Author’s response to reviews

Title: Predictors for major cardiovascular outcomes in stable ischaemic heart disease (PREMAC): Statistical analysis plan for data originating from the CLARICOR (clarithromycin for patients with stable coronary heart disease) trial

Authors:

Per Winkel (pwinkel@ctu.dk)
Janus Jakobsen (jcj@ctu.dk)
Jørgen Hilden (jhilden37@mail.dk)
Theis Lange (thlan@sund.ku.dk)
Gorm Jensen (gorm.boje.jensen@regionh.dk)
Erik Kjøller (kjoller@dadlnet.dk)
Ahmad Sajadieh (amad.Sajadieh@regionh.dk)
Jens Kastrup (jens.kastrup@regionh.dk)
Hans Kolmos (h.j.kolmos@dadlnet.dk)
Anders Larsson (anders.larsson@akademiska.se)
Johan Arnlov (johan.arnlov@medsci.uu.se)
Christian Gluud (cgluud@ctu.dk)

Version: 2 Date: 04 Jan 2017

Author’s response to reviews:

Re.: DAPR-D-16-00011

Predictors for major cardiovascular outcomes in stable ischaemic heart disease (PREMAC): statistical analysis plan for data originating from the CLARICOR (clarithromycin for patients with stable coronary heart disease) trial

Per winkel; Janus jakobsen; Jørgen Hilden; Theis Lange; Gorm Jensen; Erik Kjøller; Ahmad Sajadieh; Jens Kastrup; Hans Kolmos; Anders Larsson; Johan Ärnlöv; Christian Gluud
Dear Editor,

Thank you for your comments on our manuscript "Predictors for major cardiovascular outcomes in stable ischaemic heart disease (PREMAC): statistical analysis plan for data originating from the CLARICOR (clarithromycin for patients with stable coronary heart disease) trial" (DAPR-D-16-00011).

We are happy that you find it of potential interest and have below responded to the comments raised by your peer reviewer # 2.

On behalf of all authors,

Yours sincerely

Per Winkel

Please find each of the reviewer’s questions and our answers in the following.

Reviewer #2: I thank the authors for their clarifications and additions, which improved the paper. A few things remain unclear for me.

Answer Thank you. We are pleased that you find that the manuscript has been improved. And thank you for all your additional and fine comments. End of answer.

Comment 1- What will be done with treatment arm should be discussed more at the beginning. Now it is only mentioned just before the discussion.

Answer comment 1 Thank you. We totally agree. The topic is now mentioned in the abstract (see page 3) and further discussed in the manuscript on page 10 first section. End of answer

Comment 2 What is the key analysis: the one using only patients in the placebo arm, or the one including all patients?
Answer comment 2 Thank you for turning our attention to this ambiguity. We have decided only to use the data from the placebo patients in our search for new biochemical predictors. See pages 3 and 10. See also page 13 last section and page 14 first two lines. End of answer

Comment 3 What do you mean with the comment on out-of-hospital vs in-hospital deaths?


and found that the effect of clarithromycin is confined to out of hospital cardiovascular deaths. We now refer to this paper. End of answer

Comment 4 - PH and linearity assumptions: will you perform 1 joint PH test, or will you perform a separate test for each covariate? Will you then perform a supremum test for linear functional form for each continuous covariate?

Answer comment 4 We will perform one joint ph test for the covariates and a supremum test for linear functional form for each continuous covariate. See page 10 under the heading ‘The proportional hazards assumption’ End of answer

Comment 5- Definition of C statistic: the authors reply that C is a frequency of correct triage, but their explanation then correctly shows that it is a probability (not a frequency). I still think that the description 'frequency of correct two-patient triage decisions' is very confusing. It is the proportion of all patient pairs where the predicted survival is better for the patient who survived longer (Royston & Altman 2013; Steyerberg 2009).

Answer comment 5 We have changed the definition accordingly. See page 12 second section under the heading ‘Illustration of predictive impact’ End of answer

Comment 6 I struggle with seeing this in the context of a triage decision, because it is a clinically irrelevant situation (as often indicated by Pepe's group).
Answer comment 6 Thank you. We have now confined the discussion to budget constraint issues. See page 12 second section under the heading ‘Illustration of predictive impact’ End of answer

Comment 7- The new section about interaction between sex and other risk factors is not clear. What do you mean with 'in these situations'? Are this based on known interactions from the literature, or is it based on interactions terms added to the models? If so, will you only assess interaction terms for new markers?

Answer comment 7 Yes, we agree, it is far from being clear. It is based on interaction terms added to the model of each individual outcome and the model of the combined outcome. We will for each outcome test for interaction between sex and each clinical predictor, each standard biochemical predictor, and each advanced biochemical predictor using Bonferroni adjusted P values and use a threshold of 0.05. See page 11 first section. End of answer

Comment 8- About the combined analysis: the authors state that "consequences of leaving out some of the selected quantities will be examined". What does this mean? Will some sort of variable selection be done on the identified biochemical markers?

Answer comment 8 No, variable selection will not be attempted. Thank you for pointing out this inconsistency. As we write in the abstract ‘However, due to the potential selection bias we do not feel that it is advisable to try to rank identified biochemical predictors relative to each other nor to use the results for predictive purposes’ See page 11 the last two of the highlighted sections which will now be removed. End of answer

Comment 9 - The authors mention P<0.01 as a criterion: is this only for the combined analysis, or is 1% the alpha level used for all tests?

Answer comment 9 Thank you. It is only for the combined analysis, i.e., the analysis of each advanced biochemical predictor when adjustments have been made by the standard predictors. The hypothesis generating approach concerns the identification of potential biochemical predictors where we have elected to be more liberal. But then any clinically implausible results will be discussed and the conclusions tempered accordingly.

However, to prevent our models from being too data driven we have elected within each of the six outcomes examined to use Bonferroni adjustment when assessing the ph assumption and the functional form of each continuous covariate. The same approach will be taken when assessing the interactions with sex. See page 10 the section on 'proportional hazards assumptions', page 10
section on ‘flow of analyses’ first section, page 11 first section, and page 11 last section. End of answer

Comment 10- Calibration checks: what are 'subpopulations of interest'? More generally, I do not think that calibration is very relevant for exploratory work to identify potentially prognostic markers.

Answer comment 10 Thank you for pointing this out. We agree. We have removed the calibration checks. See page 12 section on ‘calibration checks’ which now will be removed. End of answer

Comment 11- Net benefit paragraph in discussion: the authors state that they are "not able to measure the two types of benefit using the same unit of measurement". However, they will calculate change in true positives and false positives using a risk cut-off of 25%. Then, using the Net Benefit measure from Vickers and Elkin (Med Decis Making 2006; see also Vickers et al, BMJ 2016) this cut-off can be used to combine true and false positives into a single number.

Answer comment 11 We agree that the net benefit method is a relevant tool in the present context. However, the net benefit graph serves to illustrate, rather than solve, the “same units” problem we are facing. A comment on this has been added in the Discussion; see p. 16 third section. End of answer

Comment 12- 'wasteful of insight' is a strange formulation. Also, Bonferroni does not 'disregard all but the lowest p-value'?

Answer comment 12 Thank you. Wasteful of insight means that a mechanical rule only guided by P values disregards the clinical judgment of the plausibility of the results. Bonferroni disregards P values that fulfill the condition that P*m > 0.05 where m is the number of null hypotheses tested. We have elected to remove the discussion. See last line on page 11 and first 4 lines on page 12. End of answer

Comment 13- Regarding missing values: Little's MCAR test will be used. I'm not very familiar with this test, but it seems to assume multivariate normality? Which variables will you include? Irrespective of the result, would it not be better to perform imputation to avoid loss of data?

Answer comment 13 Thank you for making us aware of this problem. In the analysis of each advanced biochemical predictor we will include the laboratory variables (i.e., the standard
biochemical predictors plus the advanced biochemical predictor in question) plus all other variables in the model in Little’s test. This has now been explained.

If missing is MCAR no biases are introduced by restricting to complete case analysis. The multiple imputation in itself introduces some random variability. On the other hand, the increase of the sample size accomplished by the multiple imputation improves the standard errors all things else being equal (which they are not). Furthermore, it is our experience that simpler analyses (i.e., without imputation) are easier to communicate. So if there are plausible reasons to believe that data are MCAR we consider it a definite advantage to avoid the complicated process of multiple imputation where a number of assumptions about the data have to be made and many errors may creep in if one is not quite careful. Please, see page 13 first section.

End of answer