Author's response to reviews

Title: Radiographic progression in early rheumatoid arthritis patients following initial combination versus step-up treat-to-target therapy in daily clinical practice: results from the DREAM registry

Authors:

Laura Steunebrink (lmm.steunebrink@mst.nl)
Letty Versteeg (g.a.versteeg@utwente.nl)
Harald Vonkeman (h.vonkeman@mst.nl)
Peter Ten Klooster (p.m.tenklooster@utwente.nl)
Monique Hoekstra (mo.hoekstra@isala.nl)
Mart van de Laar (m.vandelaar@mst.nl)

Version: 1 Date: 17 Nov 2017

Author's response to reviews:

BMC Rheumatology

November 17, 2017

BRHM-D-17-00012

Dear Dr. James Mockridge,

We thank you for your interest in our manuscript entitled "Radiographic Progression in Early Rheumatoid Arthritis Patients Following Initial Combination Versus Step-up Treat to Target Therapy in Daily Clinical Practice; Results from the DREAM registry"

Please find enclosed the revised version of our manuscript and our response to each point raised by the reviewers. The reviewers had valuable suggestions and we agreed with the feedback which was given. The manuscript was modified accordingly. In addition, the entire manuscript was checked carefully for additional errors. Changes in the manuscript are highlighted in red, furthermore a detailed point-to-point response is attached.

We thank the reviewers for their constructive criticism and for their efforts in helping to improve our manuscript.
We hope that the revised version will be acceptable for publication and look forward to hearing from you.

Also on behalf of the other authors,

Sincerely yours,

Laura Steunebrink, MSc

Arthritis Center Twente, Department of Rheumatology & University of Twente
Medisch Spectrum Twente
PO Box 50 000
7500 KA Enschede
The Netherlands
Tel: + 31 53 487 2450
E-mail: lmm.steunebrink@mst.nl

Response to the comments of the reviewers

Reviewer 1 Susanna Proudman

While well-written and utilising appropriate statistical analyses, the design of this study has significant short-comings. It was not a randomised study even though a case-control design from two observational cohorts was used. There were differences (acknowledged by the authors) in the era in which study was undertaken, the timing of assessment and dose adjustments, use of glucocorticoids, maximal MTX dose used and timing of TNF inhibitor use if needed.

1. The scorers of radiographs were not blinded to group. Were they authors on the manuscript or a third party?

Response: Trained scorers, including the first two authors, assessed different radiographs over time. This scoring assessment is indeed not blinded with respect to strategy and has been noted as a limitation in the discussion section (page 15).

2. Was the study adequately powered given the study design? Was it a retrospective design?

Response: It is a retrospective design, it is not a-priori powered. A post-hoc power analysis showed that with a sample size of 128 patients per strategy we had > 80% power to detect a small to moderate difference ($d = 0.39$) in progression between both strategies using a 2-sided Mann-Whitney u test with an alpha of 0.05.

This information was added to the discussion (page 16).
3. The introduction states "whether initial combination therapy and the subsequent shorter time to remission, as compared to initial step-up monotherapy, also results in better radiological outcomes has not yet been studied" yet it doesn't differ significantly from other studies using initial vs step up combination eg the BeSt study, raising questions about what this study adds to the literature.

Response: This is a very relevant remark. The major strength of this study is the use of real life data from consecutive patients in daily clinical practice, recently diagnosed with RA, who were being treated according to a state-of-the-art T2T remission induction protocol. Other studies, such as the BeSt-study, have used much stricter inclusion criteria. Patients should have active disease with ≥6 of 66 swollen joints, ≥6 of 68 tender joints, and either an erythrocyte sedimentation rate (ESR) ≥28 mm/hour or a global health score of ≥20 mm on a 0–100-mm visual analog scale. Also, in the BeST study, patients are not treated more aggressively in all arms like in the current comparison. Arm 1 consists, for example, of consecutive monotherapy. So there is always a switch to a new medication instead of the new medication being added. In addition, the target in the BeSt was less strict (DAS44 <2.4 instead of DAS28 <2.6 in our study). Our data confirm those previous findings in patients which are much more representative for current daily clinical practice. This is now better explained in the discussion on page 16.

4. What were the criteria for discontinuing DMARDs?

Response: If remission was achieved with DMARDs and/or biologicals, while maintaining remission for at least 6 months, medications were tapered and possibly discontinued starting from the biologicals and continued with the DMARDs.

This is now explicitly stated and explained for the two separate strategies on page 6.

5. What was the difference in mean dose of steroids in each group?

Response: Median dose of triamcinolone at baseline in strategy I was 80 mg, 75% of the patients (N = 3/4) received 80 mg triamcinolone and 25% of the patients (N = 1/4) received 120 mg triamcinolone. The median dose of triamcinolone at baseline in strategy II was 120 mg, 92.5% of the patients (N = 62/67) received 120 mg triamcinolone and 7.5% of the patients (N = 5/67) received 80 mg triamcinolone at baseline. This information was added to the results in the section on descriptive characteristics on page 9.

6. Were there differences in the proportions of patients achieving remission in each group?

Response: The remission rates in the current strategies were extensively reported in a previous publication. First remission within 6 months was achieved in 63.3% (81/128) of the patients in cohort II versus 48.4% (62/128) in cohort I (p= 0.02). First remission within 12 months was achieved in 77.3% (n= 99/128) of the patients in cohort II versus 71.9% (92/128) in cohort I (P =
Median time to first remission was 17 (95% CI = 13.2 – 20.8) weeks in cohort II versus 27 (95% CI = 20.7 – 33.3) weeks in cohort I (P = 0.04) [1].

The global remission rates at 12 months are now added in the background of the strategies (page 3).

Reviewer 2 Daniel F McWilliams

This study shows that people with RA starting treatment in 2006 experienced worse radiographic progression than those that started in 2012. As the periods of recruitment in 2006 and 2012 coincided with different treatment regimens, the authors have concluded that these are the cause of the different rates of radiographic progression. In order to try and control for the lack of randomisation, the cases were matched into pairs at baseline for the purposes of analysis. Almost all cases were matched in this manner, which either suggests that the early RA cases were very similar in 2006 and 2012, or that the matching criteria were slack. From what I can see, the matching appears to be satisfactory. The authors have also tried to control for early treatment started by GPs, by excluding people who had started DMARDs or more than 10mg of prednisolone per day. Despite the matching process, differences were observed at baseline between the 2 study groups. The 2012 group had slightly worse DAS28, but better HAQ at baseline. Other medications were also different at baseline. So one group was not more severe or active at baseline for all criteria. It is possible that this reflects differences in the treatment pathways between 2006 and 2012.

1. My main opinion about this study is that it supports the findings of previous clinical trials, but does not add much extra knowledge. Instead there is greater potential for unknown confounders and channelling bias to influence findings, as the study was observational. Can the authors discuss how this manuscript might add to the clinical trial data? For example, were the study populations more representative of the RA general population?

Response: Thank you for this very relevant remark, which was also made by reviewer 1. This study primarily adds to the literature something to the literature because the current study confirm previous these findings which were demonstrated in much more strictly selected patients in patients that are more representative for those patients seen in daily clinical practice (including those with less active disease and with co-morbidities). This study showed very similar or even better results than those found in previous RCTs, in which stricter inclusions and controlled conditions were used. For example, in the BeSt study the different treatment arms were managed more aggressively than in daily clinical practice. This was addiotnally explained in the discussion on page 16.

2. The recruitment start times of 2006 and 2012 need to be mentioned in the abstract. As many readers will only read the abstract, this major limitation needs to be clear.
Response: We add the times of 2006 and 2012 in the abstract (page 2).

3. The recruitment time periods from start to finish should be listed in the Methods.

Response: Data from 2006 until 2012 were used for strategy I, and data from 2012 and 2013 were used for strategy II. This was added to the section ‘dataselection and study design’ on page 4.

4. The authors acknowledge the weaknesses of the study in the Discussion. I think that they could improve the manuscript (and also show their study's limitation) by adding a graph of baseline and follow up radiographic scores for each Strategy. The authors could show their main outcome variable in an attractive way, and also show the years/time periods of Strategy I and Strategy II.

Response: Thank you for this valuable suggestion. We added a figure with JSN, Erosions, and total SHS score for each strategy at baseline and at 12 months to better illustrate radiographic damage in both strategies and added the years of inclusion in this figure.

“The data are presented as the mean ± standard error of mean

5. If the cases from Strategies I and II were paired, why are there no paired analyses in the Results?

Response: Thank you for this question. Patients were case-controlled-matched for the selection and therefore considered as independent groups. Therefore, independent tests were performed to compare outcomes in both strategies over time.

6. If the cases were matched more stringently in post hoc sensitivity analyses, would the general findings be replicated? Although the n= numbers might become much lower, perhaps the univariate findings could be replicated in this way?

Response: We indeed did not use more strict matching criteria, because it resulted in too small groups and too little power. For reasons of comparison, we preferred to keep the groups exactly equal groups to a previously published article on the disease activity outcomes in both strategies.

7. Can the authors add the VAS-General Health question wording into the manuscript please? Page 6, line 29.

Response: The exact wording van the VAS-GH was added on page 6.
8. The radiographic scoring process needs to be described in more detail. Were all of the radiographs scored by the same 2 people? Were the scorings performed separately in 2006 and 2012? Can we have details of reliability assessments of the observers?

Response: The two readers evaluated the radiographs together, based on consensus, therefore we could not calculated the interobserver reliability scoring of the scores. We added this information on page 7. However, many previous have already shown that SHS scores are generally very reliable. For instance, Boini and Guillemin (2001), Guillemin et al. (2005) confirmed that the interrater reliability of the SHS is generally excellent. Boini and Guillemin (2001) represented the reliability value for Sharp van der Heijde scoring method is \( r = 0.90 \). Guillemin et al (2005) represented the interrater reliability is \( r = 0.93 \) [2, 3]

9. Can we see details of how missing variables were handled? I think that "complete case" analysis was performed. So can we know the final n= for the multivariable regression models please.

Response: we used complete cases, and no imputations were performed. N= 159 in multivariable regression of minimally clinically important difference (MCID5). Within the multivariable regression analysis N = 221. This information was added to tables 3 and 4.

Minor suggestions:

1. This reviewer has a strong preference for 3 decimal places in p values...

Response: We have changed all p-values into 3 decimal p values.

2. …and also for percentages to have no decimal places.

Response: We have adjusted the percentages with no decimal places

3. Baseline SHS score and HAQ in Table 1 have different numbers of decimal places in the different groups.

Response: We adjusted the numbers of decimal places.

References
