health assessment; HAQ= Health Assessment Questionnaire; RAQoL = rheumatoid arthritis quality of life questionnaire.

Table 2: Prediction of sustained remission (DAS28CRP<2.6 from week 16 to 104) using previous DS.

Logistic regression predicting sustained remission based on	Baseline		Week 8		Week 16	
	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value
Multivariate models adjusted for gender, age, and treatment group						
Patient factor	0.64 (0.48-0.85)	0.002	0.50 (0.37-0.68)	<0.0001	0.42 (0.32-0.56)	<0.0001
Clinical factor	0.79 (0.57-1.12)	0.18	0.46 (0.28-0.76)	0.003	0.46 (0.26-0.83)	0.01
Laboratory factor	0.96 (0.61-1.50)	0.85	0.75 (0.32-1.77)	0.51	0.57 (0.25-1.29)	0.18
Discordance score	0.67 (0.50-0.89)	0.006	0.42 (0.31-0.57)	<0.0001	0.38 (0.28-0.50)	<0.0001

OR= odds ratio, CI= confidence interval

Table 3: Prediction of RA-related quality of life and self-efficacy based on previous discordance score.

Linear regression predicting RA-QoL at	Based on discordance score at week 16, corrected for age, gender, treatment group, and sustained remission.			
	Intercept (SE)	Beta (SE) for DS	p-value for DS	Adjusted R square (95% CI)
RAQoL (0-30)				
Week 52	3.15 (1.49)	20.76 (1.78)	<0.0001	0.29 (0.21-0.37)
Week 104	2.17 (1.47)	21.38 (1.74)	<0.0001	0.31 (0.23-0.40)
ASES				
Week 52				
Total (2-20)	7.90 (0.41)	-3.48 (0.53)	<0.0001	0.13 (0.06-0.21)
Pain (1-10)	3.94 (0.26)	-1.59 (0.34)	<0.0001	0.07 (0.02-0.14)
Other symptoms (1-10)	3.97 (0.23)	-1.88 (0.28)	<0.0001	0.14 (0.07-0.23)
Week 104				
Total (2-20)	7.50 (0.43)	-3.81 (0.50)	<0.0001	0.17 (0.10-0.25)
Pain (1-10)	3.59 (0.28)	-1.88 (0.33)	<0.0001	0.11 (0.05-0.19)
Other symptoms (1-10)	3.92 (0.24)	-1.94 (0.28)	<0.0001	0.16 (0.09-0.25)

SE= standard error, CI= confidence interval, DS= discordance score, RAQoL= rheumatoid arthritis quality of life questionnaire, ASES= Arthritis Self-Efficacy Scale

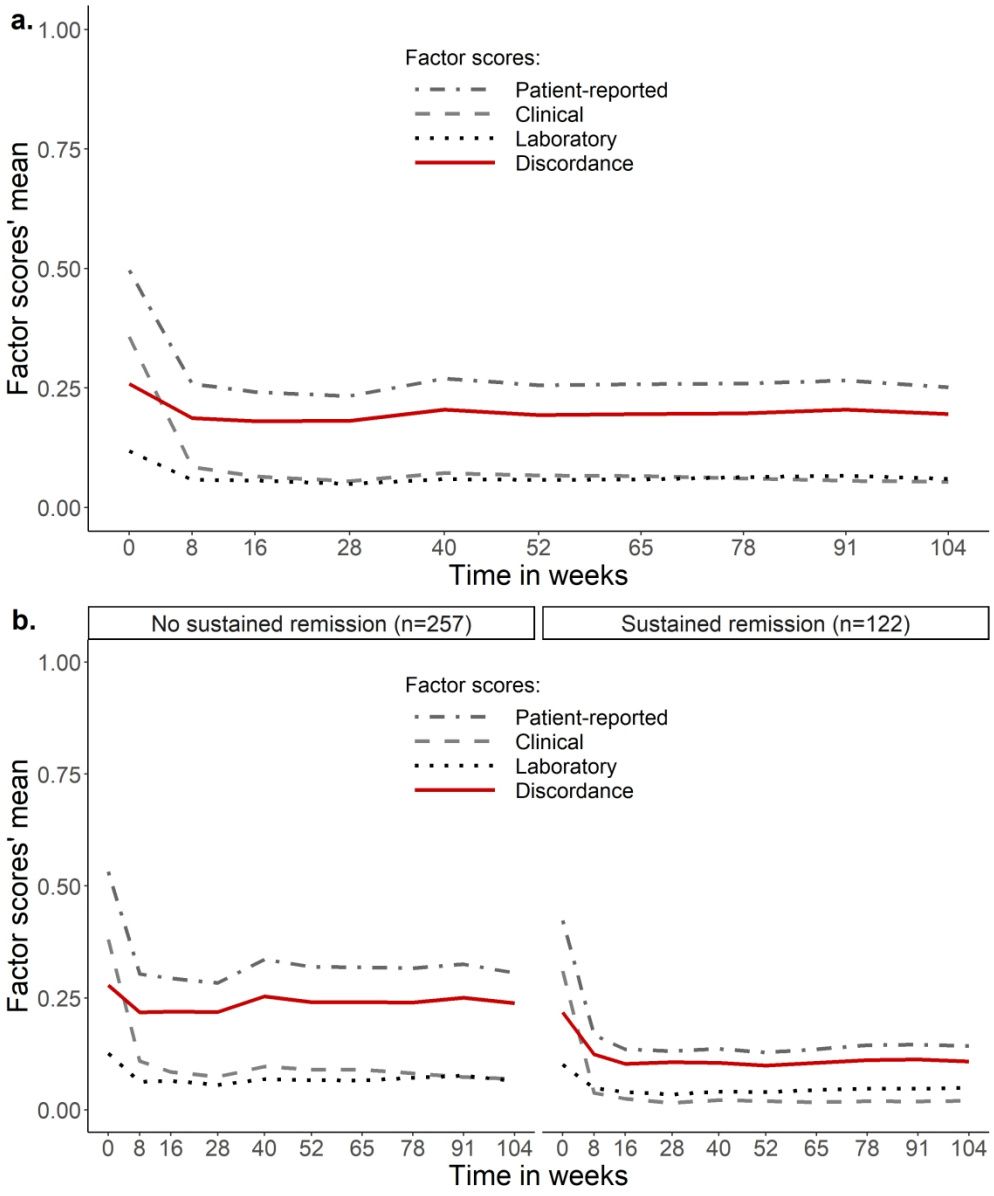

Figure 1

635x762mm (118 x 118 DPI)

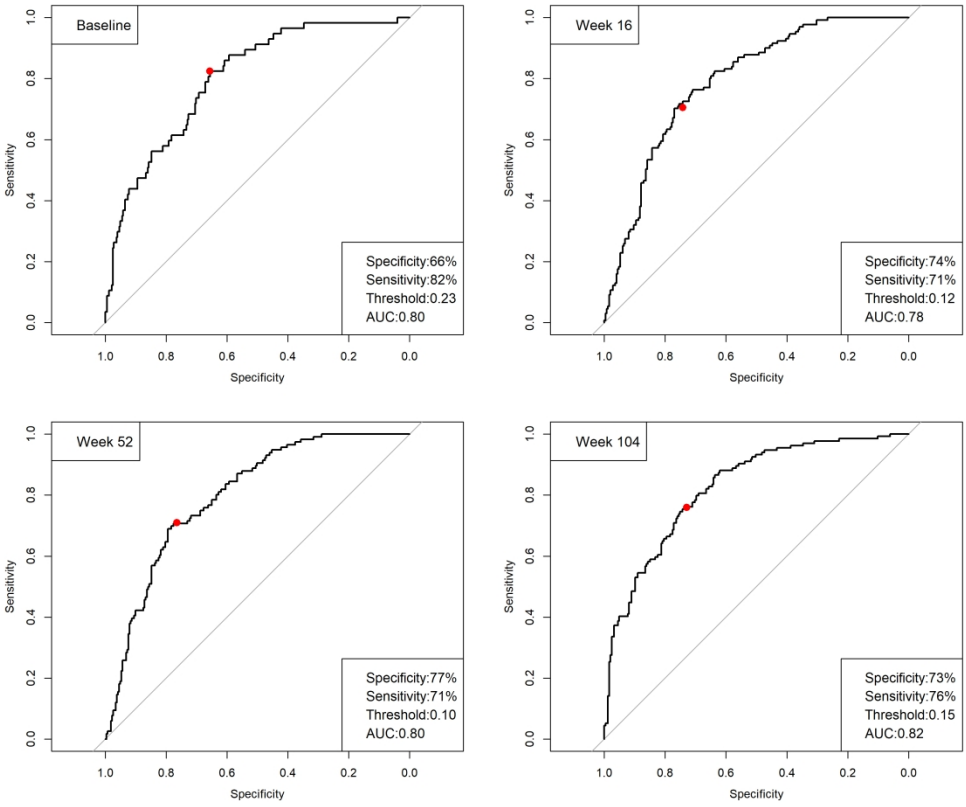

Figure 2

381x317mm (236 x 236 DPI)