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@dadolnet.dk)
Ahmad Sajadieh (amad.Sajadieh@regionh.dk)
Jens Kastrup (jens.kastrup@regionh.dk)
Hans Kolmos (h.j.kolmos@dadolnet.dk)
Anders Larsson (anders.larsson@akademiska.se)
Johan Arnlov (johan.arnlov@medsci.uu.se)
Christian Gluud (cgluud@ctu.dk)

Version: 1 Date: 24 Nov 2016

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 Arnlov; 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 you and your peer reviewers.

On behalf of all authors,

Yours sincerely

Per Winkel

Editorial comment 1  In general, more details regarding the rationale and statistical analyses (including whether and how clustering of subjects within centres will be addressed) are needed. Besides that, I strongly recommend to specify the role(s) of each individual author at the end of the manuscript. End of editorial comment 1

Answer: We have tried to improve our explanations of the rationale of the study see answers to the reviewers below. Clustering of participants within centres will be addressed by using the SAS instruction “strata center;” This method assumes that the observations are conditionally independent within a center and the coefficients of the covariates are the same across centers (the patients were recruited from five hospitals. Please, see page 12 section 1 in the “Flow of analyses” section of the statistical analyses). The roles of the authors are now specified at the end of the manuscript (please, see page 19 “Authors contribution”). We have followed the European Medical Writers Association (EMWA) guidelines. Answer finished.
Reviewer #1:

The authors describe a statistical analysis plan for the data that were obtained for the CLARICOR trial in patients with stable coronary heart disease (CHD). The availability of register data that cover 10 years follow-up allows the authors to study the incremental value of several biochemical quantities over standard clinical predictors to predict cardiovascular events and all-cause mortality in patients who already had a CHD event.

Comment 1. The use of existing data from a randomized controlled trial has some advantages such as the dedicated collection of blood samples and other data, although it may suffer some problems concerning generalizability because of selective inclusion of patients. In this specific trial, there are no data on a specific group of patients, namely those who entered the stable state and then died before the end of follow-up. This results in a selective group of patients and I believe this might affect the selection of biochemical quantities that have prognostic information on top of standard clinical predictors.

Answer We fully agree and this issue has already been raised in the current manuscript. To make this point even clearer we have now up front in the abstract stressed that the study is an exploratory hypothesis generating study (see page 2 first section and page 17 section 2 ) and the primary objective is to identify potential biochemical predictors. However, note that this reservation concerns the suitability of the CLARICOR data for the purpose of forming prognosis based on clinical data and special tests obtained AT ENTRY into the stable state. If instead, one wants to prognosticate a patient who is seen LIVING IN HIS/HER stable state (say, at a random time point some 1/2 to 6 years after entry to that state), which is equally clinically relevant task, then data obtained at CLARICOR randomization probably match this task. End of answer.

Comment 2. Besides that, the authors describe a very relevant topic where research is certainly necessary. An overview of what is already known on this topic is however lacking. This makes it difficult to put this research into the context of existing knowledge.

Furthermore, I am interested to read something about the availability of the advanced biochemical quantities in current (and future) medical settings.

Answer We are happy to see that the reviewer finds our research relevant and necessary. On page 8 last section page 9 and first section of page 10 an overview is now presented and the availability of the various assays is commented on. The assays used in the studies are all produced by commercial companies and are marketed. Only calprotectin is in late development stage. We expect that the calprotectin will be on the market at the time of publication of data with the calprotectin assay. We have thus added to the manuscript: The assays used in the studies are commercially available (see page 8 the section II advanced biochemical predictors) End of answer.
With respect to their analysis plan, the main topic of this paper, I have some major concerns.

Answer We thank the reviewer for airing these concerns, which we address below. We think it has improved the clarity of our manuscript. End of answer

Comment # 3- I suggest the authors not to select the biomarkers only on significant changes in hazard ratio, but also on their predictive ability. For this they might look at changes in performance measures such as discrimination (e.g. c-statistic or D-statistic) and calibration (e.g. O/E ratio). Furthermore, relatively new performance measures such as net reclassification improvement (NRI) or integrated discrimination improvement (IDI) might be considered to select which biochemical quantities have incremental predictive value on top of standard predictors.

Answer For inferential purposes we will use the coefficients generated by the Cox proportional hazards model as recommended by Kerr KF et al [Net reclassification indices for evaluating risk-prediction instruments: a critical review. Epidemiology. 2014;25:114-121]. Thus we will identify a potentially useful biochemical predictor using this criterion (see page 12 first section in the section “Flow of analyses” in the statistical analysis section). Apart from that we will assess the discrimination (see page 13 first two sections of “illustration of predictive impact”). In each of our applications there are two risk categories (occurrence of event and non-occurrence of event). When there are two risk categories, the components of net reclassification indices are the same as the changes in the true-positives and false-positives rates. As recommended by Kerr et al we will retain these descriptive terms. To assess the calibration we will calculate the survival adapted Hosmer-Lemeshow chi-square statistic (May S and Hosmer DW. Lifetime Data Anal. 1998; 4(2):109-120) for the old model and each new model augmented by a candidate biochemical predictor (see section “calibration checks” on page 13).

In our group we are fully aware of the shortcomings of NRI, IDI and similar performance statistics, and we have been actively engaged in developing the area both theoretically and via Monte Carlo studies [Hilden & Gerds Statistics in Medicine. 2013; Pepe & al. UW Department of Biostatistics Working Paper Series. 2013 Paper 392]. See also Comment # 4. End of answer

Comment # 4- The authors describe no efforts to validate their results, either internally or externally.

Answer We don’t have the data needed for an external validation and we have deliberately refrained from an internal validation of the results since the scope of the paper is a qualitative selection of potentially useful predictors. However, since our material represents a selective
inclusion of patients we know that in all likelihood an external validation would fail – numerically speaking. Furthermore, our intention is NOT that the reader should use our results for predictive purposes in the clinical routine. For such reasons measures of predictive ability will also receive less attention than simple significance calculations. End of answer

Comment # 5- Is it possible to obtain information on blood pressure? I think this is an important predictor of cardiovascular events and should definitely part of the standard clinical predictors.

Answer Unfortunately this information is not available. And neither are BMI, or ventricular ejection rate. We only have interview information about the presence of hypertension yes/no. We admit this is a drawback. This has now been mentioned and emphasized in the discussion section (see page 16 second section). End of answer

Comment # 6- Treatment decisions are usually based on one prediction model. Do the authors also consider analyses in which they combine all outcomes into a composite outcome?

Answer This is a very good point. This approach has now been added to the statistical section (see page 11 last section). We now intend to make one primary analysis which includes all outcomes in combination. If this analysis identifies one or several candidate markers each individual component of the outcome will be analysed. End of answer

Comment # 7- The part about skewness in predictors is not completely clear to me. How can this explain the results of the trial and how will these analyses be done?

Answer The result of the CLARICOR trial was that the intervention significantly increased the cardiovascular and all cause mortality. This may be due to an imbalance between the experimental and the control group caused by chance. Therefore we want to identify all additional prognostic predictors and repeat the comparison between the two intervention groups when these additional predictors have been included as co-variates. This is now described on page 15 last section before the beginning of the Discussion section. End of answer

One apparently minor but very crucial point:

Comment # 8- I believe multiple biochemical quantities can be described in one publication instead of separate publications for every biochemical quantity. I hope the authors consider this and at least remove the sentence saying they plan to describe every separate biochemical quantity in an individual publication (page 8, row 42).
Comment # 9 In summary, I believe the authors describe a relevant topic, although the methods they use are very basic and not up to date according to current standards.

Answer. Thank you. By amending the manuscript due to the comments from the Editor and the reviewers, we hope that we have been able to bring the manuscript up to current standards. End of answer

Reviewer #2:

Comment # 1 This manuscript describes a statistical analysis plan of an interesting topic that nicely fits the scope of the journal. In that sense this protocol submission is valuable, but before it can be acceptable the manuscript needs to clarify and explain several issues.

Answer Thank you for your kind words. End of answer

Comment # 2 1. Overall I think the problems and suggested methods are not carefully described, with a few unclear sentences/paragraphs. Examples:

a. P9: The part from 'the set of candidate predictors' to 'redundancy may be an issue' is unclear. Further, this study is mainly seen as hypothesis generating, and this should be more clear upfront (e.g. in the abstract).

b. The last paragraph on p9 is unclear to me, there seems to be a issue with the measurement of GFR?

c. Are the first 2 paragraphs on p11 still on the issue of selection bias?

Answer The part on candidate predictors has now been removed and rewritten (see page 12 “flow of analyses” first three sections). The fact that the study is mainly seen as a hypothesis generating study is now more upfront presented in the abstract see page 2 first three lines. We want to use creatinine as a marker of renal function while selecting and combining the predictors among the candidate biochemical predictors. Then we will address the problem of GFR measurement within the model evolving from that analysis (see page 14 under supplementary exploration the second section) The first two sections on page 11 in the original manuscript were still on the issue of selection bias (please, see last section on page 16 and first three sections on page 17 in revised manuscript) End of answer
Comment # 3 2. The authors mention that they will control for standard predictors, but I have not seen a description of the evidence from the literature about what standard predictors are in this context. The authors only state that they consider standard predictors to be those demographic/hospital/biochemical variables available in most Western hospital laboratories. This is very thin as an explanation, certainly given the fact that they seem to label >25 variables as standard predictors. This needs further justification.

Answer We now emphasize in the text (see page 7 last section and the first three sections on page 8) that standard predictors is only a collective term used by us in this particular study to refer to those baseline quantities that were available to us during the CLARICOR study and are either established prognostic predictors or proxies of such predictors that are missing in our material. They refer to quantities generally considered to be of prognostic value [Wong ND, Epidemiological studies of CHD and the evolution of preventive cardiology. Nat. Rev Cardiol 2014; 11: 276-289] i.e., the clinical quantities (sex, age, smoking history, history of AMI, hypertension and diabetes), the biochemical lipid quantities, C-reactive protein and creatinine, all readily available in most hospitals. Information about post-infarction heart failure and post-infarction angina pectoris was missing. Therefore we added information about the current medication as proxy information instead. The lack of information about left ventricular ejection fraction may be partially or completely compensated since Solomon SD et al [Solomon SD, Claggett B, Desai AS, Packer M, Zile M, Swedberg K, et al. Influence of Ejection Fraction on outcomes and Efficacy of Sacubitril/Valsartan (LCZ696) in Heart Failure with Reduced Ejection fraction. Circ Heart Failure2016;9:e002744. doi: 10.1161] found that age, sex, hypertension, prior AMI, creatinine, diuretics, digoxin, and mineralocorticoid receptor antagonist were related to left ventricular ejection fraction, all quantities that we have included within the group referred to as “standard predictors”. Please, see page 16 section 2).

End of answer

3. This is a statistical analysis plan, and in this respect I would expect a more detailed account of some of the planned analyses.

Answer Thank you. We agree and above and below we explain how we have amended our manuscript giving more detailed account. End of answer

Comment # 4 a. The discussion mentions the rarity of missing values, but there is no information as to how many missing values there are or as to how this issue is addressed (complete cases, imputation, …). Also, it is stated that biochemical values may be missing because collection of
blood specimens was not carried out 'for one reason or another'. That is no convincing argument to state that missing values are most likely MCAR.

Answer There are almost no missing values in the data from the CLARICOR trial. (see [10] table 1. Similarly follow-up was complete apart from 0.5% emigrations and missing persons. However, the values of 24 biochemical quantities have been added to the CLARICOR data and these data have missing values. The set of placebo patients with one or more missing biochemical quantity values includes 210 patients out of 2199 placebo treated patients giving 210/2199 = 9.5% patients having one or more biochemical quantity values missing. The percentage for individual biochemical quantities ranges between 1.6 and 5.7%. We will use Little's test to decide whether a multiple imputation or a complete case analysis should be used (Please, see the first section on page 16 in the discussion section and the section on “missing values” on page 14) End of answer

Comment # 5 b. The issue of skewness remains too vague throughout the text. What is specifically meant with this, and how will this specifically be investigated?

Answer Please, see answer to reviewer 1’s comment # 7 End of answer

Comment # 6 c. How are the PH and linearity assumptions addressed?

Answer The ph assumption of all variables and the functional forms of continuous variables will be tested using procedures derived from cumulative sums of martingale-based residuals over follow-up time and/or covariate values. (see Lin DT, Wei LJ and Ying Z Biometrika (1993),80, 357-72). The test quantity will be a Kolmogorow-type supremum test P threshold = 0.05/number of covariates in the analysis (please, see page 12 the section on the proportional hazards assumption). End of answer

Comment # 7 d. Given the amount of standard predictors, are there enough events?

Answer Using the rule of thumb by Peduzzi et al [J Clin Epidemiol (1996);49:1373-9] the number of covariates must not exceed (number of patients with least common event)/10. CVD has the lowest number of rare events which gives 364/10 = 36 covariates (which should be compared to 25 covariates (24 ‘standard predictors’ + 1 advanced biochemical predictor). UAP has 397 events/10 = 39 covariates etc. We clearly owe the reviewers & Editor this calculation, but have decided not to include it in the manuscript End of answer
Comment # 8 e. P8: is it necessary to have a significant effect when used alone AND after correction for standard predictors? What if it is only significant after correction, or what if effects before and after correction are in the opposite direction?

Answer Very good point. Thank you. We now only let us be guided by the effect after correction; any marked discrepancy between the two analyses will prompt a closer look. (Please see page 12 first two sections in the flow of analyses section where the revised rule is described). End of answer

Comment # 9 f. What do you mean with lives saveable? How is this quantified?

Answer Please, see the second section of “illustration of predictive impact” on page 14. End of answer

Comment # 10 4. Is it necessary to have six outcomes? Would you expect strongly different predictors?

Answer Please, see answer to reviewer # 1’s comment # 6. End of answer

Comment # 11 5. C statistic does not represent a frequency of correct triage decisions.

Answer In fact it does – but in a setting that would require too many words to specify. We have added the adjective “two-patient” because the reviewer may have been thinking of “single-patient triage” in the emergency room (admit or let go). [Two patients arrive in a critical state at the same time, but there is only equipment for treating one of them (a triage scenario). Unbeknownst to the doctors, one of the two patients will die unless treated, the other will survive anyhow. Using a risk calculation, the doctors must treat the one with greater risk of being destined to die unless treated. Here, C = Prob{correct choice | situation described}. End of answer

Comment # 12 6. Second sentence in the methods section of the abstract is difficult, I suggest to reword.

Answer The sentence has been reworded as follows: “It included data from 4372 stable coronary artery disease patients who were alive by October 1999 and had been diagnosed with myocardial infarction or unstable angina pectoris during 1993 to 1999 in Copenhagen. They were
randomized during the period October 1999 to April 2000” (see page 2 second section). End of answer

Comment # 13 7. Write C in full in abstract.
Answer C has now been changed to Celsius. (See page 2 third section). End of answer

Comment # 14 8. What is meant with 'we plan to assess each of the advanced biochemical quantities in individual publications and then assess their combined effect in final publication'? The authors cannot really suggest to write a separate paper for every advanced biochemical predictor, and then a final paper? More than 10 advanced predictors are listed.
Answer The sentence has been removed. End of answer

Reviewer #3:
Comment # 1 This is an interesting paper which aims to investigate new predictors for several cardio-vascular events in a future publication. Paper writing is clear-cut and concise presenting a database yielded from a previous clinical trial. The authors exposed planned selection strategy to test new predictors' reliability and efficiency.
Answer Thank you for the very kind words. End of answer

Authors may consider following remarks regarding mainly methods issues:
Comment # 2- it would be important to report and clarify pathophysiological mechanisms and/or prior findings related to selected new predictors
Answer Please, see answer to reviewer 1 comment # 2 End of answer

Comment # 3- since more than 10 predictors will be tested in time-to-event analyses using 5 endpoints, p-values threshold should be corrected to the value of 0.001 = 0.05/ (12*5) due to multiple testing
Answer We have combined the 5 endpoints into one outcome. We will use the P values of the analyses as a data-reducing device where a threshold of P < 0.01 is used at each decision node on our way to identify a candidate biochemical predictor. We regard the set of selected candidate
biochemical predictors as the primary result of this essentially hypothesis-generating effort. In such contexts we think, formal use of Bonferroni corrections is wasteful of insight, (it disregards all but the lowest P-value, although other small P-values may carry more biological plausibility). But of course we will mention and discuss the multiplicity issue. Especially when facing biologically implausible results with low P values (please, see page 13 last section of “flow of analyses”). End of answer

Comment # 4- additional tools to C-index can be used to assess discrimination ability such as net reclassification improvement (NRI); Kerr KF, Wang Z, Janes H, McClelland RL, Psaty BM, Pepe MS. Net reclassification indices for evaluating risk prediction instruments: a critical review. Epidemiology. 2014 Jan;25(1):114-21. doi:0.1097/EDE.0000000000000018. Review. PubMed PMID: 24240655; PubMed Central PMCID: PMC3918180.

Answer Thank you very much. We have now revised the corresponding section in the paper (please, see page 13 section on “Illustration of predictive impact”) in accordance with the critical review to which you have referred us. End of answer

Comment # 5- Statistical analyses should be stratified using sex, as in main CVD risk equations, which may also be used in the present study depending on the outcome of interest (Wong ND. Epidemiological studies of CHD and the evolution of preventive cardiology. Nat Rev Cardiol. 2014 May;11(5):276-89. doi: 10.1038/nrcardio.2014.26. Review. PubMed PMID: 24663092.)

Answer Thank you very much. We will stratify by sex in the sense that if an effect of a predictive covariate is suspected to depend on the sex we will include an interaction between sex and the covariate in the statistical model (Please, see page 12 second section in “flow of analyses”) End of answer

Comment # 6- In sensitivity analyses, tough authors planned to use cox models given the data (time-to-event), competing risks cases in related to death events should be considered (depending on the outcome of interest).

Answer A competing risk approach (of whatever kind) would certainly be needed to produce a full picture of the web of imaginable courses of events, but our aim is to quantify the overall impact of predictors on the mortality/morbidity hazards which (in proportional hazards contexts) form the building blocks of competing-risk descriptions. End of answer