PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Trajectory of intensive treat-to-target disease modifying drug regimen in an observational study of an early rheumatoid arthritis cohort
AUTHORS	Thomas, Ranjeny; White, Douglas; Pahau, Helen; Duggan, Emily; Paul, Sanjoy

VERSION 1 - REVIEW

REVIEWER	Hill, Catherine University of Adelaide, Rheumatology Unit The Queen Elizabeth Hospital
REVIEW RETURNED	05-May-2013

THE STUDY	I have answered 'no' to a number of questions above. I have outlined in the reasons in the attached file which addresses the deficient areas of the STROBE checklist. My main areas of reticience is the likelihood of selection bias related to this study created by the exclusion of >50% of the cohort due to inadequate data collection. This manuscript requires a statistical reviewer before consideration for publication. It is an interesting premise and one that influences daily practise in treating RA patients.
RESULTS & CONCLUSIONS	Although I have 'no' to several of these questions, it is not clear that there is a yes/no answer to them, rather that the results and conclusions need to be further clarified. This is a highly regarded research group so the results are 'credible' but generalisability of the results and impact on clinical practice is unlikely, rather this work is hypothesis generating about decision making in the treatment of RA.
REPORTING & ETHICS	There are a number of areas that do not meet the STROBE checklist for cohort studies. I have outlined these in the attached document for the authors to address.
GENERAL COMMENTS	I have attached a summary document with the issues related to the manuscript which I have listed under the STROBE checklist. This is an interesting manuscript that raises the possibility that response to DMARD therapy at a time point as early as 4 weeks predicts treatment response at 12 months. This differs from other previous analysis which have looked at 3 months as the initial response time. However, due to the large number of participants excluded (>50%) from the initial data set due to inadequate data, the risk of bias is high. In addition, no comparative analysis between those included

	and excluded has been undertaken. I have further comments regarding the manuscript which I have outlined below as part of the STROBE Statement Checklist (To be included in reports of cohort studies). The STROBE checklist did not appear to be in the files that I received. Further information should be included under the following checklist item numbers: No 5. Setting. Further information about the referral pattern of the clinic would be helpful. No 6. Further information about the eligibility criteria. Patients were recruited consecutively from the Rheumatology Clinic. Did participants give individual written informed consent? How was 'active' RA defined? More than 50% of participants were excluded due to lack of 12 month data. Did this mean that participants had to have had visits at each of the 9 visits to be included in this analysis? No 7. Further definition of potential confounders. Was symptom duration at first visit recorded? Although it is within the scope of good clinical practice in an RA clinic to measure BP, it does not necessarily follow that this is a confounder. In addition, disease duration is a potential confounder and the duration of symptoms was not included in the models. Is there any information on SES variables which may impact disease outcomes in RA, in addition to the ability to present at each visit? No 9. Bias. No description of any efforts to address potential sources of bias was addressed. As more than 50% of patients were not included in the analysis, this is potentially a major source of bias. A statistical comparison needs to be made of those who were included in this time-dependent analysis and those who were excluded due to inadequate data points for analysis. No 10. No sample size calculation is included. It is not clear why the investigators decided to cease data collection at this time (2008) . No 11. The statistical methods appear sound and overall well

	reported. However, I am not familiar with all of the statistical methods that have been used and would recommend a separate statistical reviewer. It states that imputations were made for missing data. However, it is not clear from the methods how much missing data was allowed in this study for the participants to be included. No 13. No flow diagram was included and may have been useful. Almost 50% of the cohort were excluded due to incomplete data and a flow chart would help the reader interpret the result. A flow chart would also allow authors to show how many participants progressed to other DMARDs and biologic therapy during the study. No 14. Descriptive data. A comparison of the baseline characteristics of the participants who were included in the final analysis and those who were excluded due to incomplete data observations would be appropriate. Table 1 does not explain how smoking was defined. I am not clear the relevance of BP, eGFR and glucose to the current analysis. I understand that cardiovascular risk factors were of importance to other analyses performed on this dataset but probably do not need to be included here. There is no mention of other co-morbidities. This is likely to be of relevance in attending the clinic and ability of the clinician to escalate therapy. There is no indication of the number of participants with missing data for each variable of interest. No 15. Outcome data. Disease outcomes were included. However, no adverse event data was included. It is possible that those who were excluded due to lack of availability of 12 month data had a different disease trajectory and have not been included in this study and increase bias. No 16. Main results. Unadjusted comparisons were shown in Tables 3 and 4. These show the relevant regression analyses but do not show the reasoning behind these choices. It is not clear why these confounders were adjusted for in the analysis. A sensitive analysis should be done to look at the 4 week DAS scores of the participants included in this current analysis and those who were excluded for lack of 12 month data (although it wouldn't be the full 107, I am presuming that a % would have had 4 week data). No 17. Other analysis. There is no subgroup analysis undertaken. It

	is likely to be of interest to compare those with <6 and >6 months symptom duration at presentation. The analysis looking at factors affecting the response trajectory (page 11 of 31) only appears to include a very limited number of factors and do not include disease duration. Could the authors explain why these were chosen and other clinical characteristics not? In addition, as 22% already had low DAS (<=3.2) at baseline, a subgroup analysis excluding these patients may be helpful to further interpret the results. No 18. Limitations. This needs further discussion particular in terms of bias introduced by excluding those with incomplete data observation. Limitation in terms of use of 1987 classification. The slow incremental methotrexate actually means that patients were only taking 10mg/week at first visit (4 weeks) which may be subtherapeutic. Could the steep drop in the DAS scores at 4 weeks be explained by regression to the mean? And the fall over the 3 months is a more reliable and accurate measure. Patients with RA are most likely to present to GP and gain specialist referral when disease activity at its maximal. The proportion of patients with low disease activity did not significantly alter between 6 (25%) and 12 months (29%). Could the authors comment/speculate on reasons for this in the setting of a TTT strategy. No 21. Generalisability. This is discussed but is a major limitation of the work. References: Reference 3. Bakker et al. is incomplete The reference list does not appear to be consistent with the way that BMJ Open presents references.

REVIEWER	Raimon Sanmarti MD Senior consultant rheumatologist Arthritis Unit

	Rheumatology Service Hospital Clinic Barcelona, Spain I declare no conflict of interests.
REVIEW RETURNED	05-May-2013

THE STUDY	This is an study analysing the time-dependant therapeutic response to a DMARD combination strategy using the treat to target aprroach in patients with early RA. Although the study is interessting I have some major concerns that limits the conclusions raised by the authors. 1. The whole population included is 101 patients with early RA. However 107 patients were not included in the basis of data unavailable for the analysis. This proportion of withdrawals seems to me very high and a selection bias may be important. . Are the baseline characterisics of theses patients similar to thodse observed in the group of patients finally included?. 2. There are 12 patients in remission at sudy entry. How do the authors coud explain this finding if these patients have not initiated DMARD therapy?.I think that only patients with active disease at baseline should be included in the study 3. Although it is mentioned in the discussion section, the decrease in inflammatory activity after only 4 weeks is very surprising. How the authors could explain this rapid response to DMARDs? 4. The use as the cut-off of the median of DAS28 reduction at 4 weeks for the statistical analysis is not clear. Pehaps a decrease in 0.6 according to EULAR criteria may be more informative. Minor comments -DAS28 is preferable than the other definitons. - please provide if there are differences in DMARDs used between patients with and withouth significant reductions of DAS28 at week 4. - what does it means exactly interaction CRP-weight? - Please provide the disease duration (months) in Table 1 and introduce this variable in the regression analysis. - RF 88.1% and ACPA 50.5%. This difference in percentages is very surprising. - Data at week 4 are not presented in table 2 and may be of interest. - Information raised in Table 3 may be ommited. - anti CCP> 6: can the authous explain this cutoff and provide the methods used for CCP determination.

REVIEWER	Brooks, Peter University of Melbourne
REVIEW RETURNED	15-May-2013

GENERAL COMMENTS	It is good - provides new data and confirms other data . The authors should be asked to speculate why such rapid responses are seen . Also do they have any data comparing the responses of patients who were seen within 4 weeks of symptoms to those who were first seen after a longer period . The Guidelines of the Australian Pharmaceutical Benefits scheme for use of these agents should be at least referenced if not summarised in the text

REVIEWER	Dr Arvind Chopra, Md

	Director and Chief Rheumatologist, Center for Rheumatic Diseases, Pune, India. I have no conflict of interest.
REVIEW RETURNED	16-May-2013

THE STUDY	This study is really exploratory in answering the research question of prediciting response to such an intensive DMARD combo in patients with very early RA. A control is required.
RESULTS & CONCLUSIONS	These were early RA patients with not so moderately severe disease (moderate DAS, ESR and CRP). Early RA can responds favorably to somewhat minimal regimens (with pulse steroids). As of now, preciese prognostication for very ealry RA is difficult and there are several publications that do not support combo DMARD. THis is a good clinic based study with limited sample size which can at the best generate hypothesis for controlled evaluation.
GENERAL COMMENTS	A good real to life concept and approach. I think one need to remember that these patietns had fairly early RA and in this cohort prognostication is difficult and over zealous therapeutics common. Such a study needs control. Also the baseline parameters were that of moderate activity diseases. Kinldy reflect over the role of steroids in your study. Do you thing mon DMARD with pulse steroid would suffice. You ought to blend your conclusions with more humility based on well accepted and know limitations of such an observational study. At its best, your study can generate hypothesis. The control should be for early RA irrespective of response to DMARD in your kind of strategy. You could do a subset analysis for patients with high DAS.

VERSION 1 – AUTHOR RESPONSE

Reviewer: Catherine Louise Hill
Comments to the Author
I have outlined in the reasons in the attached file which addresses the deficient areas of the STROBE checklist. My main areas of reticience is the likelihood of selection bias related to this study created by the exclusion of >50% of the cohort due to inadequate data collection.
This manuscript requires a statistical reviewer before consideration for publication.
It is an interesting premise and one that influences daily practise in treating RA patients.
Resp: See response to point 6 below.
This is a highly regarded research group so the results are 'credible' but generalisability of the results and impact on clinical practice is unlikely, rather this work is hypothesis generating about decision making in the treatment of RA.
Resp: The discussion has been modified to reflect this on page 14.
Attached Strobe checklist comments
This is an interesting manuscript that raises the possibility that response to DMARD therapy at a time point as early as 4 weeks predicts treatment response at 12 months. This differs from other previous analysis which have looked at 3 months as the initial response time. However, due to the large number of participants excluded (>50%) from the initial data set due to inadequate data, the risk of bias is high. In addition, no comparative analysis between those included and excluded has been undertaken. I have further comments regarding the manuscript which I have outlined below as part of the STROBE Statement Checklist (To be included in reports of cohort
studies). The STROBE checklist did not appear to be in the files that I received. Further information should be included under the following checklist item numbers:
No 5. Setting. Further information about the referral pattern of the clinic would be helpful.

Resp: revised page 6
No 6. Further information about the eligibility criteria. Patients were recruited consecutively from the Rheumatology Clinic. Did participants give individual written informed consent? How was 'active' RA defined? More than 50% of participants were excluded due to lack of 12 month data. Did this mean that participants had to have had visits at each of the 9 visits to be included in this analysis?
Resp: The previous version was misleading, and excluded patients in fact comprised 37% of those diagnosed with RA, as 49 of the 206 patients referred to the clinic did not have RA. Methods description now revised, pages 6-7.
No 7. Further definition of potential confounders. Was symptom duration at first visit recorded? Although it is within the scope of good clinical practice in an RA clinic to measure BP, it does not necessarily follow that this is a confounder. In addition, disease duration is a potential confounder and the duration of symptoms was not included in the models. Is there any information on SES variables which may impact disease outcomes in RA, in addition to the ability to present at each visit?
Resp: Symptom duration was recorded and was not associated with change in DAS at week 4 from baseline. The level of education achieved was recorded as a measure of SES and was significantly associated with reduction in DAS (text revised pages 8, 12, 16).
No 9. Bias. No description of any efforts to address potential sources of bias was addressed. As more than 50% of patients were not included in the analysis, this is potentially a major source of bias. A statistical comparison needs to be made of those who were included in this time-dependent analysis and those who were excluded due to inadequate data points for analysis.
Resp. Unfortunately a subanalysis of the 49 excluded RA patients who had a baseline visit and at least one other visit was not possible due to the small number of paired baseline and 4w DAS measurements in these patients. A comment has been made on page 14 to this effect.
No 10. No sample size calculation is included. It is not clear why the investigators decided to cease data collection at this time (2008).
Resp. Due to the exploratory nature of this study it was not possible to undertake a sample size calculation. Data collection ceased due to insufficient funding after 2008.
No 11. The statistical methods appear sound and overall well reported. However, I am not familiar with all of the statistical methods that have been used and would recommend a separate statistical reviewer. It states that imputations were made for missing data. However, it is not clear from the methods how much missing data was allowed in this study for the participants to be included.
Resp. A comment has been included page 9.
No 13. No flow diagram was included and may have been useful. Almost 50% of the cohort were excluded due to incomplete data and a flow chart would help the reader interpret the result. A flow chart would also allow authors to show how many participants progressed to other DMARDs and biologic therapy during the study.
Resp. Information on exclusions and reasons are included on page 7. As decision making is dynamic regarding changes to DMARDs according to target criteria, changes would be very complicated to include in a flow chart.
No 14. Descriptive data. A comparison of the baseline characteristics of the participants who were included in the final analysis and those who were excluded due to incomplete data observations would be appropriate. Table 1 does not explain how smoking was defined. I am not clear the relevance of BP, eGFR and glucose to the current analysis. I understand that cardiovascular risk factors were of importance to other analyses performed on this dataset but
probably do not need to be included here. There is no mention of other co-morbidities. This is likely to be of relevance in attending the clinic and ability of the clinician to escalate therapy.
There is no indication of the number of participants with missing data for each variable of interest.
Resp. Smoking ascertainment is outlined on page 8. Descriptive data for excluded patients now included in Table 1.
No 15. Outcome data. Disease outcomes were included. However, no adverse event data was included. It is possible that those who were excluded due to lack of availability of 12 month data had a different disease trajectory and have not been included in this study and increase bias.

Resp. A comment has been added to page 14.
No 16. Main results. Unadjusted comparisons were shown in Tables 4 and 5. These show the relevant regression analyses but do not show the reasoning behind these choices. It is not clear why these confounders were adjusted for in the analysis. A sensitive analysis should be done to look at the 4 week DAS scores of the participants included in this current analysis and those who were excluded for lack of 12 month data (although it wouldn't be the full 107, I am presuming that a % would have had 4 week data).
Resp. Table 4 represents an analysis at week 4 and at week 12 of variables affecting outcome. Table 5 includes baseline and time-varying variables over 52 weeks. Time-varying analysis was not be carried out at the early time points. The suggested sub-analysis could not be meaningfully carried out using these data.
No 17. Other analysis. There is no subgroup analysis undertaken. It is likely to be of interest to compare those with <6 and >6 months symptom duration at presentation. The analysis looking at factors affecting the response trajectory (page 11 of 31) only appears to include a very limited number of factors and do not include disease duration. Could the authors explain why these were chosen and other clinical characteristics not?
Resp. We have now included symptom duration and education level in the analysis (page 12). Characteristics chosen were based on clinically meaningful variables appropriate to the timepoint analysed (as above). Symptom duration and education level have now been included also.
In addition, as 22% already had low DAS (<=3.2) at baseline, a subgroup analysis excluding these patients may be helpful to further interpret the results.
Resp. The regression analysis with time-varying risk factors was adjusted for the baseline DAS, thus eliminating differences contributed by baseline DAS.
No 18. Limitations. This needs further discussion particular in terms of bias introduced by excluding those with incomplete data observation. Limitation in terms of use of 1987 classification. The slow incremental methotrexate actually means that patients were only taking 10mg/week at first visit (4 weeks) which may be subtherapeutic.
Could the steep drop in the DAS scores at 4 weeks be explained by regression to the mean? And the fall over the 3 months is a more reliable and accurate measure. Patients with RA are most likely to present to GP and gain specialist referral when disease activity at its maximal. The proportion of patients with low disease activity did not significantly alter between 6 (25%) and 12 months (29%). Could the authors comment/speculate on reasons for this in the setting of a TTT strategy.
No 21. Generalisability. This is discussed but is a major limitation of the work.
References:
Reference 3. Bakker et al. is incomplete
The reference list does not appear to be consistent with the way that BMJ Open presents references.
Resp: Discussed page 13-14, and 15. Corrected referencing

Reviewer: Raimon Sanmarti MD
This is a study analysing the time-dependant therapeutic response to a DMARD combination strategy using the treat to target aprroach in patients with early RA. Although the study is interessting I have some major concerns that limits the conclusions raised by the authors.
1. The whole population included is 101 patients with early RA. However 107 patients were not included in the basis of data unavailable for the analysis. This proportion of withdrawals seems to me very high and a selection bias may be important. . Are the baseline characterisics of theses patients similar to thodse observed in the group of patients finally included?.
Resp: see response to first reviewer above
2. There are 12 patients in remission at sudy entry. How do the authors coud explain this finding if these patients have not initiated DMARD therapy?.I think that only patients with active disease at baseline should be included in the study 3. Although it is mentioned in the discussion section, the decrease in inflammatory activity after only 4 weeks is very surprising. How the authors could explain this rapid response to DMARDs?

Resp: DAS <2.6 is defined as minimal residual disease and not remission. It is possible (indeed not uncommon in early RA patients) to have sufficient joint count/VAS/ESR to have a diagnosis of RA but still have DAS <2.6. Thus these patients do have active disease, but this activity is at a low level. The explanation is likely due to a combined mechanism of action of the drugs and the fact that the patients are in the early stage of disease, which is most responsive to treatment (added comment page 17).
4. The use as the cut-off of the median of DAS28 reduction at 4 weeks for the statistical analysis is not clear. Pehaps a decrease in 0.6 according to EULAR criteria may be more informative.
Resp. This is a cut-off determined based on the observed distribution of the DAS, simply for the purpose of stratifying response outcome in this treated population. The EULAR criteria are unlikely to be suitable for this skewed patient population, limiting the robustness of the inference.
Minor comments
-DAS28 is preferable than the other definitons.
Resp. The DAS28 was used in this study.
- please provide if there are differences in DMARDs used between patients with and withouth significant reductions of DAS28 at week 4.
Resp. All patients received the same protocol of triple DMARD therapy between 0 and 4 weeks
- what does it means exactly interaction CRP-weight?
Resp. It means that the joint dynamics of changing weight and changing CRP over time are interacting to affect the outcome i.e. DAS.
- Please provide the disease duration (months) in Table 1 and introduce this variable in the regression analysis.
Resp. Done, see above
- RF 88.1% and ACPA 50.5%. This difference in percentages is very surprising.
Autoantibody frequencies vary between populations e.g. related to genotype and this may explain the difference. The anti-CCP were tested by our clinical laboratory on a previous generation of the test, which was less sensitive than the generation currently used, and this might also explain the relatively low ACPA frequency.
- Data at week 4 are not presented in table 2 and may be of interest.
- Information raised in Table 3 may be ommited.
Resp. Week 4 data are presented in Figs 1 and 2. We would like to keep Table 3 for completeness
- anti CCP> 6: can the authous explain this cutoff and provide the methods used for CCP determination.
Resp: included page 8.

Reviewer: Peter Brooks
It is good - provides new data and confirms other data . The authors should be asked to speculate why such rapid responses are seen . Also do they have any data comparing the responses of patients who were seen within 4 weeks of symptoms to those who were first seen after a longer period .
Resp. Please see response to both questions above
The Guidelines of the Australian Pharmaceutical Benefits scheme for use of these agents should be at least referenced if not summarised in the text
Resp. Included (page 13)

Reviewer: Dr Arvind Chopra, Md
This study is really exploratory in answering the research question of prediciting response to such an intensive DMARD combo in patients with very early RA. A control is required.
These were early RA patients with not so moderately severe disease (moderate DAS, ESR and CRP). Early RA can responds favorably to somewhat minimal regimens (with pulse steroids). As of now, preciese prognostication for very ealry RA is difficult and there are several publications that do not support combo DMARD. This is a good clinic based study with limited sample size which can at the best generate hypothesis for controlled evaluation.

A good real to life concept and approach. I think one need to remember that these patietns had fairly early RA and in this cohort prognostication is difficult and over zealous therapeutics common. Such a study needs control. Also the baseline parameters were that of moderate activity diseases. Kinldy reflect over the role of steroids in your study. Do you thing mon DMARD with pulse steroid would suffice. You ought to blend your conclusions with more humility based on well accepted and know limitations of such an observational study. At its best, your study can generate hypothesis. The control should be for early RA irrespective of response to DMARD in your kind of strategy. You could do a subset analysis for patients with high DAS.
Resp: We have increased the discussion on the limitations of the study and indicated that it is only hypothesis-generating. Please see response re subanalysis to reviewer 1, point 17.

VERSION 2 – REVIEW

REVIEWER	Raimon Sanmarti, MD Rheumatology Service Hospital Clinic Barcelona, Spain No conflict of interest is declared.
REVIEW RETURNED	24-Jun-2013

- The reviewer completed the checklist but made no further comments.