Author's response to reviews

Title: How age and sex affect the erythrocyte sedimentation rate and C-reactive protein in early rheumatoid arthritis.

Authors:

Liseth Siemons (L.siemons@utwente.nl)
Piet M. ten Klooster (p.m.tenklooster@utwente.nl)
Harald E. Vonkeman (h.vonkeman@mst.nl)
Piet L.C.M. van Riel (Piet.vanRiel@radboudumc.nl)
Cees A.W. Glas (c.a.w.glas@utwente.nl)
Mart A.F.J. van de Laar (m.vandelaar@mst.nl)

Version: 2 Date: 9 May 2014

Author's response to reviews: see over
Referee number: 1

Referee's comments to the author(s)

This study attempts to determine the impact of age, gender and BMI on the erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP) in early rheumatoid arthritis.

The findings are of interest but such studies are challenging to perform well given all of the potential confounders. Indeed, this study has a number of limitations, listed below by section in the manuscript.

1. Introduction: “Acute phase reactants are commonly used to measure the extent of inflammation in rheumatoid arthritis (RA)...” I would recommend revising to qualify this since no acute phase reactant can be or should be used as a sole measure of disease activity; it might be better to say “Acute phase reactants are commonly used as a measure of inflammation in RA”

* Authors’ response: We agree and changed it accordingly.

2. Throughout - The language is a bit colloquial: examples – Intro: “...because these cause the red blood cells to settle down more rapidly...” – can delete the word “down”; Intro: “Although the test is easy to perform and relatively cheap...” Would revise to say “inexpensive to perform” rather than “cheap” Discussion – “How this relationship works remains unclear.” Better to say “A thorough understanding of this relationship remains uncertain” or something similar.

* Authors’ response: We agree and changed the examples accordingly and checked the full text for other colloquial phrases.

3. Throughout – the phrases “concentrations of ESR or CRP” or “ESR and CRP concentrations” are used throughout the text; but there is no “concentration” of ESR so it might be better to use the term “levels” or “results of” or “values” instead of concentration.

* Authors’ response: We agree and changed it throughout the paper into “levels”.

4. Intro: “ESR levels respond slowly to inflammatory stimuli and, thus, to changes in disease activity...” – It is true that it may take several days for ESR to change, but I would not consider that “slowness” a limitation of the test – in RA, it is not common to need assessment of inflammation that is less than a few days in duration and repeated assessments within several days are rarely necessary.

* Authors’ response: We understand the reasoning of the referee and we would like to clarify ours as well. As current RA treatments emphasize an early and aggressive approach to suppress RA disease activity as fast and as complete as possible, following strict monitoring and treatment strategies, the slow response of the ESR levels is often considered a limitation. We have rewritten the phrase to give the statement less emphasis and to clarify it more thoroughly:

“Given the current emphasis on strict and aggressive treatment strategies (Smolen et al., 2010; Combe et al., 2007) that suppress RA disease activity as early, fast and complete as possible
5. Intro: “ESR levels can be greatly influenced by, for instance, infections, malignancies.” This is mentioned as a limitation yet the ESR is likely reflecting inflammatory mediators in these conditions so it is not really a limitation of the ESR. It is doing exactly what it is “designed” to do! Rather than pose as a limitation of the test, one could say that the test’s non-specificity is a limitation (though also true of CRP and other measures of inflammation).

* Authors’ response: We understand the confusion here, and removed the part “in addition” to clarify that the non-specificity is the actual limitation of the test when measuring a patient’s disease activity. Now it reads as follows: “Also, because the ESR is a non-specific acute phase reactant of systemic inflammation, elevated levels are not necessarily (solely) due to the inflammation of the rheumatic disease. It has been shown that ESR levels can be greatly influenced by...”

6. Intro: “Most recent studies tend to favor CRP over ESR in assessing RA inflammation, as it is believed to give a better reflection of current disease activity than ESR because of its more rapid response to increases or decreases in inflammatory stimuli [4, 5, 16].” – Ref. 5 is from 1991 so would not consider it a “recent” study. Given the selected references cited, this is probably an overstatement.

* Authors’ response: We are sorry for the confusion here, since we included all the references at the end of the sentence. However, ref 16 is the ref that belongs to “recent”, so we moved that reference to the beginning of the sentence. Furthermore, to diminish the strength of this statement, we changed it into “Many studies tend to favor...”.

7. Intro: “As current RA treatment guidelines strongly emphasize early and aggressive treatment aiming at fast remission, optimal measurement of inflammation in this patient group is becoming increasingly important [24, 25].” This is not necessarily true. Ref. 24 and 25 argue for treating to target and early/aggressive treatment; they do not argue that we need better inflammatory markers. In fact, Ref. 16 makes this statement in its abstract: “Neither test (neither CRP not ESR) adds significantly to clinical measures of disease activity including joint counts and global assessments.” So, the authors’ premise of needing rapidly responsive and more accurate inflammatory markers is not compelling.

* Authors’ response: The reviewer is correct in this comment and these references should have been placed earlier in the sentence. We changed this accordingly. Nevertheless, given 1) this current emphasis on early and aggressive treatment, and 2) the fact that RA is an inflammatory disease and that a state of remission means that this inflammatory disease is “in rest”, we think it is not an overstatement to say that we need optimal measurement of inflammation. Although the acute phase reactants are not sufficient on their own, they are important and informative.

8. Methods: “...symptoms for less than a year. Only patients who had a valid baseline measure of both the ESR and the CRP were included for analyses.” The authors should consider/discuss how patients without ESR and CRP results were different (or similar) to those included, hopefully backed by data (including those with neither test, one test or the other test). The authors state that data was generated from daily practice; can they describe how often and/or why both tests were routinely
ordered? Were the ordering clinicians aware of this study from its planning stages or during patient recruitment? Did the clinicians tend to order both tests when the diagnosis wasn’t certain?

* Authors’ response: Data was collected during daily clinical practice, following a prescribed protocol. We included two references to this protocol because it has already been described elsewhere: “Patients followed a treat-to-target treatment protocol that aimed for a fast remission (DAS28 <2.6). The protocol has been described elsewhere (Vermeer et al, 2011; Siemons et al, in press).” Additionally, only 37 patients had incomplete data on one or both of the laboratory measures. We did check whether they differed from the included study sample and this was not the case for any variable. We added this to the method section: “Only patients who had a valid baseline measure of both the ESR and the CRP were included for analyses, which resulted in the exclusion of 37 patients. Independent t-tests, Kruskal Wallis tests, and Chi-Square tests showed that they did not differ from the other patients in any of the collected baseline measures that are described in the next paragraph.”

9. Methods – “...had not used DMARDs or prednisolone before.” Do the authors mean any corticosteroids when mentioning prednisolone? Did patients have no steroid treatment of any sort? Would revise to clarify this. If injectable/parenteral or oral corticosteroids of any sort were allowed, this could affect results and should be addressed.

* Authors’ response: The way it is written is correct, we really mean prednisolone. We agree that the potential use of certain medicine might have confounded the results and this should be addressed, though medication use as such was not part of our main objectives. Therefore, we included this as a study limitation in the discussion section: “A possible limitation of the current study might be the exclusion of other potential confounding variables, such as certain lifestyle factors (e.g. smoking, physical activity), dietary patterns, or medication use (e.g. estrogen, steroids, or NSAIDs).”

10. Methods/Patients – there is no mention of RA criteria or even diagnosis by a rheumatologist – how does the reader know these patients had RA or how generalizable the data are compared to their own RA population? (This becomes particularly relevant given their low TJC, SJC, ESR, CRP, RF-positivity).

* Authors’ response: We agree that this requires clarification. Patients had RA, based on a clinical diagnosis: “This observational multicenter cohort included newly, clinically diagnosed patients with RA, who experienced symptoms for less than a year and who were not in remission at the time of inclusion. Patients followed a treat-to-target treatment protocol that aimed at a fast remission (DAS28 <2.6). This protocol has been described elsewhere.”

11. Measures – “BMI was calculated as the ratio of weight and height (kg/m2)” – this should be revised to state “BMI was calculated as the ratio of weight in kilograms and the square of the height in meters.”

* Authors’ response: We agree and changed it accordingly.

12. Measures – “psychical functioning was assessed...” – should this be psychological functioning?

* Authors’ response: It should be physical functioning. We corrected this.
13. Statistical analysis – Obese individuals can have a range of BMI values so it would have been useful to have more BMI categories (e.g., >30-35, >35-40, etc) with which to analyze CRP and ESR, especially since so many study subjects were overweight. The same applies to age: <55 and >65 includes large ranges.

* Authors’ response: This would have been interesting indeed, but the groups became too small if more categories were introduced (for instance, there are only 23 patients in the category 30-35, and only 4 in the category >40 (we included this). Given the Dutch nationality of the sample we choose to split the group into 3 categories, based on commonly used cut-off points in the Netherlands. Likewise, we choose to divide age into logical, and often used, groups. These cut-off points resulted in 3 groups that were roughly the same size and were large enough to perform the analyses (n=231, n=175, and n=183, respectively).

14. Multivariate analyses – “The unstandardized betas (0.017 vs. 0.009) represent the change of the natural-log-transformed ESR and CRP concentrations, respectively, for each year of aging. Furthermore, women demonstrated average ESR levels that were significantly higher than those of the men (β=0.198, p=0.007), whereas men had significantly higher CRP levels (β=-0.182, p=0.048).” This should be revised so that the average reader can understand the impact of age and gender on results of ESR and CRP. Means would be helpful here; similarly, would consider adding “for each decade of age increase, the CRP and ESR increased by xx and yy respectively” or something similar.

* Authors’ response: We agree with the reviewer that we could clarify the meaning of these number and modified the paragraph as follows: “The unstandardized betas (0.017 vs. 0.009) represent the change of the natural-log-transformed ESR and CRP levels, respectively, for each year of aging. Transformed back to normal values, this means that for each decade of aging the ESR and CRP levels become 1.19 and 1.09 times higher, respectively. Furthermore, women demonstrated average ESR levels that were 1.22 times higher than those of the men (β=0.198, p=0.007), whereas men had 1.20 times higher CRP levels (β=-0.182, p=0.048).”.

15. Results – “….DAS28 score > 2.6 (i.e. the cut-off point for remission)…” A DAS over 2.6 suggests they were not in remission, NOT that they had active disease; in fact, it seems these patients had rather limited disease with only a few symptomatic joints and a normal average ESR and unimpressive CRP mean at baseline.

* Authors’ response: We agree that a DAS28 over 2.6 means that the patient is not in remission and rephrased the sentence: “… a DAS28 score > 2.6 (i.e. the cut-off point for not being in remission),…”

16. Discussion – “Age and gender are independently associated with the concentrations of both acute phase reactants in early RA, although the effects appear to be strongest for ESR.” The results are stated as a fact; it would be better to say that “In this study, age and gender were independently associated with elevations of CRP and ESR…” since it is rare that a single, imperfect study establishes truth.

* Authors’ response: We agree that we stated this too strongly as a fact and included “in this study” to the sentence.
17. Discussion – If this study’s conclusions are accepted, how should the average reader adjust DAS scores or even individual ESR or CRP results to adjust for age and/or gender? The variation in ESR and CRP by age and gender seem modest (per Table 4). It would be helpful to provide the reader some context.

* Authors’ response: We agree that it would be good to provide the reader some ideas, so the following is included in the conclusion section:

“It could also be argued to develop modified DAS28 scores, including an adaptation for age and gender because these are two risk factors which cannot be modified. Miller et al. [11] already proposed a simple formula for calculating age-adjusted ESR values. However, this was several decades ago, so supplementary research on the possible incorporation of gender and age adjustments in the DAS28 formula is recommended. Another solution might be to specify age and gender specific thresholds of current disease activity scores in order to make them comparable across patient subgroups of different age and gender. Yet to determine the values of such thresholds, further research is required.”

The actual development of such a new formula or such adjusted thresholds was beyond the objectives of this study, but these proposals might provide the reader some new ideas for future studies.

18. Discussion – “Perhaps the underlying processes of early RA, certain lifestyle differences, the higher prevalence of obesity in the women, hormonal factors, or differences in metabolic risk factor [10, 11, 19] might (partly) explain the higher baseline CRP levels in men in this study.” These are reasonable confounders to consider but the potential for medication treatment as a potential confounder is barely mentioned (bottom of page 8; only estrogen is mentioned); while inclusion criteria prohibited DMARD treatment, non-steroidal anti-inflammatory drug (NSAID) treatment (including low dose aspirin) and many other drugs should be considered as important potential confounders – Did the data collected include non-prescription medications (including topical steroids, inhaled steroids, other medications that might have anti-inflammatory effects)? Perhaps men or overweight individuals took NSAIDs less; perhaps low dose NSAIDs have an impact on CRP but not ESR (or vice versa) – The potential for drug-related effects is an important limitation that should be mentioned and addressed.

* Authors’ response: We agree that drug related effects are also reasonable confounders to consider. Yet our aim here was not to provide an exhaustive list of all possible confounders. As rightly raised by the reviewer, many possible confounding factors can play a role. We did include the exclusion of medication as a confounder as a limitation of the study and mentioned the potential effect of steroids and NSAIDs as well):

“A possible limitation of the current study might be the exclusion of other potential confounding variables, such as certain lifestyle factors (e.g. smoking, physical activity), dietary patterns, or medication use (e.g. estrogen, steroids, or NSAIDs).”

Unfortunately, data on non-prescription medications was not collected.

19. Discussion – “Supplementary post-hoc analyses did indeed show significantly more swollen joints, higher inflammatory values, and a more active disease according to the DAS28-ESR score in the two highest age groups (55-65 and 65 years old) compared to the youngest group (<55 years old). “ If this could explain the age-related findings, shouldn’t this temper your conclusions?

* Authors’ response: At this point we tried to give a possible explanation for our findings, it does not change the findings. So, in our opinion it also does not temper our conclusions. A significant relationship with age was found in multivariate analysis while controlling for the other variables of
disease activity that are included in the DAS28. Age is an important factor in relation to the ESR and CRP levels.

20. Discussion – “To evaluate whether the associations changed over time....” And “Follow up analyses including a patient’s rheumatoid factor [9] and physical functioning (as measured with the HAQ) were carried out....” It is best not to include extensive post-hoc analysis in the Discussion section; would either move to results or omit. Only results presented in the results section should be discussed in the Discussion section. Rather than include a detailed discussion of post-hoc analysis at the end, those that cannot be included (the “missing” potential confounders) should be mentioned and addressed as Limitations of the study; the ones that can be included should be included in a re-analysis (see below).

* Authors’ response: We agree that post-hoc analysis do not belong to the discussion section. Therefore, we decided to move the 1 year results to the result section, since these are part of the main objectives of the manuscript. The post-hoc analysis with other possible confounding factors were omitted from the manuscript, as suggested, since these follow-up analysis were not part of the main objectives of our study and don’t belong to the discussion section either. The aim was not to include all potential confounding factors (which are simple too many) but to examine the relationship of age, sex, and BMI with the ESR and CRP levels while controlling for the other DAS28 variables of disease activity. The exclusion of possible other confounding factors has been included as a limitation of the current study.

21. Discussion – since RF and anti-CCP positivity are well-known to co-migrate with more severe disease, it seems important that these variables be included in the analysis; in this manuscript, only RF is mentioned as part of a post-hoc analysis in the Discussion section.

* Authors’ response: Similar to the previous comment, these follow-up analysis were not part of the main objectives of our study, so we decided to omit these analyses from the manuscript and included the omission of potential important confounders as a limitation of the current study.

22. Discussion – given the post-hoc analysis finding physical functioning as an important confounder, wouldn’t it be best to re-analyze including this variable (along with physical activity, anti-CCP and RF)?

* Authors’ response: Similar to the previous two comments, these follow-up analysis were not part of the main objectives of our study, so we decided to omit these analyses from the manuscript and included this as a limitation of the current study.

23. Discussion – “However, since the inflammatory markers are two of the most reliable components of disease activity measures as the DAS28 (unpublished observations) [31], further research is recommended.” It is inappropriate to include Ref. 31 which is a submitted and as yet unpublished paper (written by most of the same authors as this paper) to back up this contention. And, as presented, this statement is pure opinion; in fact, if the CRP and ESR are so reliable, why would the current study be necessary? This statement should be supported with a different argument (with other references) or omitted.

* Authors’ response: Appropriate remark. Since the study we refer to is not yet published, we omitted this statement.
24. Table 1 – Only 61% of the study population is female; the authors should address why there wasn’t a larger female predominance as found in most other studies

* Authors’ response: This 61% s not very unusual for a (Dutch) early RA population and comparable to the proportion reported in other studies. See for instance:


25. Table 1 – If disease onset was 1 year or less, these patients were quite old for new onset RA compared with most other series; again, the authors should address this.

* Authors’ response: Early RA patients of 57 years old are not very unusual for a (Dutch) early RA population. See for instance:


26. Table 1 – The mean baseline ESR was 21 mm/hr; if the average age was 57, this ESR is normal - that seems unexpected for new RA that is untreated; it raises the question of whether these patients were taking a medication (e.g., an NSAID) that might have lowered their ESR; the baseline CRP is unimpressive as well. And the TJC and SJC are quite low as well for what is billed as new, active, untreated RA. The authors should address this.

* Authors’ response: This is indeed a notable finding and it should be addressed. These findings are not surprising per se, since the patients visit the rheumatologist with complaints (i.e. active disease), but this takes place at a very early stage in the disease course (i.e. the disease does not fully express itself yet). As nicely explained by Hazes and Luime (The epidemiology of early inflammatory arthritis 2011:7;381-390), this can be compared with the growth of a plant. In short, at first the genetic potential for the development of the disease is present, then symptoms gradually develop, and
eventually it can take its definite form and is fully classifiable as RA. Thus, this early in the disease course, the disease of the patients might still be developing, though enough symptoms can be present for a classification and for inclusion in the cohort. After all, despite the low ESR/CRP levels and the reasonably low amount of affected joints, the DAS28 does point to a disease that is not in remission. Nevertheless, we agree that this is a notable finding and addressed this in the discussion section:

“Furthermore, it is a notable finding that the acute phase reactants as well as the number of affected joints were relatively low in the recruited patient sample, even though they were not in remission (DAS28 >2.6). Yet this is not very surprising given the early stage of their disease. The disease might still be developing when the patient comes to visit the clinic, but enough symptoms may be present for classification and for inclusion in the cohort.”

27. Table 3 – If I’m reading the Table correctly, TJC28 is not correlated with ESR but SJC28 is correlated with ESR; and the same is true for CRP – This seems odd for untreated RA and should be addressed.

* Authors’ response: This is indeed remarkable, but not uncommon. It is a well-known problem of the tender joint count that it correlates weakly with other variables of RA disease activity, including variables like pain (Sokka T: Assessment of pain in rheumatic diseases. Clinical and Experimental Rheumatology 2005; 23:S77-S84), but also with variables like the ESR (e.g. Leeb BF, et al.: Disease Activity Measurement of Rheumatoid Arthritis: Comparison of the Simplified Disease Activity Index (SDAI) and the Disease Activity Score Including 28 Joints (DAS28) in Daily Routine. Arthritis Care & Research 2005: 53;56-60 / Leeb BF, et al.: The DAS28 in rheumatoid arthritis and fibromyalgia patients. Rheumatology 2004; 43;1504–1507).

Furthermore, as addressed in the previous comment, the disease might still be developing at the moment patients come to visit the clinic. So this might also (partly) explain the low correlations between the variables at this point in the disease. Thus, although it is certainly remarkable, we don’t think it is an odd observation.
Reviewer number:2

Referee's comments to the author(s)

This paper explores a research question that several others have tackled previously, albeit with conflicting results. Therefore, a strong study is needed to more definitely answer the research question that others have previously addressed. In its present state, this paper does not resolve the issues.

There are a number of things that could be improved to strengthen the results of this study.

Major:
1. Why was BMI categorized instead of analyzed as a continuous variable? Why wasn't a category for underweight (BMI < 18.5 kg/m2) included? BMI relationships are often “U” shaped, so continuous BMI should be analyzed using methods that allow for non-linear associations, such as splines.

   * Authors’ response: Good question. BMI was categorized, because its relationship was not linear (and also not U-shaped). The use of dummy’s is common in such a situation. Given the Dutch nationality of the sample we choose to split the group into 3 categories, based on commonly used cut-off points in the Netherlands. There was no category of underweighted people because there were only 8 patients with a BMI<18.5 kg/m2. (We included this.)

2. What is the basis for the statement at the end of page 6 that the differences between ESR and CRP decreased with age? Is this statement based on statistical tests or just visual inspection of the table?

   * Authors’ response: This statement is based on visual inspection of the table, since the interaction was not significant. To emphasize that this is not a fact, we phrased it as: “... APPEARED to gradually decrease with age” and we added the Table number (Table 4) to make clear that you can observe this decrease in the table.

3. It could be helpful to include columns in Table 1 for BMI groups to show which characteristics may be associated with BMI, since BMI is univariately associated with CRP, but not multivariately associated.

   * Authors’ response: Although we understand the request, it would not be appropriate to include this information in this table, since this table just gives a descriptive table of the population, before the reader has seen any results.

4. In Table 4 some of the SDs are large relative to the means. This can indicate non-normality. Means and inter quartile ranges could be more informative for skewed data.

   * Authors’ response: ESR and CRP levels had indeed non-normal distributions. We changed the mean (sd) into median (IQR).
5. Did the log transformations actually result in normally distributed ESR and CRP data. If not, different analysis methods should be used to allow for nonlinear relationships between ESR, CRP, age and BMI.

* Authors’ response: Yes, it did normalize the distribution, and more importantly, it did normalize the residuals.

Minor:
6. On page 4 under measures, "psychical" should be "physical".

* Authors’ response: That is correct, we changed it accordingly.

7. At the end of page 6, a p-value for interaction should be provided to support the statement that there was no significant association between age and sex.

* Authors’ response: We agree and included the p-value. The p-value for the interaction between age and gender was 0.222 in the ESR-model and 0.665 in the CRP-model, respectively.

8. The term "sex" should be used instead of "gender". Sex is commonly refers to the biological definition of male / female and gender refers to the social definition.

* Authors’ response: We agree and changed it accordingly.
Study Summary: This study attempts to examine which inflammatory marker (ESR vs CRP for use in DAS28 scores) is least affected by age, gender, and BMI in patients with early RA (patients with <1 year symptoms in multicenter Dutch DREAM registry).

General Comments:
Although this relatively brief manuscript is well written and should be of clinical significance to practicing rheumatologists, the present study as designed and written is not necessarily original but may certainly contribute to the medical literature in regards to their stated study objective/title question. However, if this manuscript can be significantly revised to include longitudinal clinical treatment/outcomes data discrepancies depending on the use of ESR vs CRP in DAS28 scores, this expanded research design/methodology then would perhaps make this manuscript more original and clinically relevant to the rheumatology community.

Specific Comments:
1. On page 5/15, Results/Table 1, consider adding in the data for mean RA disease duration and percent positive for ACPA/anti-CCP (cyclic citrullinated peptide) for this early RA cohort.

* Authors’ response: We added the information in Table 1.

2. Although this study cohort may not necessarily be extremely large but likely adequate, may consider confirming/adding in some sample size calculations.

* Authors’ response: We understand why this could have been useful. However, given the descriptive character of the manuscript and the large size of the sample we don’t think this would add much to the current study as it is.

3. On page 17, Table 3, consider adding in the 95% confidence interval for Beta (also for Table 2). May also consider reporting the R2 for the ESR and CRP models as a footnote in Table 3.

* Authors’ response: We included the confidence intervals for the beta’s in Table 2, 3, and 5 and we reported the R2 of the multivariate analyses beneath the tables, as requested.

4. On page 8, Discussion, should either consider integrating the “post-hoc analyses” data into the Results section or providing the data as a Supporting/Additional online file/Table. Similarly, the multivariate analyses data repeated at 1 year should either be integrated into the Results section or provided as a Supporting/Additional online file/Table. Moreover, if this manuscript can be revised to include longitudinal clinical treatment/outcomes data discrepancies depending on the use of ESR vs CRP in calculating DAS28 scores correlated with their multivariate analyses, this may perhaps make this manuscript much more original and clinically relevant to the rheumatology community.

* Authors’ response: We agree that the post-hoc analysis on physical functioning do not belong to the discussion section. In line with the comment of one of the other reviewers, we decided to omit these analyses from the manuscript because these follow-up analysis weren’t part of the main objectives of
our study. We included the omission of possible additional relevant confounders as a limitation of the current study. The aim was not to include all possible confounding factors (which could be enormous). Instead, we were particularly interested in the relationship of age, sex, and BMI with the ESR and CRP levels.

We did, however, move the 1-year multivariate analyses to the result section as suggested, because these are associated with our study objectives.

5. A few grammatical corrections: On page 4, change “psychical functioning” to “physical functioning”. On page 7, consider changing “Ethnicity related differences across population … explain these deviating results.” to “Ethnic differences across population … explain these diverse results.”

* Authors’ response: We agree and changed it accordingly.
Reviewer number: 4

Referee’s comments to the author(s)

Comments to the author:
This paper aims to determine whether ESR or CRP are more affected by age and gender in early RA, in a cross-sectional study carried out in patients enrolled in a Dutch early RA cohort, the DREAM registry. Associations between ESR and CRP, and age, gender, and BMI were evaluated, and the relation of the associations to disease activity was further evaluated. This is important to understand given the prominence of inflammatory markers in assessing RA disease activity and the importance of accurately assessing disease activity. The major finding of the study was that age and gender, but not BMI, were associated with ESR and CRP levels, independent of disease activity. However there are some concerns that need to be addressed before this is ready for publication.

Major concerns:
1. The methodology appears appropriate to the question, and several of the patients’ baseline characteristics are clearly described, however smoking and use of steroids were likely collected, could have influenced the results and were not included. What was the process to decide what variables should be included, as the question of interest seems to be what affects acute phase reactants in early RA.

  * Authors’ response: The aim of the current study was not to include all possible confounding factors (which are simple too many), but to examine the relationship of the (relatively) stable factors of age, sex, and BMI with the ESR and CRP value, while controlling for the other variables of disease activity that are included in the DAS28. Nevertheless, we do recognize smoking and medication use as possible other confounding factors and included their exclusion as a limitation of the current study.

2. A major finding of the study demonstrates that age and gender but not BMI were independently associated with CRP and ESR levels. However, the mean of the highest BMI group was only 33. When presenting the demographic data, suggest describing your population based on proportions that fit into the WHO classification categories including the highest (BMI>40) group and the very low BMI (BMI <18.5). The lack of association with BMI may relate to the lack of morbidly obese or very thin patients in this cohort. Additionally please comment if age and BMI were co linear and whether this was why BMI fell out of the model?

  * Authors’ response: No collinearity was found. Given the Dutch nationality of the sample we choose to split the group into 3 categories, based on commonly used cut-off points in the Netherlands. There was no category of underweighted people because there were only 8 patients with a BMI<18.5 kg/m2. Likewise the groups became too small if more categories were introduced (for instance, there are only 23 patients in the category 30-35, and only 4 in the category >40. We did include a paragraph to pay attention to this in the discussion section: “In contrast to previous studies, no independent association between BMI and acute phase reactants was found at baseline. Perhaps the proportion of obese patients in this study was too small to detect these effects (15.2% of the men and 21.7% of the women). Likewise, only 8 patients were underweighted with a BMI <18.5 kg/m2 and only 4 patients had morbid overweight with a BMI >40 kg/m2.”.
3. Detailed analysis of the relationship of components of the DAS was interesting and informative, but other confounders such as smoking are not included. Take for example smoking. Le-Ha et al showed a relationship between smoking and C-RP levels. Le-Ha C. et al. (Gender and the Active Smoking and High-Sensitivity C-Reactive Protein Relation in Late Adolescence. J Lipid Res. 2014 Feb 27.). Presumably the intent was to find out what possible factors could influence acute phase reactants yet important variables such as smoking were not assessed. Usually an assessment of smoking is part of usual care. The authors need to confirm if this was collected and if so suggest adding this to the analysis as it would be very interesting if it is related to CRP or ESR.

* Authors’ response: Some data about smoking was collected at baseline, however smoking was not part of the main objective of our study and, as such, was not included in the current data set that was used. The aim was to include all possible confounding factors, but to examine the relationship of age, sex, and BMI with the ESR and CRP value while controlling for the other variables of disease activity that are included in the DAS28. Nevertheless, we do recognize smoking as a possible other confounding factor and included its exclusion as a limitation of the current study.

4. Similarly patients often come into studies like these having been exposed to some therapy, most often steroids. This is another co-variate that could easily be added to the analysis. Was this considered or possible? If you only used patients not exposed to drugs this should be clarified. If patients who had been on steroids were used this could have affected CRP and ESR causing you to miss association with other variables. It seems these is probably sufficient statistical power to examine other potential confounders (e.g. steroids use, smoking, co-morbidities). Any variables that are collected in DREAM that could affect ESR or CRP should have been considered and been included in the model.

* Authors’ response: Consistent with our previous response, the aim was not to include all possible confounding factors (which are simple too many), but to examine the relationship of the (relatively) stable factors of age, sex, and BMI with the ESR and CRP value, while controlling for the other variables of disease activity that are included in the DAS28. Nevertheless, we do recognize medication use and smoking as possible other confounding factors and included their exclusion as a limitation of the current study.

5. Do you think the relationship between other components of the DAS was lost because of the relatively low disease activity in your subjects? What proportion were in different disease activity states? This relates back to the question of therapies patients were exposed to at the time of study entry, which is presumably when this analysis was done.

* Authors’ response: The non-significance of the TJC is indeed remarkable, but not uncommon. It is a well-known problem of the tender joint count that it correlates weakly with other variables of RA disease activity, including variables as, for instance, pain (Sokka T: Assessment of pain in rheumatic diseases. Clinical and Experimental Rheumatology 2005; 23:S77-S84), but also with acute phase reactants like the ESR (e.g. Leeb BF, et al.: Disease Activity Measurement of Rheumatoid Arthritis: Comparison of the Simplified Disease Activity Index (SDAI) and the Disease Activity Score Including 28 Joints (DAS28) in Daily Routine. Arthritis Care & Research 2005: 53;56-60 / Leeb BF, et al.: The DAS28 in rheumatoid arthritis and fibromyalgia patients. Rheumatology 2004; 43:1504–1507). Also, given the early state of the disease, the disease might still be developing at the moment patients come to visit the clinic. This is nicely explained by Hazes and Luime (The epidemiology of early inflammatory arthritis 2011:7:381-390). They state this can be compared with the growth of a plant. In short, at first the genetic potential for the development of the disease is present, then symptoms gradually develop, and eventually it can take its definite form and is fully classifiable as RA. Thus, this
early in the disease course, the disease of the patients might still be developing, though enough symptoms can be present for a classification and for inclusion in the cohort. Likewise, correlations might still be building up, which might also (partly) explain the low correlations between the variables at this point in the disease. So, although it is remarkable and it might be due to the early stage of the disease or the relatively low disease activity (which still pointed to non-remission with a DAS28 > 2.6!), we don’t think it’s an odd finding. Nevertheless, we added a phrase on the relatively low disease activity (i.e. number of affected joints and relatively low levels of the ESR and CRP) in the discussion section: “Furthermore, it is a notable finding that the acute phase reactants as well as the number of affected joints were relatively low in the recruited patient sample, even though they were not in remission (DAS28 >2.6). Yet this is not very surprising given the early stage of their disease. The disease might still be developing when the patient comes to visit the clinic, but enough symptoms may be present for classification and for inclusion in the cohort.”

Minor considerations:
1. Title: Initially BMI was also included as a covariate of interest but his becomes non-significant yet it remains in the title. The title thus becomes misleading and BMI should probably be removed. Suggest modifying the title: How age, gender affect the erythrocyte sedimentation rate and C-reactive protein in early Rheumatoid Arthritis.

* Authors’ response: Good suggestion. We changed it accordingly.

2. Introduction-- although the rationale for the study is argued well, it would be useful to clearly state the research question in the introduction/background section.

* Authors’ response: We included the aim of the study at the end of the background section: “To date, however, it remains unclear which inflammatory marker is affected most strongly by the external effects of age, sex, and BMI in patients with early RA and with this study we aim to provide more insight into these associations.”

3. Methods- the patients are not clearly defined- although they are part of a cohort it would be nice to know if they meet ACR/EULAR 2010/ ACR 1987 diagnostic criteria. Are all patients from baseline. Is this a cross sectional analysis of patients at study entry. We’re patient characteristics compared or representative of the entire registry? How were patients recruited in DREAM? Were they receiving steroids or had they recently received steroids? What proportion if subjects had normal ESR and CRP. What proportion of subjects fell in each category?

* Authors’ response: Patients were recruited in the clinic. The remission induction cohort is an observational cohort, including patients that went to see a rheumatologist in daily clinical practice and that were newly diagnosed with RA, based on a clinical judgment (we added this information to the method section).
All patients had at least a baseline measure, which was used for analyses: “Only patients who had a valid baseline measure of both the ESR and the CRP were included for analyses.” The treatment protocol has been described previously and we added the corresponding reference: “This protocol has been described elsewhere (Vermeer et al., 2011; Siemons et al., in press).”
4. Results—Although the SF-36 results are included, there is little mention of them in the analysis. The discussion becomes too speculative in the comments regarding activity, abdominal fat and increased inflammatory markers. While interesting, there is little presented to support this in this study. The potential inferences drawn from the HAQ results would have more weight if the results were presented more thoroughly in the results section. Wellbeing shows up in a table but is not defined in tables or manuscript. Is this a component of the SF36 or sub component? Was well being included in table 1?

* Authors’ response: We agree with the reviewer that some things became too speculative. Therefore, and because the post-hoc analysis on physical functioning (i.e. the HAQ) were not part of the main objectives of our study, we decided to omit these post-hoc analyses from the manuscript and included the omission of potential important confounders as a limitation of the current study. This makes the manuscript less speculative.

The SF36 physical and mental health scores in table 1 are the two components summary scores of the SF36. Wellbeing is the same as “general health”. For consistency, we made sure that all words “wellbeing” were replaced by “general health”.

5. Discussion—Introducing new longitudinal BMI data in the discussion section is confusing in the absence of a description of the actual data—such as the proportion included in the follow up evaluations.

* Authors’ response: We moved the analyses on the 1-year data to the result section and included the corresponding beta’s and p-values in an extra table (table 5).

6. Do not recommend basing final concluding statement referring to unpublished results.

* Authors’ response: We agree. Since that article is still under review we removed it from this manuscript.

7. Do you think it is reasonable to compare ESR and CRP results from a population of inflammatory arthritis to those results from populations where CRP or ESR are not expected to be high? Have there been other studies looking at CRP and ESR levels in inflammatory conditions that examined what factors might have influence these?

* Authors’ response: This is a relevant question that we would like to address. At first, we tried to connect our results to both “general” literature as well as specific “RA” literature (like the study of Radovits et al. or Ranganath et al.) to give the reader a broad picture of the (previously) found relationships.

In addition to this, we included the following phrase to make clear that a lot of research is performed in other populations than in patients with inflammatory conditions and that this might explain some deviating results: “Nevertheless, the exact underlying mechanisms behind these dependencies remain unknown and all these previous studies were conducted in healthy community populations instead of in samples with active disease.”

8. Minor grammatical and spelling errors should be corrected: e.g. for example in the measures section refer to data as singular yet data are plural. ....data WERE collected...; “psychical function” is probably supposed to be physical function.

* Authors’ response: That’s correct, we changed it accordingly.