Citation	<p>Van der Elst K, Verschueren P, Stouten V, Pazmino S, De Groef A, De Cock D, Joly J, Moons P, Westhovens R</p> <p>Patient-Reported Outcome Data From an Early Rheumatoid Arthritis Trial: Opportunities for Broadening the Scope of Treating to Target.</p> <p>Arthritis Care Res (Hoboken). 2019 Dec;71(12):1566-1575</p>
Archived version	Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher
Published version	https://doi.org/10.1002/acr.23900
Journal homepage	https://onlinelibrary.wiley.com/journal/21514658
Author contact	<p>your email Philip.Moons@med.kuleuven.be</p> <p>your phone number + 32 (0)16 373315</p>
IR	

(article begins on next page)

Patient reported outcome data from the Care in Early Rheumatoid Arthritis trial:

Opportunities for broadening the scope of treating to target

Kristien Van der Elst, PhD^{1,2}, Patrick Verschueren, PhD^{*1,2}, Veerle Stouten, MSc², Sofia Pazmino, MSc², An De Groef, PhD^{3,4}, Diederik De Cock, PhD², Johan Joly, MSc¹, Philip Moons, PhD^{5,6,7}, René Westhovens, PhD^{1,2}

¹University Hospitals Leuven, Department of Rheumatology; ²KU Leuven - University of Leuven, Department of Development and Regeneration, Skeletal Biology and Engineering Research Center; ³Patient research partner; ⁴KU Leuven – University of Leuven, Department of Rehabilitation Sciences and University Hospitals Leuven, Department of Physical Medicine and Rehabilitation; ⁵KU Leuven - University of Leuven, Department of Public Health and Primary Care, Academic Center for Nursing and Midwifery, Leuven, Belgium; ⁶University of Gothenburg, Institute of Health and Care Sciences, Gothenburg, Sweden; ⁷University of Cape Town, Department of Paediatrics and Child Health, Cape Town, South Africa

***Corresponding author:** Patrick Verschueren. Mailing address: Department of Rheumatology, University Hospitals Leuven, Herestraat 49, B-3000 Leuven, Belgium. E-mail: patrick.verschueren@uzleuven.be; Tel: +32-16-342541; Fax: +32-16-342543

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.

Word count: Abstract: 250 (maximum: 250), Total: 3549 (maximum: 3800)

ABSTRACT

Objective: Treating early, intensively and to target leads to rapid disease control, preventing joint damage and loss of function in early rheumatoid arthritis (RA). We report the effect of such approach on patient-reported outcomes and explore the contribution of rapid and persistent disease control on wellbeing after 1 year of treatment.

Methods: This study is part of the Care in early RA trial, a prospective, 2-year, investigator-initiated, randomized controlled trial rooted in daily practice and implementing the treat-to-target principle. SF-36 and IPQ-R data were collected prospectively. We defined 4 clinical response profiles based on speed and consistency of the treatment response within the first year, defined as DAS28CRP<2.6. Linear regression analyses including these response profiles and treatment type were constructed to predict SF-36 dimensions vitality, social functioning, role emotional, mental health, and IPQ-R illness perception subscales consequences, treatment control and illness coherence at year 1.

Results: 333 patients were available for the main analyses, including 140 early persistent responders. Variation in each of the psychosocial outcomes at year 1 was explained mostly by their baseline values, followed by the clinical response profiles. Patients with an early persistent response reported significantly higher vitality, more positive beliefs about disease consequences and treatment effect. Treatment type did not matter.

Conclusion: Rapid and persistent disease control and not treatment type was associated with favorable patient reported health and illness perceptions at year 1, but baseline psychosocial variables mattered most. Our data indicate opportunities to broaden the scope of the treat-to-target principle in early RA.

SIGNIFICANCE AND INNOVATIONS

- It is well-established that time is a valuable thing in rheumatoid arthritis (RA) management, at least in view of the classical outcomes of disease.
- The present study demonstrated that a rapid and persistent initial clinical response is important also from the patient perspective and for patient valued outcomes, broadening the applicability of the window of opportunity theory in early RA.
- Ultimately, psychosocial patient reported factors were even more related to the overall outcome in an early, intensive and treat-to-target setting, while treatment regimen was not.
- Our findings could inspire initiatives to broaden the scope of treat-to-target strategies in early RA by integrating the patient's psychosocial profile.

Rheumatoid arthritis (RA) is a chronic inflammatory disease, traditionally described as progressive and very disabling. This view of an inevitably unfavorable prognosis continues to shape its public perception (1). Today, the future for RA patients is looking much brighter. Aside from the growing availability of effective therapeutic agents, 3 treatment principles shaped the transformation towards better clinical outcomes: treating early, intensively and to target (2-4). In view of the window of opportunity theory, the goal of early RA treatment is to reach remission as soon as possible, since this initial treatment response is a strong predictor of long-term remission, radiographic damage and functional capacity (5, 6). Pragmatic early RA trials demonstrated that applying current treatment principles in a daily practice setting is effective in terms of these key outcomes, including levels of DAS28CRP <2.6 (28-joint disease activity score using C-reactive protein level) in up to 70% of patients (7).

Despite good disease control at the group level, not all patients with early RA have a sufficient treatment response or enjoy a smooth trajectory towards response, while they prefer to return to normal as quickly as possible (8). Few studies linked early clinical response to outcomes valued by patients (9, 10), and showed that patient-reported outcomes (PROs) of patients in early remission reached values close to normal (11). However, the great heterogeneity in response to treatment, both in terms of clinical outcomes and in how patients perceive their disease and health, cannot be captured at the group level. This diversity among patients entails the need for an alternative way of using PROs, if we want to study questions that are meaningful to patients with early RA (12).

Newly diagnosed patients have preconceived illness beliefs and are not familiar with their disease and treatment yet. The initial clinical response might to a large extent determine how patients adapt to their new illness. For example, patients with an

insufficient or delayed clinical response can make at some stage the switch from believing the disease is reversible and manageable, to believing it will be irreversible. Illness perceptions have been shown to be barriers or facilitators of adjustment to illness, and therefore can be determinants of heterogeneity in functioning and health (13-15). Previous research showed that patient personalities, different stressors, and social support at the early disease stage contributed to future anxiety and depressed mood in early RA (16).

Making use of the Care in Early RA (CareRA) trial, a pragmatic randomized study implementing current RA treatment recommendations (7, 17), we compared the effect of different treatment regimens on PROs. However, the main aim of this study was to explore the relative contributions of the different treatment combinations studied, the initial clinical response and specific baseline psychosocial characteristics on patient reported health and illness perceptions 1 year after treatment initiation for early RA. We assumed that a rapid and stable clinical response would improve future aspects of psychosocial functioning and explored the contributions of personal psychosocial factors on their outcome.

PATIENTS and METHODS

Setting

The present observational study is an integral part of the CareRA trial, a Flemish, multicenter, prospective, 2-year, investigator-initiated, randomized controlled pragmatic trial that implements an early intensive treatment in a treat-to-target approach in all patients with early RA (7, 17). Using a risk stratification algorithm based on classical prognostics factors (i.e., erosions, rheumatoid factor and/or anticitrullinated protein antibody and DAS28CRP calculated at screening), patients

were randomized to one of the following treatment groups: COBRA Classic: 15 mg methotrexate (MTX) weekly, 2 g sulfasalazine daily and a weekly step-down scheme of oral prednisone (60-40-25-20-15-10-7.5 mg QD); COBRA Slim: 15 mg MTX weekly with a weekly step-down scheme of oral prednisone (30-20-12.5-10-7.5-5 mg QD); COBRA Avant-garde: 15 mg MTX weekly, 10 mg leflunomide daily and a weekly step-down scheme of oral prednisone (30-20-12.5-10-7.5-5 mg QD); from week 6 or 7 onwards, a low maintenance dose of prednisone was continued to week 28 and then tapered until discontinuation at week 34; Tight Step-up: 15 mg MTX weekly, no oral glucocorticoids allowed. For details see Figure 1. Treatment was adjusted to a target of low disease activity (DAS28CRP ≤ 3.2). The COBRA Slim treatment including monotherapy with methotrexate and temporary glucocorticoids provided the best balance between effectivity, safety and feasibility compared to the other treatment arms in this study setting. This strategy was meanwhile explicitly adopted in The European League Against Rheumatism 2016 recommendations for treating RA (2).

Alongside the traditional disease evaluation measures, implementation aspects were studied (18-20), and patients completed a set of self-reported questionnaires (Figure 1). Here, we focused on data from 2 questionnaires: 36-item Short Form Health Survey (SF-36) and the Revised Illness Perception Questionnaire (IPQ-R).

The CareRA trial was approved by the leading ethics committee (University Hospitals Leuven) after consulting the ethics committees of participating centers. All participants gave written informed consent.

The initial clinical response profiles

The disease activity trajectory in the first year of treatment was determined based on achieving DAS28CRP < 2.6 at week (W) 16 or not, and persistence of this early

response until W52. Four profiles of initial clinical response were defined: 1) the 'persistent responders' (n=140) had an early clinical response at W16, which was sustained at every clinical visit between W16 and W52; 2) the 'secondary failures' (n=92) had an early clinical response, but lost this response between W16 and W52; 3) the 'delayed responders' (n=71) had a late clinical response, which was obtained between W16 and W52; and 4) the 'non-responders' (n=30) had never a clinical response throughout the first treatment year. We selected W16 to evaluate the speed of response, to have disease activity not being blurred by the effect of the initial higher dosages of glucocorticoids and because it is known that the level of disease activity at this time point is predictive for response to treatment after 1 year (9, 21).

Measures/instruments

Descriptive variables included disease activity (DAS28CRP), allocated treatment (Figure 1), presence of depression (i.e., having a diagnosis of depression and/or taking anti-depressants at baseline), age, gender, body mass index (BMI) at baseline, alcohol intake, smoking and employment status at screening, and perceived social support at baseline using the subscales *supportive interactions* and *negative interactions* of the Social Support List, in which higher scores indicate more experience of supportive or negative interactions (22). The PROs of interest (i.e., perceived health and illness perceptions) were collected through valid and reliable Dutch translations of SF-36 (version 1) and IPQ-R (23, 24). Questionnaires were paper-based and completed by patients at home. For the main research question that focused on psychosocial aspects, we selected the corresponding SF-36 subscales of *vitality*, *social functioning*, *role emotional* and *mental health* (25). Higher SF-36 scores indicate better perceived health (26). Furthermore, 3 IPQ-R illness perception subscales were selected based

on their relevance for our research question: *consequences, treatment control and illness coherence*. Higher scores indicate greater perceived consequences, and a stronger sense of treatment control and illness coherence (27).

Data analysis

The role of treatment

This analysis included SF-36 reporting of the total CareRA sample according to randomized treatment groups. Mean scores for all 8 dimensions of the SF-36 at baseline, W16, W52 and W104 were plotted on spidergrams for each treatment group. Differences in improvement from baseline at W16, W52 and W104 were tested separately for the 3 high-risk treatment groups using Kruskal-Wallis test, and for the 2 low-risk treatment groups using Mann-Whitney U test.

The role of initial clinical response profiles

Sample characteristics and baseline differences between initial clinical response profiles

Patients who were allocated to a profile, completed the first treatment year, and had a DAS28CRP available at W16 and at least once between W16 and W52 (W52 included). For the purpose of this analysis, SF-36 and IPQ-R had to be completed minimally on 1/4 measurement moments allowing eventual missing data imputation. Descriptive data were reported as absolute numbers and proportions, or as means with standard deviations (SD) and medians with interquartile range (IQR). Patients were not randomized but assigned to 1 of 4 profiles based on their initial clinical response, and therefore we tested baseline differences between profiles on all study

variables using 1-way analysis of variance or Kruskal-Wallis tests according to data distribution for the continuous variables, and χ^2 tests for categorical variables.

The association between the initial clinical response and psychosocial outcomes after 1 year of treatment

We performed linear regression analyses to investigate the association between the initial clinical response (predictor variable) and the predefined set of SF-36 and IPQ-R subscales (7 outcome variables). As the clinical response profile is a categorical variable, dummy coding was applied where the persistent responders were chosen as reference category against which the secondary failures, delayed responders and non-responders were compared (according to our hypothesis). In the multiple regression analyses, the effect of initial clinical response on W52 psychosocial outcomes was adjusted for age, sex, presence of depression, allocated treatment, baseline DAS28CRP and baseline value of the predicted variable. To deal with not normally distributed data, especially for SF-36 social functioning and SF-36 role emotional, robust regression based on bootstrapping with 1000 bias-corrected and accelerated replications was applied, providing bootstrap confidence intervals and significance values that give an accurate estimate of the true population beta values (28, 29). Evaluation of PROs after 1 year of treatment was chosen as the primary outcome cut point.

Missing data

To deal with missing data, we first applied the method recommended in the SF-36 users' manual (26), and in the Dutch IPQ-R scoring syntax before calculating the subscale scores. If SF-36 and IPQ-R subscale scores could not be calculated because

of too many missing items, we used the Expectation-Maximization algorithm (EM) (30). In our study sample, the proportion missingness for the 7 selected SF-36 and IPQ-R subscale scores ranged between 9.3–10.2% at W16, 22.2–22.8% at W52 and 30.3–31.2% at W104. These missing scores were imputed using 1 EM imputation model, including available scores of all SF-36 and IPQ-R subscales at baseline, W16, W52 and W104 for the total CareRA sample. No imputations were made for missing scores at baseline (1.8–2.4%). To examine baseline differences between participants with and without the studied PRO data available at W52, Mann-Whitney U tests were computed. Sensitivity analysis using SF-36 and IPQ-R scores without EM imputation were performed for the regression analyses.

A P-value of 0.05 was used as cut-off for statistical significance, with p-values being corrected per regression model using the Holm-Bonferroni method to reduce the risk for inflating type 1 error (31). Data were analyzed using IBM SPSS Statistics for Windows, version 25 (Armonk, NY: IBM Corp.). There was no specific power calculation for this explorative study alongside the CareRA trial.

Patient involvement

Our interdisciplinary research team was strengthened by a patient research partner (AG), involved in study design, data interpretation and writing-up of the results.

RESULTS

The role of treatment

Patient characteristics of all CareRA patients have been previously reported (17). All treatments resulted in improved SF-36 domain scores in the high-risk treatment groups

(Figure 2) and in the low-risk treatment groups (Figure 3). Change from baseline to W16, W52 and W104 did not significantly differ between the 3 high-risk treatment groups, nor between the 2 low-risk treatment groups for any of the 8 SF-36 domain scores (data on file).

The role of initial clinical response profiles

Sample characteristics and baseline differences between initial clinical response profiles

Of the 379 randomized patients in the CareRA trial, 333 patients (87.9%) had the required dataset for this particular analysis available (Supplementary file 1). The selected patients were not significantly different from those excluded (data on file). Of those 333, 259 patients had SF-36 and IPQ-R data available at W52 (77.8%). Patients for which these PRO data were missing at W52 only scored on average 5.5 units less on mental health at baseline than those with W52 PRO data available ($p=0.042$).

Table 1 presents the baseline characteristics and Table 2 the PRO data organized by initial clinical response profile. As patients were not randomized but subdivided post hoc according to their initial clinical response, baseline differences were observed for BMI, patient reported vitality, social functioning, mental health, perceived consequences and negative social interactions, as well as for DAS28CRP. Separate analysis of the 4 DAS28CRP components revealed that items based on patient reporting and patient-physician interaction were different between profiles, while the laboratory measure CRP was not.

The association between the initial clinical response and psychosocial outcomes after 1 year of treatment

Table 2 displays the average W52 scores for perceived health and illness perceptions by initial clinical response profile. Firstly, the univariate regression results showed that the proportion of the variability explained by the response profiles ranged between 6.3%–12.6% for perceived health outcomes and between 3.1%–12.7% for illness perception outcomes (Supplementary file 2). Secondly, in almost all multiple regression models were the response profiles independently associated with the psychosocial outcomes at year 1 of treatment (Figure 4; for details see Supplementary file 3). Only illness coherence remained unexplained by initial clinical response. Depending on the model, contributing confounders were the baseline value of the outcome variable, baseline disease activity and presence of depression. The baseline values had the highest contribution to the overall variability of the W52 outcome variable, which was in general followed by the response profiles (see standardized beta values in Figure 4). The allocated treatment did not contribute at all, nor did sex and age. Alcohol intake, smoking and employment status and BMI were evaluated as potential confounders, but either were not different at baseline between response groups or did not significantly contribute to the multivariate models we tested.

Figure 4 shows that compared to early and persistent responders, patients without response scored mostly worst, while the difference in group means often was smallest between the persistent responders and the secondary failures. Patients who enjoyed an early and persistent clinical response reported higher vitality, more positive beliefs about disease consequences and a higher belief in the effect of treatment at W52, as compared to patients who had a relapse after initial response or had a delayed response or no response at all. Moreover, compared to persistent responders, delayed responders and non-responders reported more interference with social activities and more problems with work or other daily activities because of emotional problems.

Patients with a relapse or no response, scored less well on mental health at W52 compared to persistent responders.

Sensitivity analysis using the W52 PRO scores without EM imputation showed a similar trend in effect sizes and similar study conclusions (Supplementary file 4).

DISCUSSION

As hypothesized, speed and stability of the initial clinical response was independently associated with aspects of psychosocial functioning 1 year after RA treatment initiation. Also baseline psychosocial characteristics predicted future psychosocial outcomes of patients with early RA, while allocated treatment did not contribute to the variability in patient reported health and illness perceptions. Evolution of SF-36 scores in the total CareRA patient population did not differ between treatment groups as one could expect based on the similar effect on disease activity in this treat-to-target trial.

It is well-established that there is no time to waste in early RA management. Remarkably, limited attention has been given to what the early clinical response profile could mean for long-term outcomes other than disease activity, joint damage and functionality. Preliminary evidence in RA showed that time could also be of essence from the patient's point of view (9-11, 32). Our study adds to this evidence by sub-defining patients with early RA according to speed and persistency of their clinical response. This methodology allowed us to conclude that patients with a rapid and persistent initial clinical response scored better on aspects of psychosocial functioning after 1 year of treatment. Hence, from the patient perspective reaching good disease control is important but also the clinical trajectory leading to this target. In a previous longitudinal qualitative study, we already demonstrated that a rapid and sustained response was highly valued by patients (8). Similarly, 'feeling good as soon as possible

for as long as possible' was highlighted in the report of the 10th Outcome Measures in Rheumatology (OMERACT) about response criteria (33). Although it might seem self-evident that patients with a favorable clinical response profile would also report a beneficial effect for outcomes that matter most to them, also in this group unmet needs may exist requiring further study (34).

Already at baseline patients differed for DAS28 components other than CRP depending on their later clinical response profiles, suggesting disease activity scoring is influenced by a patient-specific reporting style. Recently, Michelsen and colleagues concluded that psychosocial factors (i.e., depression and anxiety) were associated to more subjectively weighted measures of disease activity and not acute phase reactants and swollen joint count (35), which was confirmed in an early RA cohort by Boer et al (36). Both studies identified baseline psychosocial distress as a strong negative predictor of remission in early RA. This fits with our results, confirming baseline levels as the most powerful predictors of the explored psychosocial outcomes, as previously shown (37, 38). Nonetheless, in contrast to these studies, in daily practice psychosocial variables almost never get the necessary attention in management of disease. Unfortunately, even compliance with current guidelines on pharmacological treatment in RA and the treat-to-target principles is challenging in clinical practice (39).

There is growing awareness that RA pharmacotherapy alone may not be the solution for relevant health outcomes like psychosocial wellbeing (40). Following the definition of precision medicine, treatment should be targeted to a broad range of individual needs, based on genetic, biomarker, phenotypic and psychosocial characteristics (41). Nevertheless, actual investments are mainly aiming for the development of new diagnostics, prognostics and therapeutics. Our findings suggest

that to benefit from the window of opportunity in all its aspects, timely patient education and counseling might be important as well as personalized support for patients at risk of a more difficult RA trajectory due to personality-related and social aspects. It will however be a challenge to translate this into the therapeutic setting. The response rate in this pragmatic trial illustrates the difficulty of evaluating psychosocial status by means of questionnaires and a more multidimensional evaluation of newly diagnosed patients, including the chosen measures reported here, can probably only be realized by interdisciplinary teams working towards timely implementation of a wider scope of interventions tailored to patients' needs.

Our results were based on a DAS28CRP <2.6 and could have been different when using other definitions of disease control. However, applying DAS28CRP <2.6 resulted in a distribution with good face validity and sufficient patient numbers per response profile. Moreover, in the CareRA trial the cut-off of DAS28CRP ≤ 3.2 was used for therapeutic steering, which results in a better effectiveness of the treat-to-target strategy (42), and might correspond more closely to what patients consider to be acceptable. Our study included only particular aspects of the broad concept of psychosocial functioning. For example, work participation, which was added to the most recent treat-to-target recommendations (43), could not reliably be extracted from our data. Because data were collected in a setting mimicking clinical practice, no reminders were sent to encourage patients to return their questionnaires. As in similar studies we encountered missing PRO data, which might hamper the generalizability of our results. On the other hand, PROs were studied within a pragmatic trial setting close to clinical practice, making study conclusions generally applicable for current disease management. Future studies on the effect of early and sustained disease control on patients' future psychosocial functioning could include even more long-term time-

points while considering integrating 'planned missingness' in the study design (44), or perhaps computer adaptive testing to improve data completion (45). Prevalence of depression was low in our early RA sample. There is a risk we underestimated depression as potential confounder since we did not formally evaluate patients but used classification criteria with a high specificity but low sensitivity for the diagnosis of depression. We also recognize the potential relevance to our psychosocial outcomes of measures we did not include as confounders, such as socioeconomic status and social support (46, 47). However, our study was conducted in Flanders, a region where standards of life are generally good with a low unemployment rate and only minor socioeconomic inequality. Moreover, the Belgian healthcare system is easily accessible and includes a comprehensive social security. Eventually, perceived social support was not added as confounder to the regression models because of the small numerical difference between patients with different clinical response profiles and the lack of evidence about its clinical meaningfulness. A methodological strength of this study is that we analyzed PRO data by defining clinically relevant response subgroups. For example, our data revealed that suffering a relapse after a favorable initial response can have implications for future psychosocial functioning and is important to consider in early disease management. Moreover, the patient was at the heart of this study, together with a patient researcher, a condition of clinically useful research (12, 48).

In conclusion, this study provides initial evidence in favor of a psychosocial window of opportunity for early RA management. Our data showed that a rapid and sustained response was independently associated with positive aspects of psychosocial wellbeing at year 1 of treatment. Baseline psychosocial status however

contributed most, while treatment type did not. These findings could be a source of inspiration for initiatives to broaden the scope of treat-to-target strategies in early RA.

Acknowledgements

We are very grateful to all patients, who were willing to complete the questionnaires during a crucial period of their disease. Furthermore, we would like to thank the local rheumatology staff for data collection, the CareRA team for data monitoring, and Anikó Lovik for providing statistical advice.

Contributors

The following authors were involved in this study: Kristien Van der Elst (KE), Patrick Verschueren (PV), Veerle Stouten (VS), Sofia Pazmino (SP), An De Groef (AG), Diederik De Cock (DC), Johan Joly (JJ), Philip Moons (PM), René Westhovens (RW). KE, RW, PV, PM, AG, VS and SP made substantial contributions to the conception or design of the study. KV analyzed the data. KV, RW and PV drafted the manuscript. PM, AG, VS, SP, DC and JJ were involved in revising the manuscript critically for important intellectual content.

REFERENCES

1. Simons G, Mason A, Falahee M, Kumar K, Mallen CD, Raza K, et al. Qualitative Exploration of Illness Perceptions of Rheumatoid Arthritis in the General Public. *Musculoskeletal Care*. 2017;15(1):13-22.
2. Smolen JS, Landewe R, Bijlsma J, Burmester G, Chatzidionysiou K, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis*. 2017;76(6):960-77.
3. Combe B, Landewe R, Daien CI, Hua C, Aletaha D, Alvaro-Gracia JM, et al. 2016 update of the EULAR recommendations for the management of early arthritis. *Ann Rheum Dis*. 2017;76(6):948-59.
4. Singh JA, Saag KG, Bridges SL, Jr., Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Arthritis Care Res (Hoboken)*. 2016;68(1):1-25.
5. Landewe RB, Boers M, Verhoeven AC, Westhovens R, van de Laar MA, Markusse HM, et al. COBRA combination therapy in patients with early rheumatoid arthritis: long-term structural benefits of a brief intervention. *Arthritis Rheum*. 2002;46(2):347-56.
6. Markusse IM, Akdemir G, Dirven L, Goekoop-Ruiterman YP, van Groenendael JH, Han KH, et al. Long-Term Outcomes of Patients With Recent-Onset Rheumatoid Arthritis After 10 Years of Tight Controlled Treatment: A Randomized Trial. *Ann Intern Med*. 2016;164(8):523-31.
7. Verschueren P, De Cock D, Corluy L, Joos R, Langenaken C, Taelman V, et al. Effectiveness of methotrexate with step-down glucocorticoid remission induction (COBRA Slim) versus other intensive treatment strategies for early rheumatoid arthritis

in a treat-to-target approach: 1-year results of CareRA, a randomised pragmatic open-label superiority trial. *Ann Rheum Dis*. 2017;76(3):511-20.

8. Van der Elst K, Meyfroidt S, De Cock D, De Groef A, Binnard E, Moons P, et al. Unraveling Patient-Preferred Health and Treatment Outcomes in Early Rheumatoid Arthritis: A Longitudinal Qualitative Study. *Arthritis Care Res (Hoboken)*. 2016;68(9):1278-87.

9. Verschueren P, Esselens G, Westhovens R. Predictors of remission, normalized physical function, and changes in the working situation during follow-up of patients with early rheumatoid arthritis: an observational study. *Scand J Rheumatol*. 2009;38(3):166-72.

10. Kristiansen TM, Primdahl J, Antoft R, Horslev-Petersen K. Everyday life with rheumatoid arthritis and implications for patient education and clinical practice: a focus group study. *Musculoskeletal Care*. 2012;10(1):29-38.

11. Heimans L, Wevers-de Boer KV, Koudijs KK, Visser K, Goekoop-Ruiterman YP, Harbers JB, et al. Health-related quality of life and functional ability in patients with early arthritis during remission steered treatment: results of the IMPROVED study. *Arthritis Res Ther*. 2013;15(5):R173.

12. Ioannidis JP. Why Most Clinical Research Is Not Useful. *PLoS Med*. 2016;13(6):e1002049.

13. Petrie KJ, Jago LA, Devcich DA. The role of illness perceptions in patients with medical conditions. *Curr Opin Psychiatry*. 2007;20(2):163-7.

14. Groarke A, Curtis R, Coughlan R, Gsel A. The role of perceived and actual disease status in adjustment to rheumatoid arthritis. *Rheumatology (Oxford)*. 2004;43(9):1142-9.

15. Karademas EC, Dimitraki G, Papastefanakis E, Ktistaki G, Repa A, Gergianaki I, et al. Adaptation to inflammatory rheumatic disease: Do illness representations predict patients' physical functioning over time? A complex relationship. *J Behav Med.* 2018;41(2):232-42.
16. Evers AW, Kraaimaat FW, Geenen R, Jacobs JW, Bijlsma JW. Longterm predictors of anxiety and depressed mood in early rheumatoid arthritis: a 3 and 5 year followup. *J Rheumatol.* 2002;29(11):2327-36.
17. Verschueren P, De Cock D, Corluy L, Joos R, Langenaken C, Taelman V, et al. Methotrexate in combination with other DMARDs is not superior to methotrexate alone for remission induction with moderate-to-high-dose glucocorticoid bridging in early rheumatoid arthritis after 16 weeks of treatment: the CareRA trial. *Ann Rheum Dis.* 2015;74(1):27-34.
18. Meyfroidt S, van Hulst L, De Cock D, Van der Elst K, Joly J, Westhovens R, et al. Factors influencing the prescription of intensive combination treatment strategies for early rheumatoid arthritis. *Scand J Rheumatol.* 2014;43(4):265-72.
19. Meyfroidt S, Van der Elst K, De Cock D, Joly J, Westhovens R, Hulscher M, et al. Patient experiences with intensive combination-treatment strategies with glucocorticoids for early rheumatoid arthritis. *Patient Educ Couns.* 2015;98(3):384-90.
20. De Cock D, Meyfroidt S, Joly J, Van der Elst K, Westhovens R, Verschueren P, et al. A detailed analysis of treatment delay from the onset of symptoms in early rheumatoid arthritis patients. *Scand J Rheumatol.* 2014;43(1):1-8.
21. Aletaha D, Funovits J, Keystone EC, Smolen JS. Disease activity early in the course of treatment predicts response to therapy after one year in rheumatoid arthritis patients. *Arthritis Rheum.* 2007;56(10):3226-35.

22. van Sonderen E. Het meten van sociale steun met de Sociale Steun Lijst-Interacties (SLL-I) en Sociale Steun Lijst-Discrepanties (SSL-D): een handleiding. Groningen: Noordelijk Centrum voor Gezondheidsvraagstukken, Rijksuniversiteit Groningen; 1993.
23. Heijmans M, Rijken M. Validation of the IPQ-R in a large, representative Dutch sample. Book of abstracts of the 17th Conference of the European Health Psychology Society, 'European health psychology. Gender, culture and health', 24–27 September 2003, Kos, Greece, p.14.
24. Aaronson NK, Muller M, Cohen PD, Essink-Bot ML, Fekkes M, Sanderman R, et al. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. *J Clin Epidemiol*. 1998;51(11):1055-68.
25. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30(6):473-83.
26. Ware JE, Jr., , Kosinski, M., Gandek, B. SF-36® Health Survey: Manual & Interpretation Guide. Lincoln, RI: QualityMetric, 2005, c1993, 2000.
27. Moss-Morris R WJ, Petrie KJ, Horne R, Cameron LD, Buick D. The revised illness perception questionnaire (IPQ-R). *Psychol Health* 2002;17:1–16.
28. Field A. *Discovering Statistics Using IBM SPSS Statistics*. 4th ed. Los Angeles: Sage Publications Ltd 2013:419-426.
29. Walters SJ, Campbell MJ. The use of bootstrap methods for analysing Health-Related Quality of Life outcomes (particularly the SF-36). *Health Qual Life Outcomes*. 2004;2:70.
30. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc* 1977;39:1-38.

31. Aickin M, Gensler H. Adjusting for multiple testing when reporting research results: the Bonferroni vs Holm methods. *Am J Public Health*. 1996;86(5):726-8.
32. Kristiansen TM, Primdahl J, Antoft R, Horslev-Petersen K. It means everything: continuing normality of everyday life for people with rheumatoid arthritis in early remission. *Musculoskeletal Care*. 2012;10(3):162-70.
33. Strand V, Boers M, Idzerda L, Kirwan JR, Kvien TK, Tugwell PS, 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(8):1720-7.
34. Steunebrink LMM, Oude Voshaar MAH, Taal E, Vonkeman HE, Zijlstra TR, van de Laar M. Determinants of Perceived Health Nonimprovement in Early Rheumatoid Arthritis Patients With Favorable Treatment Outcomes. *Arthritis Care Res (Hoboken)*. 2018;70(4):510-5.
35. Michelsen B, Kristianslund EK, Sexton J, Hammer HB, Fagerli KM, Lie E, et al. Do depression and anxiety reduce the likelihood of remission in rheumatoid arthritis and psoriatic arthritis? Data from the prospective multicentre NOR-DMARD study. *Ann Rheum Dis*. 2017;76(11):1906-10.
36. Boer AC, Huizinga TWJ, van der Helm-van Mil AHM. Depression and anxiety associate with less remission after 1 year in rheumatoid arthritis. *Ann Rheum Dis*. 2018.
37. Benka J, Nagyova I, Rosenberger J, Calfova A, Macejova Z, Middel B, et al. Social support and psychological distress in rheumatoid arthritis: a 4-year prospective study. *Disabil Rehabil*. 2012;34(9):754-61.
38. van Lankveld W, Naring G, van't Pad Bosch P, van de Putte L. The negative effect of decreasing the level of activity in coping with pain in rheumatoid arthritis: an

increase in psychological distress and disease impact. *J Behav Med.* 2000;23(4):377-91.

39. Gvozdenovic E, Allaart CF, van der Heijde D, Ferraccioli G, Smolen JS, Huizinga TW, et al. When rheumatologists report that they agree with a guideline, does this mean that they practise the guideline in clinical practice? Results of the International Recommendation Implementation Study (IRIS). *RMD Open.* 2016;2(1):e000221.

40. Matcham F, Galloway J, Hotopf M, Roberts E, Scott IC, Steer S, et al. The impact of targeted Rheumatoid Arthritis pharmacological treatment on mental health: A systematic review and network meta-analysis. *Arthritis Rheumatol.* 2018.

41. Jameson JL, Longo DL. Precision medicine--personalized, problematic, and promising. *N Engl J Med.* 2015;372(23):2229-34.

42. Verschueren P, Westhovens R. The use of glucocorticoids in early rheumatoid arthritis. *Rheumatology (Oxford).* 2018;57(8):1316-7.

43. Smolen JS, Breedveld FC, Burmester GR, Bykerk V, Dougados M, Emery P, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. *Ann Rheum Dis.* 2016;75(1):3-15.

44. Little TD, Jorgensen TD, Lang KM, Moore EW. On the joys of missing data. *J Pediatr Psychol.* 2014;39(2):151-62.

45. Witter JP. The Promise of Patient-Reported Outcomes Measurement Information System-Turning Theory into Reality: A Uniform Approach to Patient-Reported Outcomes Across Rheumatic Diseases. *Rheum Dis Clin North Am.* 2016;42(2):377-94.

46. Puolakka K, Kautiainen H, Mottonen T, Hannonen P, Hakala M, Korpela M, et al. Predictors of productivity loss in early rheumatoid arthritis: a 5 year follow up study. *Ann Rheum Dis*. 2005;64(1):130-3.
47. Evers AW, Kraaimaat FW, Geenen R, Jacobs JW, Bijlsma JW. Pain coping and social support as predictors of long-term functional disability and pain in early rheumatoid arthritis. *Behav Res Ther*. 2003;41(11):1295-310.
48. de Wit MP, Berlo SE, Aanerud GJ, Aletaha D, Bijlsma JW, Croucher L, et al. European League Against Rheumatism recommendations for the inclusion of patient representatives in scientific projects. *Ann Rheum Dis*. 2011;70(5):722-6.

TABLES

Table 1. Descriptive variables at trial baseline for the sample of patients with early rheumatoid arthritis (n=333) subdivided according to their initial clinical response profile

Variables	Persistent responders (n=140)	Secondary failures (n=92)	Delayed responders (n=71)	Non- responders (n=30)	P-value
Age, years	52.7±13.4	52.3±13.1	52.0±11.3	51.7±14.8	0.974 ^a
DAS28CRP (0-9.4)	4.5±1.1	4.6±1.4	5.3±1.3	5.2±1.1	<0.001 ^a
28SJC (0-28)	6.7±4.9	6.7±5.7	8.7±5.6	8.1±5.4	0.029 ^b
28TJC (0-28)	7.5±5.2	8.4±6.4	11.1±6.8	10.2±6.3	0.001 ^b
VAS PGA (0-100)	50.9±23.8	51.9±21.9	64.6±21.0	61.0±21.6	<0.001 ^b
CRP, mg/L	7.7±15.7	11.6±26.0	13.4±28.1	13.5±35.9	0.712 ^b
Symptom duration, weeks	31.2±30.5	42.2±68.3	33.6±34.5	43.6±76.9	0.773 ^b
Allocated treatment					0.061 ^c
COBRA Classic	42 (30.0)	20 (21.7)	13 (18.3)	9 (30.0)	
COBRA Slim (HR)	33 (23.6)	35 (38.0)	15 (21.1)	5 (16.7)	
COBRA Avant- garde	36 (25.7)	22 (23.9)	19 (26.8)	7 (23.3)	
Tight Step-up	12 (8.6)	8 (8.7)	15 (21.1)	6 (20.0)	
COBRA Slim (LR)	17 (12.1)	7 (7.6)	9 (12.7)	3 (10.0)	
Sex, women	95 (67.9)	65 (70.7)	45 (63.4)	23 (76.7)	0.570 ^c
Depression, yes	6 (4.3)	6 (6.5)	5 (7.0)	1 (3.3)	0.755 ^c
Alcohol intake, yes	88 (62.9)	50 (54.3)	41 (57.7)	12 (40.0)	0.123 ^c

Smoking status (ever)	70 (50.0)	58 (63.0)	45 (63.4)	15 (50.0)	0.117 ^c
Employed at screening	68 (48.6)	49 (53.3)	36 (50.7)	16 (53.3)	0.901 ^c
BMI (kg/m ²)	25.6±3.8	27.0±4.6	27.1±4.2	27.3±4.9	0.037 ^b
SSL Supportive interactions (34- 136), n=324)	79.1±15.7	76.4±15.4	79.0±17.3	80.1±18.6	0.858 ^b
SSL Negative interactions (7-28), n=324	9.1±2.9	9.4±3.2	10.9±4.3	9.7±3.5	0.047 ^b

Note. Values reported are mean± standard deviation for continuous variables and absolute numbers (proportions) for categorical variables. P-values were obtained using: ^a 1-way analysis of variance test, ^b Kruskal-Wallis test, ^c χ^2 -statistic. Alcohol intake: consumption of any form of alcohol; BMI: body mass index; DAS28CRP: 28-joint disease activity score using C-reactive protein level; 28SJC: 28 swollen joint count; 28TJC: 28 tender joint count; VAS: Visual Analog Scale; PGA: Patient Global Assessment; High-risk (HR)/ low-risk (LR) treatment groups: patients with/without a poor prognostic profile based on the presence of classical prognostic factors; SSL: Social Support List.

Table 2. Patient-reported outcomes at trial baseline and at year 1 of treatment for the sample of patients with early rheumatoid arthritis (n=333) subdivided according to their initial clinical response profile

Variables	Persistent responders (n=140)	Secondary failures (n=92)	Delayed responders (n=71)	Non- responders (n=30)	P-value
SF-36 Vitality (0-100)					
Baseline, n=326	52.9±19.0	46.4±20.2	40.6±19.2	47.0±22.2	<0.001 ^a
Week 52, n=333	64.4±17.0	54.4±17.7	53.8±18.1	46.3±18.0	<0.001 ^a
SF-36 Social functioning (0-100)					
Baseline, n=327	69.1±25.6	60.9±26.3	51.6±27.9	57.6±24.4	<0.001 ^b
Week 52, n=333	83.0±19.6	74.0±21.2	70.8±20.2	56.4±18.1	<0.001 ^b
SF-36 Role emotional (0-100)					
Baseline, n=324	69.9±42.9	59.3±42.9	58.5±45.2	61.7±45.0	0.097 ^b
Week 52, n=333	84.9±29.6	68.7±39.3	61.3±40.2	55.4±37.9	<0.001 ^b
SF-36 Mental health (0-100)					
Baseline, n=327	65.7±17.3	61.5±18.5	55.8±22.2	63.6±17.9	0.005 ^a
Week 52, n=333	75.2±15.5	67.4±17.3	66.4±16.5	63.4±15.1	<0.001 ^a
IPQ-R Consequences (6-30)					

Baseline, n=325	19.2±4.1	18.9±3.9	21.0±4.4	19.2±4.6	0.010 ^a
Week 52, n=333	15.9±4.2	17.7±4.1	19.2±4.7	20.8±3.7	<0.001 ^a
IPQ-R Treatment					
control (5-25)					
Baseline, n=326	18.9±2.2	18.7±2.3	18.0±2.7	19.2±2.3	0.055 ^b
Week 52, n=333	18.7±2.2	17.5±2.4	17.2±2.3	17.3±2.1	<0.001 ^b
IPQ-R Illness					
coherence (5-25)					
Baseline, n=325	17.6±3.8	17.1±3.6	16.1±4.9	17.2±3.4	0.175 ^c
Week 52, n=333	18.5±3.4	17.2±3.7	16.9±3.9	17.2±2.7	0.004 ^a

Note. Values reported are mean± standard deviation for continuous variables, and absolute numbers (proportions) for categorical variables. P-values were calculated at baseline only using: ^a 1-way analysis of variance (ANOVA) test, ^b Kruskal-Wallis test, ^c Welch's ANOVA test; SF-36: 36-item Short Form Health Survey; IPQ-R: Revised Illness Perception Questionnaires.

FIGURE LEGENDS

Figure 1. Overview of the CareRA study design. BL: baseline (i.e., randomization to treatment group); W: week; High-risk group: patients with a poor prognostic profile based on the presence of classical prognostic factors (i.e., erosions, rheumatoid factor and/or anticitrullinated protein antibody and DAS28CRP calculated at screening); Low-risk group: patients without a poor prognostic profile; PROs: Patient-Reported Outcomes; n=379: randomized patients in the CareRA trial; COBRA Classic: 15 mg methotrexate (MTX) weekly, 2 g sulfasalazine daily and a daily oral prednisone bridging scheme (weekly step-down from BL until W7 (60-40-25-20-15-10-7.5 mg), continuing 7.5 mg from W7 to W28, and then tapering down weekly until discontinuation at W34); COBRA Slim: 15 mg MTX weekly with a daily oral prednisone bridging scheme (weekly step-down from BL until W6 (30-20-12.5-10-7.5-5 mg), continuing 5 mg from W6 to W28, and then tapering down weekly until discontinuation at W34); COBRA Avant-garde: 15 mg MTX weekly, 10 mg leflunomide daily and a daily oral prednisone bridging scheme (weekly step-down from BL until W6 (30-20-12.5-10-7.5-5 mg), continuing 5 mg from W6 to W28, and then tapering down weekly until discontinuation at W34); Tight Step-up: 15 mg MTX weekly, no oral glucocorticoids allowed.

Figure 2. Mean 36-item Short Form Health Survey (SF-36) domain scores at baseline, week 16, week 52 and week 104 plotted on spidergrams for each high-risk treatment group within the CareRA trial. Patients with early rheumatoid arthritis stratified to a high-risk profile and then randomized to 1:3 treatment groups: (A) COBRA Classic, (B) COBRA Slim and (C) COBRA Avant-garde. High-risk group: patients with a poor prognostic profile based on the presence of classical prognostic

factors. For week 16, week 52 and week 104: missing data were imputed using the Expectation-Maximization method.

Figure 3. Mean 36-item Short Form Health Survey (SF-36) domain scores at baseline, week 16, week 52 and week 104 plotted on spidergrams for each low-risk treatment group within the CareRA trial. Patients with early rheumatoid arthritis stratified to a low-risk profile and then randomized to 1:2 treatment groups (A): Tight Step-up and (B) COBRA Slim. Low-risk group: patients without a poor prognostic profile based on the presence of classical prognostic factors. For week 16, week 52 and week 104: missing data were imputed using the Expectation-Maximization method.

Figure 4. Multivariate linear regression models of the contribution of the initial clinical response and other predictor variables (age, sex, baseline depression, baseline DAS28CRP, treatment regimen and baseline outcome value) to psychosocial outcomes after 1 year of early RA treatment. Graphs show adjusted standardized beta values and significant p-values (* $p < 0.05$ corrected with Holm-Bonferroni) based on 1000 bootstrap samples; BL: baseline; LR: low-risk; HR: high-risk. See supplementary file 3 for details about 95% confidence intervals, and supplementary file 2 for details about the univariate analyses.

Fig 1

Fig 2

Fig 3

Fig 4

Supplementary file 1. Flow chart of the patient selection for the analysis of the role of initial clinical response profiles

BASELINE

Supplementary file 2. Univariate linear regression analyses

2A. Univariate linear regression models of the contribution of initial clinical response to patient-perceived health (SF-36) and illness perceptions (IPQ-R) subscales after 1 year of early RA treatment

Psychosocial outcome at year 1	Number of patients	Persistent responders	Secondary failures	Delayed responders	Non-responders	Adjusted R ²
SF-36 subscales						
Vitality (0-100)	n=326	reference	B = -10.02 (-14.41, -5.18)	B = -10.64 (-15.89, -4.70)	B = -18.12 (-25.31, -10.41)	10.0%
Social functioning (0-100)	n=327	reference	B = -8.91 (-14.26, -3.28)	B = -12.14 (-17.78, -6.09)	B = -26.56 (-34.57, -18.77)	12.6%
Role emotional (0-100)	n=324	reference	B = -16.17 (-25.69, -6.95)	B = -23.58 (-33.90, -13.49)	B = -29.50 (-43.65, -16.75)	8.0%
Mental health (0-100)	n=327	reference	B = -7.81 (-12.18, -2.97)	B = -8.78 (-13.36, -4.08)	B = -11.77 (-18.19, -5.63)	6.3%
IPQ-R subscales						
Consequences (6-30)	n=325	reference	B = 1.83 (0.74, 2.80)	B = 3.33 (1.98, 4.66)	B = 4.99 (3.33, 6.38)	12.7%
Treatment control (5-25)	n=326	reference	B = -1.13 (-1.77, -0.47)	B = -1.42 (-2.1, -0.70)	B = -1.38 (-2.22, -0.52)	6.6%
Illness coherence (5-25)	n=325	reference	B = -1.32 (-2.24, -0.38)	B = -1.62 (-2.66, -0.62)	B = -1.35 (-2.37, -0.22)	3.1%

Note. SF-36: 36-item Short Form Health Survey; IPQ-R: Revised Illness Perception Questionnaire; reference: the beta values show the relative difference between each clinical response group and the persistent responders group (i.e., the group we chose as the reference category); B: adjusted unstandardized beta weights; 95% bias corrected and accelerated confidence intervals (CI) reported in parentheses. CI and significance values based on 1000 bootstrap samples. Significant results in bold ($p < 0.05$, not adjusted for multiple testing).

2B. Univariate linear regression models of the contribution of the baseline value of the respective dependent outcome of patient-perceived health (SF-36) and illness perceptions (IPQ-R) subscales after 1 year of early RA treatment

Psychosocial outcome at year 1	Number of patients	Baseline value of the dependent outcome	Adjusted R ²
SF-36 subscales			
Vitality (0-100)	n=326	B = 0.46 (0.37, 0.55)	25.2%
Social functioning (0-100)	n=327	B = 0.30 (0.23, 0.38)	14.6%
Role emotional (0-100)	n=324	B = 0.35 (0.25, 0.45)	16.9%
Mental health (0-100)	n=327	B = 0.51 (0.42, 0.61)	34.1%
IPQ-R subscales			
Consequences (6-30)	n=325	B = 0.56 (0.46, 0.67)	26.7%
Treatment control (5-25)	n=326	B = 0.26 (0.14, 0.37)	6.4%
Illness coherence (5-25)	n=325	B = 0.53 (0.44, 0.63)	35.2%

Note. SF-36: 36-item Short Form Health Survey; IPQ-R: Revised Illness Perception Questionnaire; reference: the beta values show the relative difference between each clinical response group and the persistent responders group (i.e., the group we chose as the reference category); B: adjusted unstandardized beta weights; 95% bias corrected and accelerated confidence intervals (CI) reported in parentheses. CI and significance values based on 1000 bootstrap samples. Significant results in bold ($p < 0.05$, not adjusted for multiple testing).

Supplementary file 3. Multivariate linear regression models of the contribution of initial clinical response to patient-perceived health (SF-36) and illness perceptions (IPQ-R) subscales after 1 year of early RA treatment

Variable	Adjusted B (95% CI)	SE B	β	P-value	Adjusted R ² of the model
SF-36 Vitality (0-100), n= 326					34.5%
(Constant)	31.03 (16.06, 47.13)	8.08		0.001	
Persistent responders vs Secondary failures	-6.73 (-10.85, -2.43)	2.19	-0.16	0.005	
Persistent responders vs Delayed responders	-8.29 (-13.43, -3.39)	2.51	-0.19	0.004	
Persistent responders vs Non-responders	-17.65 (-23.37, -12.05)	3.02	-0.27	0.001	
Age	0.01 (-0.14, 0.16)	0.08	0.01	0.891	
Sex	-1.49 (-5.53, 2.87)	1.99	-0.04	0.461	
Presence of depression	-1.82 (-8.67, 5.14)	3.44	-0.02	0.584	
Baseline disease activity (DAS28CRP)	3.13 (1.63, 4.64)	0.78	0.22	0.001	
COBRA slim (low-risk) vs COBRA Classic	-2.45 (-7.84, 2.44)	3.00	-0.06	0.422	
COBRA slim (low-risk) vs COBRA Slim (high-risk)	-3.99 (-9.72, 1.36)	3.04	-0.10	0.190	
COBRA slim (low-risk) vs COBRA Avant-garde	-5.26 (-11.32, -0.29)	3.06	-0.13	0.095	
COBRA slim (low-risk) vs Tight Step-up	0.84 (-5.40, 6.65)	3.38	0.02	0.816	
Baseline SF-36 Vitality	0.46 (0.35, 0.56)	0.05	0.50	0.001	
SF-36 Social functioning (0-100), n= 327					25.3%
	Values	Adjusted B (95% CI)	SE B	β	P
(Constant)		56.37 (39.20, 73.49)	8.4		0.001
Persistent responders vs Secondary failures		-6.13 (-11.22, -0.98)	2.58	-0.13	0.015
Persistent responders vs Delayed responders		-9.01 (-14.72, -3.49)	2.90	-0.18	0.003
Persistent responders vs Non-responders		-24.80 (-32.36, -17.82)	3.83	-0.33	0.001
Age		-0.08 (-0.26, 0.08)	0.09	-0.05	0.349
Sex		1.71 (-2.48, 5.61)	2.19	0.04	0.440
Presence of depression		-8.84 (-15.88, -2.45)	3.43	-0.09	0.009
Baseline disease activity (DAS28CRP)		2.95 (0.92, 5.09)	0.98	0.18	0.003
COBRA slim (low-risk) vs COBRA Classic		-5.62 (-13.19, 1.38)	3.38	-0.12	0.101
COBRA slim (low-risk) vs COBRA Slim (high-risk)		-7.17 (-14.04, -0.02)	3.57	-0.15	0.042
COBRA slim (low-risk) vs COBRA Avant-garde		-9.00 (-16.24, -1.53)	3.62	-0.19	0.014
COBRA slim (low-risk) vs Tight Step-up		-6.02 (-13.69, 1.36)	3.84	-0.09	0.108
Baseline SF-36 Social functioning		0.31 (0.23, 0.40)	0.04	0.40	0.001

SF-36 Role emotional (0-100), n= 324						23.4%
	Values	Adjusted B (95% CI)	SE B	β	P	
(Constant)		62.21 (37.02, 87.83)	13.03		0.001	
Persistent responders vs Secondary failures		-11.97 (-22.16, -2.06)	4.73	-0.15	0.016	
Persistent responders vs Delayed responders		-21.06 (-31.23, -11.69)	5.23	-0.24	0.001	
Persistent responders vs Non-responders		-26.67 (-42.84, -10.62)	7.76	-0.20	0.002	
Age		0.09 (-0.22, 0.34)	0.15	0.03	0.524	
Sex		-1.32 (-8.45, 6.33)	3.77	-0.02	0.723	
Presence of depression		-22.45 (-36.85, -6.63)	7.89	-0.13	0.004	
Baseline disease activity (DAS28CRP)		0.85 (-2.37, 3.79)	1.56	0.03	0.585	
COBRA slim (low-risk) vs COBRA Classic		-6.53 (-16.77, 5.38)	5.96	-0.08	0.277	
COBRA slim (low-risk) vs COBRA Slim (high-risk)		-3.84 (-15.82, 8.89)	6.10	-0.05	0.532	
COBRA slim (low-risk) vs COBRA Avant-garde		-8.47 (-20.53, 4.37)	6.04	-0.10	0.157	
COBRA slim (low-risk) vs Tight Step-up		1.60 (-12.44, 15.43)	7.19	0.01	0.832	
Baseline SF-36 Role emotional		0.32 (0.22, 0.41)	0.05	0.38	0.001	
SF-36 Mental health (0-100), n= 327						38.4%
	Values	Adjusted B (95% CI)	SE B	β	P	
(Constant)		35.50 (22.26, 48.79)	6.85		0.001	
Persistent responders vs Secondary failures		-5.96 (-9.29, -2.45)	1.82	-0.16	0.001	
Persistent responders vs Delayed responders		-4.92 (-9.08, -0.77)	2.08	-0.12	0.019	
Persistent responders vs Non-responders		-11.48 (-17.28, -5.05)	2.89	-0.19	0.001	
Age		0.03 (-0.10, 0.16)	0.07	0.03	0.636	
Sex		0.13 (-2.92, 3.43)	1.61	0.00	0.949	
Presence of depression		-4.50 (-10.74, 1.84)	3.45	-0.06	0.192	
Baseline disease activity (DAS28CRP)		1.24 (-0.33, 2.83)	0.74	0.09	0.095	
COBRA slim (low-risk) vs COBRA Classic		0.49 (-4.53, 5.48)	2.51	0.01	0.861	
COBRA slim (low-risk) vs COBRA Slim (high-risk)		0.52 (-4.40, 5.16)	2.74	0.01	0.840	
COBRA slim (low-risk) vs COBRA Avant-garde		-3.00 (-7.92, 1.55)	2.56	-0.08	0.238	
COBRA slim (low-risk) vs Tight Step-up		1.75 (-3.69, 7.01)	3.06	0.03	0.558	
Baseline SF-36 Mental health		0.50 (0.40, 0.59)	0.05	0.57	0.001	

IPQ-R Consequences (6-30), n= 325						37.4%
	Values	Adjusted B (95% CI)	SE B	β	P	
(Constant)		7.10 (3.69, 10.55)	1.71		0.001	
Persistent responders		2.03 (1.12, 3.00)	0.46	0.20	0.001	
vs Secondary failures						
Persistent responders		2.74 (1.49, 3.99)	0.61	0.25	0.001	
vs Delayed responders						
Persistent responders		5.37 (4.16, 6.47)	0.59	0.33	0.001	
vs Non-responders						
Age		-0.01 (-0.04, 0.01)	0.02	-0.04	0.360	
Sex		-0.39 (-1.30, 0.58)	0.46	-0.04	0.384	
Presence of depression		0.97 (-0.96, 2.67)	0.85	0.05	0.239	
Baseline disease activity (DAS28CRP)		-0.38 (-0.70, -0.05)	0.17	-0.11	0.027	
COBRA slim (low-risk)		1.10 (-0.28, 2.46)	0.68	0.11	0.109	
vs COBRA Classic						
COBRA slim (low-risk)		0.80 (-0.58, 2.13)	0.70	0.08	0.261	
vs COBRA Slim (high-risk)						
COBRA slim (low-risk)		1.17 (-0.07, 2.38)	0.67	0.11	0.086	
vs COBRA Avant-garde						
COBRA slim (low-risk)		0.39 (-1.27, 2.07)	0.82	0.03	0.605	
vs Tight Step-up						
Baseline IPQ-R Consequences		0.57 (0.45, 0.69)	0.05	0.52	0.001	
IPQ-R Treatment control (5-25), n= 326						11.7%
	Values	Adjusted B (95% CI)	SE B	β	P	
(Constant)		13.28 (10.59, 16.12)	1.43		0.001	
Persistent responders		-1.05 (-1.64, -0.52)	0.31	-0.20	0.003	
vs Secondary failures						
Persistent responders		-1.33 (-2.15, -0.64)	0.37	-0.23	0.001	
vs Delayed responders						
Persistent responders		-1.59 (-2.45, -0.80)	0.43	-0.19	0.001	
vs Non-responders						
Age		0.00 (-0.02, 0.02)	0.01	-0.00	0.992	
Sex		0.26 (-0.29, 0.83)	0.28	0.05	0.339	
Presence of depression		-0.04 (-0.75, 0.71)	0.38	-0.00	0.911	
Baseline disease activity (DAS28CRP)		0.14 (-0.07, 0.35)	0.11	0.07	0.202	
COBRA slim (low-risk)		-0.57 (-1.42, 0.36)	0.45	-0.11	0.210	
vs COBRA Classic						
COBRA slim (low-risk)		-0.87 (-1.80, 0.03)	0.45	-0.16	0.051	
vs COBRA Slim (high-risk)						
COBRA slim (low-risk)		-0.75 (-1.64, 0.15)	0.45	-0.14	0.100	
vs COBRA Avant-garde						
COBRA slim (low-risk)		-0.04 (-0.96, 0.99)	0.47	-0.01	0.944	
vs Tight Step-up						
Baseline IPQ-R Treatment control		0.26 (0.14, 0.37)	0.06	0.26	0.001	

IPQ-R Illness coherence (5-25), n= 325					36.3%
	Values	Adjusted B (95% CI)	SE B	β	P
(Constant)		9.38 (6.23, 12.15)	1.59		0.001
Persistent responders vs Secondary failures		-1.03 (-1.84,-0.22)	0.41	-0.13	0.009
Persistent responders vs Delayed responders		-0.98 (-1.94, 0.03)	0.49	-0.11	0.042
Persistent responders vs Non-responders		-1.21 (-2.33, -0.15)	0.51	-0.09	0.012
Age		-0.01 (-0.03, 0.02)	0.01	-0.02	0.723
Sex		-0.05 (-0.69, 0.68)	0.36	-0.01	0.899
Presence of depression		-1.59 (-2.94, -0.25)	0.64	-0.10	0.008
Baseline disease activity (DAS28CRP)		0.13 (-0.17, 0.42)	0.15	0.05	0.374
COBRA slim (low-risk) vs COBRA Classic		-0.48 (-1.53, 0.68)	0.60	-0.06	0.461
COBRA slim (low-risk) vs COBRA Slim (high-risk)		-0.23 (-1.29, 1.0)	0.61	-0.03	0.713
COBRA slim (low-risk) vs COBRA Avant-garde		0.07 (-1.0, 1.23)	0.60	0.01	0.908
COBRA slim (low-risk) vs Tight Step-up		0.17 (-1.14, 1.54)	0.76	0.02	0.830
Baseline IPQ-R Illness coherence		0.52 (0.42, 0.62)	0.05	0.58	0.001

Note. SF-36: 36-item Short Form Health Survey; IPQ-R: Revised Illness Perception Questionnaire; B: adjusted unstandardized beta weights; 95% bias corrected and accelerated confidence intervals (CI) reported in parentheses. CI, standard errors and significance values based on 1000 bootstrap samples; β : standardized beta values; SE: standard error. Original p-values are shown, i.e. unadjusted for multiple testing.

Supplementary file 4. Sensitivity analysis including the comparison of parameter estimates in the multivariate regression models predicting psychosocial outcomes at year 1 of early rheumatoid arthritis treatment, based on complete cases only versus imputed cases using the Expectation Maximization method

Variable	Missing Week 52	Adjusted B (95% CI) Not imputed population	P- value	Adjusted B (95% CI) Imputed population	P- value
SF-36 Vitality (0-100)	22.5%	n= 258		n=333	
Persistent responders		reference		reference	
Secondary failures		-7.70 (-12.58, -3.39)	0.004	-6.73 (-10.85, -2.43)	0.005
Delayed responders		-8.25 (-13.17, -3.35)	0.006	-8.29 (-13.43, -3.39)	0.004
Non-responders		-15.94 (-22.55, -8.76)	0.001	-17.65 (-23.37, -12.05)	0.001
SF-36 Social Functioning (0-100)	22.2%	n= 259		n=333	
Persistent responders		reference		reference	
Secondary failures		-4.86 (-11.44, 1.11)	0.125	-6.13 (-11.22, -0.98)	0.015
Delayed responders		-7.77 (-13.92, -1.39)	0.026	-9.01 (-14.72, -3.49)	0.003
Non-responders		-25.44 (-36.41, -16.12)	0.001	-24.80 (-32.36, -17.82)	0.001
SF-36 Role Emotional (0-100)	23.1%	n= 256		n=333	
Persistent responders		reference		reference	
Secondary failures		-9.88 (-21.13, 1.22)	0.088	-11.97 (-22.16, -2.06)	0.016
Delayed responders		-23.29 (-35.47, -10.84)	0.002	-21.06 (-31.23, -11.69)	0.001
Non-responders		-29.27 (-50.43, -7.45)	0.009	-26.67 (-42.84, -10.62)	0.002
SF-36 Mental health (0-100)	22.5%	n= 258		n=333	
Persistent responders		reference		reference	
Secondary failures		-6.93 (-11.46, -2.21)	0.004	-5.96 (-9.29, -2.45)	0.001
Delayed responders		-4.81 (-9.64, -0.20)	0.044	-4.92 (-9.08, -0.77)	0.019
Non-responders		-10.62 (-17.62, -3.66)	0.005	-11.48 (-17.28, -5.05)	0.001
IPQ-R Consequences (16-30)	22.5%	n= 258		n=333	
Persistent responders		reference		reference	
Secondary failures		2.22 (0.87, 3.60)	0.002	2.03 (1.12, 3.00)	0.001
Delayed responders		2.89 (1.48, 4.23)	0.001	2.74 (1.49, 3.99)	0.001
Non-responders		5.92 (4.22, 7.47)	0.001	5.37 (4.16, 6.47)	0.001
IPQ-R Treatment control (5-25)	22.5%	n= 258		n=333	
Persistent responders		reference		reference	
Secondary failures		-1.14 (-1.87, -0.42)	0.005	-1.05 (-1.64, -0.52)	0.003
Delayed responders		-1.30 (-2.19, -0.42)	0.006	-1.33 (-2.15, -0.64)	0.001
Non-responders		-1.60 (-2.86, -0.42)	0.007	-1.59 (-2.45, -0.80)	0.001
IPQ-R Illness coherence (5-25)	22.8%	n= 257		n=333	
Persistent responders		reference		reference	
Secondary failures		-0.84 (-1.85, 0.14)	0.105	-1.03 (-1.84, -0.22)	0.009
Delayed responders		-0.78 (-1.87, 0.29)	0.181	-0.98 (-1.94, 0.03)	0.042
Non-responders		-0.98 (-2.27, 0.27)	0.102	-1.21 (-2.33, -0.15)	0.012

Note. CC: Complete cases; EM: Expectation-Maximization; SF-36-item Short Form Health Survey; IPQ-R: Revised Illness Perception Questionnaire; reference: the beta values show the relative difference between each clinical response group and the persistent responders group (i.e., the group we chose as the reference category B: adjusted unstandardized beta weights; 95% bias corrected and accelerated confidence intervals (CI) reported in parentheses. CI and significance values based on 1000 bootstrap samples; Regression analyses (enter method) were adjusted for age, gender, baseline disease activity, presence of depression, baseline value of the outcome variable, and allocated treatment. Original p-values are shown, i.e. unadjusted for multiple testing.