Author's response to reviews

**Title:** The diagnostic accuracy of anti-CCP antibody in rheumatoid arthritis: Impact of reporting quality on pooled diagnostic estimates

**Authors:**

Elias Zintzaras (zintza@med.uth.gr)
Afroditi Papathanasiou (apapath@med.uth.gr)
Dimitrios Ziogas (dziogas@bidmc.harvard.edu)
Michael Voulgarelis (mvoulgar@med.uoa.gr)

**Version:** 3  **Date:** 8 January 2012

**Author's response to reviews:**

Comment 1: One of the problems in the manuscript is that only 4 studies are reported before the introduction of the STARD statement. So it is probably difficult to make a comparison before and after. However it would be great to see what items are sufficiently addressed by the studies and which ones need improvement. It would also be helpful to know what journals endorse the STARD statements and which ones do not.

Reply:

In Results, the following reading was added:

“The 103 eligible articles were published in 35 journals during the period 2003-2010 (four articles were published in 2003, i.e. before the introduction of the STARD statement).”

In Discussion, the following reading was added:

“Focusing on the four “pre-STARD” publications, we found similar results regarding the identified overall reporting quality. Seven of the 22 items of the checklist such as the expertise of study’s executers (25%) and blinding of readers (0%), dates of recruitments (25%), presence of recommended flow diagram (25%), time intervals from index tests to the reference standard (25%), adverse events from the performing tests (0%) and handling of indeterminate data (0%) were reported in less than 50% of studies (2 of 4). However in our main analysis, the percentages of reporting of these items were increased after STARD introduction, but there is still room for further improvement.”

Comment 2: It would also be helpful to know what journals endorse the STARD statements and which ones do not.

Reply: In Results, the following reading was added:

“The characteristics of studies included in the analysis are shown in Supplementary Table 1. The 103 eligible articles were published in 35 journals during the period 2003-2010 (four articles were published in 2003, i.e. before the introduction of the STARD statement). Of the 35 journals only 16 (45.7%)
endorse the STARD statement (Supplementary Table 2) in their guidelines to authors."

Comment 3: Another general remark is the evaluation of anti-ccp in the new criteria. Because they are part of the new criteria simple analysis of Se and Sp is out of the question due to circulation bias. It would be great if the authors could address this issue in the discussion.

Reply: In the Discussion, we noticed clearly the different sources of biases that are introduced in the analysis of Se and Sp of diagnostic accuracy studies by adding the following reading:

“However, the findings of the present synthesis (sensitivity of anti-CCP2, 71% and specificity, 96%) are compatible with those of earlier reviews (Nishimura et al [38]: sensitivity, 67% and specificity, 95%, Whiting et al [13]: sensitivity, 67%, specificity, 96%). The overall specificity (Sp) of anti-CCP antibodies for RA does not seem to differ significantly. The observed overestimation of our overall sensitivity (Se) might be due to the introduction of different sources of bias and variation in the design and conduct of diagnostic accuracy studies[39]. Although little is known about the effects of these sources of bias and variation, some of the variability in the Se and Sp between studies investigating diagnostic accuracy of anti-CCP in RA may relate to different cut-off points for positivity, lack of stratification by study design, differences in disease duration, different disease prevalence and severity, potential co-morbidity, distorted selection of participants, absent or inappropriate reference standard, partial verification bias, clinical review bias and observer or instrument variation [40]. However, the empirical evidence about the size and effect of these issues is limited.”

Comment 4: In the introduction no references are provided to other fields that did the same exercise (Pre and Post STARD). It would be helpful to discuss this and make the introduction at bit more logically in the one before last paragraph.

Reply: In the Introduction, the following reading was added:

“Nevertheless, STARD does not assess the actual quality of the research study but the reporting quality, two issues which are not necessarily correlated. Since the publication of the STARD statement in 2003, more than 200 journals mention the STARD statement in their instructions to authors. However, the quality of reporting of diagnostic accuracy studies remains suboptimal and only a slight improvement was detected overtime [23-25]. Applications of the STARD statement guidelines for assessing the quality of reporting in diagnostic accuracy studies, have been conducted in various medical fields such as in the field of diagnostic endoscopy [26], of juvenile idiopathic arthritis in peripheral joints [27], of diabetic retinopathy screening [28], of glucose monitor studies [29], of optical coherence tomography in glaucoma [30], of ultrasonography for the diagnosis of developmental dysplasia of the hip [31] and in the field of screening ultrasonography for trauma [32]. More specifically, in the field of reproductive medicine, Coppus et al. [33] comparing pre- and post-STARD periods demonstrated a poor reporting overtime of some important methodological STARD criteria. In a recent meta-analysis [13], the quality of studies investigating the diagnostic accuracy of anti-CCP antibody in RA, was evaluated by QUADAS,
an additional tool for assessing the methodological quality of diagnostic accuracy studies [34].

Comment 5: On page 7, the last paragraph starts with “Study Quality assessment with STARD”. This should be rephrased to Evaluation of STARD reporting guideline.

Reply: The title of the last paragraph on page 7 was rephrased to “Evaluation of STARD reporting guideline”

Comment 6: It is not exactly clear from the text on page 8 when a study scores positive. May a description of what items are addressed well and what items need improvement would help to understand this. It would also be good to keep the exact phrasing of the STARD items in table 1.

Reply: In the section “Evaluation of STARD reporting guideline”, we describe briefly how we evaluated the quality of included diagnostic accuracy studies. As we stated in the article, we followed the STARD Explanation and Elaboration guidance and we scored as positive the items that were reported with enough details, as the definition of item suggested. We describe comprehensively what items are addressed well and what items need improvement in the section of Results and not in the Methods section. The phrasing of item (14) was changed to the exact phrasing of the STARD items in table 1 and the paragraph was finally changed to the following one:

“All items were investigated in terms of whether they were reported, not whether they were actually carried out during the study. Items were to be scored as positive if they were reported in enough detail as the definition of item suggested, allowing the reader to judge that the item/quality criterion had been fulfilled. For example, in the case of item (14) entitled “The number of participants satisfying the criteria for inclusion that did or did not undergo the index tests and/or the reference standard; describe why participants failed to receive either test”, a flow diagram was strongly recommended. This item was coded as “yes” only when the flow diagram was given or explicitly described (i.e. the number of controls per case was specified and the matching variables were clearly stated). Alternatives responses (apart from “yes” or “no”) and unclear responses to each item were coded as negative responses.”

Comment 7: I am doubting whether you should do a pooled analysis on the median score of the STARD statement. The current analysis still suggests that you use it as an methodological quality instrument rather than a reporting guideline. I understand that you would like to have some feeling of what the impact of poor reporting would be. Are there other ways you could assess the impact of poor reporting?

Reply: The purpose of the pooled analysis on the median score of the STARD statement was to explore the impact of poor reporting. However, we believe that this is the only plausible way to assess the impact of poor reporting.

Comment 8: The heading on page 12 should be “Impact of reporting guidelines…”

Reply: The heading on page 13 was changed from “Impact of study quality on
diagnostic estimates” to “Impact of reporting guidelines”.

Comment 9: The conclusion of the summary is not equal to the conclusion of the manuscript.

Reply: The conclusion of the summary was modified as follows:

“Conclusion: The reporting quality of the diagnostic studies investigating the diagnostic accuracy of anti-CCP in RA highlights the need for further improvement. The endorsement of STARD statement might help us to overcome the different sources of bias introduced by diagnostic studies and minimize the variability of diagnostic accuracy.”