Author’s response to reviews

Title: Predictors of biologic-free disease control in patients with rheumatoid arthritis after stopping tumor necrosis factor inhibitor treatment

Authors:
Marjan Ghiti Moghadam (m.ghiti@mst.nl)
Femke Lamers-Karnebeek (femke.lamers-karnebeek@radboudumc.nl)
Harald V onkeman (h.vonkeman@mst.nl)
Peter ten Klooster (p.m.tenklooster@utwente.nl)
Janneke Tekstra (j.tekstra@umcutrecht.nl)
Barbara van Schaeybroeck (B.Schaeybroeck@asz.nl)
Ruth Klaasen (r.klaasen@meandermc.nl)
Marieke van Onna (m.vanonna@amc.uva.nl)
Henk Visser (hvisser@rijnstate.nl)
Marc Kok (kokmr@maasstadziekenhuis.nl)
Piet van Riel (piet.vanriel@radboudumc.nl)
Annemarie Schilder (am.schilder@znb.nl)
Robert Landewé (landewe@rlandewe.nl)
Mart van de Laar (M.vandeLaar@mst.nl)
Tim Jansen (tjansen@viecuri.nl)
Hein Bernelot Moens (h.moens@zgt.nl)

Version: 1 Date: 27 Feb 2019

Author’s response to reviews:
Point-by-point reply:

Veena Ranganath (Reviewer 1): The manuscript entitled "Predictors of biologic-free disease control in patients with rheumatoid arthritis after stopping tumor necrosis factor inhibitor treatment" by Giti Moghadam and colleagues is based on an interesting cohort that has been previously published off of in at least 5 other papers. Each of the variables described as independently predicting disease control after TNFi discontinuation have been described prior (disease duration, low/moderate MBDA (whole paper dedicated to this)), except for antibody type TNFi. In addition, the prediction model proposed can be improved on.

Response: The reviewer is correct in that outcomes of the POET study have already been published. Also, several different potential predictors have already been explored in other articles or in or in specific small subsamples. In this study, however, we much more thoroughly examined as many potential predictors as possible, using an optimal sample size and more thorough and robust analyses, e.g. by also examining interaction-effects with type of TNFi, which was not done before at the time of analysis. This resulted in a novel and relevant finding that type of TNFi was actually one of the strongest predictors of successful discontinuation. Interestingly, in the mean-time, Hashimoto et al. very recently (while the current study was under review) also reported this finding. We now also refer to this study in the discussion.

Minor- page 5 line 124, "We used date…" most likely should be changed to "We used data…"

Response: Thank you for pointing us to this typo, we changed this.

p.5 line 130- The inclusion criteria for the study is rather loose, only requiring 6mos in LDA and/or rheumatologists opinion. And, the cutpoint of MBDA of 44 used in the models is also rather loose, since this describes low disease activity or remission. Why wasn't MBDA remission cutpoint used? In addition, what % of patients were on prednisone at baseline in the withdrawal group? These patients should be eliminated from the cohort, since in clinical practice, prednisone would be the first medication to discontinue not the TNFi.

Response: We agree that these inclusion criteria could be considered rather loose, especially when compared to previous RCTs. However, POET was deliberately designed as a pragmatic study, trying to reflect clinical decision processes in daily clinical practice. The cut-point of 44 was used since our previous study of the MBDA already showed that patients with either low (<30) or moderate (30–44) MBDA scores showed very similar patterns of flare-free survival which were both significantly different to patients with high MBDA scores (Ghiti Moghadam et al 2018). We added the reference to this study in the analysis section to clarify this choice.

In the current sample, 26 (6%) of the patients were receiving prednisone at baseline. We fully agree with the reviewer that in clinical practice, stopping prednisone is generally the first step when patients are in remission considering their long-term side effects. This is also explicitly recommended in the published international guidelines. However, clinical experience also suggests that it proves difficult to stop prednisone in a small subset of patients. Since the POET study was a pragmatic study in real-life patients, for which the rheumatologists judged that they were eligible for stopping their bDMARD, we decided not to exclude these patients in order to preserve the representativeness of the sample instead of creating a more homogenous selected group of patients. We do now explicitly mention the
proportion of patients using corticosteroids at baseline in Table 1. We now also explicitly mention this as a limitation in the discussion (page 14, see also comment 10 by reviewer 2).

p.5- please clearly define antibody type TNFi in the methods section, what was the considered reference variable.

Response: Both antibody TNFi and receptor antagonist are now more clearly defined in the methods (page 7) and “and” was changed into “versus” to clarify the reference category (page 4-5).

p.7 line 182- Instead of using the Nagelkerke's pseudo R2 to examine the model strength it is typical to evaluate by the area under the ROC curve to summarize the model if the goal is to predict the outcome. In addition, there were only a few patients with missing baseline values (N=38), thus it is not clear what the advantage of utilizing multiple imputations is. It would probably be better to use the data available- especially when performing a prediction modeling exercise.

Response: The area under the ROC curve is indeed a better interpretable static for examining the strength of a multivariable logistic model. We calculated and now present this statistic for model strength (Statistical analysis page 7, Table 3 page 10) and deleted the R2 statistics. We now also report this statistic in the abstract. We fully agree that given the low proportion of patients with missing values and the analytic purpose of predicting, there is theoretically not much advantage of using multiple imputation. Therefore, we now present the results of the models based on observed data only (and Table 3, page 10). We now mention in the analysis section only that repeated models using imputed data showed very similar results, actually conforming that the missing values did not bias the results.

p.10 Table 2- Which was the reference for the Type of TNFi? Antibody or Receptor agonist? This is not clearly stated.

Response: The reference category is now clearly indicated in Tables 2 and 3 (variable rename in Table 2 and footnote below both tables).

Would suggest that the authors provide logistic regression model with predictors that are continuous (ie MBDA, disease duration, etc) rather than utilizing arbitrary cutpoints.

Response: We agree that dichotomizing predictors may result in loss of predictive power. For clinicians and clinical purposes, on the other hand, dichotomized predictors are generally more useful and more easily than ORs for continuous scores in their own metric. Moreover, dichotomization allows for easier comparison of ORs. We therefore kept the model results for the dichotomized predictors, but also performed additional analyses with disease duration and MBDA scores as continuous predictors. Since these provided almost identical results, these models were added as an additional file (Additional file 1).

What is the added value of MBDA as a predictor? This answer hasn't been fully considered. Suggest building model with MBDA and without it in the model. Would compare AUC with MBDA and without it.
Response: Thank you for this suggestion. The MBDA indeed did not appear to have much additional value as a predictor, as evident from only a minor decrease of the area under the ROC curve for the multivariate model with and without MBDA as a predictor (from 0.66 to 0.65). We now state this explicitly in the results section (page 8).

Also, evaluating radiographic progression as an outcome would be of value. Perhaps "RA flare" as defined may not truly delineate a RA flare, but perhaps a fibro or OA flare. Would be of interest to know the type of patient who doesn't have radiographic progression, but may still have flares?

Response: We agree that accounting for the presence of radiographic progression would have been very interesting. Unfortunately, radiographic data was not collected in the POET trial.

Why wasn't the ultrasound data also evaluated in this analysis?

Response: The ultrasound data reported by Lamers-Karnebeek et al (2017) concerned a small subsample of 259 POET patients in a selective sample of hospitals. In the current study we wanted to use an optimal and representative sample (size). Moreover, the study of Lamers-Karnebeek already indicated that US had limited value in predicting flares in these patients.

Jennifer Barton (Reviewer 2): The authors present the results of a post-hoc analysis of a pragmatic, open-label multicenter trial of patients with RA in low disease activity on a TNFi who were randomized 2:1 to either discontinue their TNFi or to continue. The present study only looks at those patients from the parent study who discontinued their TNFi to see what factors were associated with achievement of biologic-free disease control. This is a contribution to the literature in that this field of study is nascent yet important - to best determine who/when can discontinue TNFi therapy. The paper is well-written and concise however there are a few areas in need of clarification as follows:

1. In the abstract, would be helpful to know the proportion who were seropositive in the study

Response: We added the percentage of RF positive patients to the abstract.

2. Background, page 4, line 94 - please define what is meant by "persistent remission"

Response: This is an excellent point. Unfortunately, the EULAR recommendations (Smolen et al., 2014) do not explicitly define “persistent remission” in their recommendation about tapering bDMARDs. Probably because most studies on this topic also used very different inclusion criteria with respect to both the measurement and definition of remission and the required time being in remission. We added this lack of a definition for persistent remission to the introduction.

3. Methods, page 5, line 125 - please specify which disease activity measure was used to define low disease activity (LDA) [comes later in paragraph that it is by DAS28 - consider mentioning with first mention of LDA, minor point] 4. Methods, page 6, line 139 - was prednisone or corticosteroid use in
general captured, and if so, dose?

Response: As suggested, we now specify the DAS28-ESR with the first mention of LDA. Corticosteroids use at baseline was indeed captured, but not the dosage, and dealing with corticosteroids use was not part of the protocol. We now included the proportion of patients using corticosteroids at baseline in Table 1.

5. Results, page 8, line 194 - would restate the total n for the current analysis

Response: We added the total n between brackets

6. Results, page 8, line 198 - does the 251 with a physician-reported flare include the 219 who restarted their TNFi (as mentioned in the first half of the sentence)? Are these two groups overlapping or does the 251 include the 219? this is not clear as written

Response: This is a good point. These two groups did not fully overlap. The exact overlap is now added to the results (page 8).

7. Results, page 8, lines 208-2010 - consider presenting the the point estimates as opposed to just p-value

Response: The actual OR and 95% CI was added to the p-value.

8. Table 1, page 9 - would add the values for "normal BMI" and provide the range for MBDA as well as the actual cut-point values for low and moderate MBDA scores

Response: The value for normal BMI and the cut-points for MBDA scores were added to Table 1.

9. Discussion, page 12, lines 288-292 - please provide citations

Response: The relevant citations were added to the different sentences.

10. Discussion, page 13, lines 303-304 - is there an explanation for why predictors were only measured at baseline in the parent study? Would comment on role of steroids, pattern of synthetic DMARD use in the study - and if not captured, list as limitation

Response: The reasons for this are both conceptual and pragmatic. First and foremost, we wanted to examine whether there would be robust predictors for rheumatologists and patients for decision making when faced with the wish or possibility to stop a patient’s TNFi. For this purpose, “baseline” measurement are only available in such cases. Secondly, for several potential predictors we simply only had a baseline measurement available, such as for MBDA scores and erosion. For reasons of clarity, we therefore decided to only focus on the predictive value of baseline variables.
We now also specifically comment on the fact that a small proportion of patients still used corticosteroids when discontinuing their bDMARD (page 14, see also comment 3 by reviewer 1).

11. Discussion, page 13, lines 313 - could the authors present thoughts on next steps for this area of research or where these findings should lead us - in terms of practice and/or future research

Response: We added two sentences on our thoughts on relevant venues for future research in this area.

12. General comment - consider using the phrase "anti-TNF monoclonal antibodies" or monoclonal antibodies instead of TNF antibodies to describe adalimumab and others

Response: For consistency, we changed this terminology throughout the manuscript.