Author's response to reviews

Title: The Impact of Coronary Artery Disease Risk Loci on Ischemic Heart Failure Severity and Prognosis: Data from the COntrolled ROsuvastatin multiNAtional trial in heart failure (CORONA)

Authors:

Vincent G Haver (v.g.haver@umcg.nl)
Niek Verweij (n.verweij@umcg.nl)
John Kjekshus (john.kjekshus@medisin.uio.no)
Jayne C Fox (jayne.fox@astrazeneca.com)
Hans Wedel (hans.wedel@biostat.se)
John Wikstrand (john.wikstrand@wlab.gu.se)
Wiek H van Gilst (w.h.van.gilst@umcg.nl)
Rudolf A de Boer (r.a.de.boer@umcg.nl)
Dirk J van Veldhuisen (d.j.van.veldhuisen@umcg.nl)
Pim van der Harst (p.van.der.harst@umcg.nl)

Version: 2
Date: 6 October 2014

Author's response to reviews: see over
Dear editor,

On behalf of all co-authors, I herewith present you the revised manuscript entitled “The Impact of Coronary Artery Disease Risk Loci on Ischemic Heart Failure Severity and Prognosis: Data from the COntrolled ROsvastatin multiNAtional trial in heart failure (CORONA)” (MS: 4137533861302383).

Please find the responses to the raised comments by the referees below.

We are grateful to you for considering the revised manuscript for publication.

Yours sincerely,

Dr. Pim van der Harst
On behalf of the authors
Reviewer: Markus Scholz

Reviewer's report:

Major issues:

1. The CAD GWAS of Samani et al. (2007) used to select SNPs of interest in this study is outdated. Several new loci were found in recent meta-analyses comprising several more cases and controls (e.g. Schunkert, Nat. Gen. 2011). The authors should explain their choice, especially why they ignored more recent findings. On that note, GWAS references are outdated too.

The reviewer is correct that we tested the first reported variants only. The analyses are based on the SNPs from Samani et al. (2007, 2009)[1, 2] because they are the most robustly associated SNPs with stronger effect sizes compared to the newly identified SNPs. Budget and time-constraints unfortunately do not allow us to easily continue further genotyping in the CORONA cohort (DNA is due to informed consent requirements based at AstraZeneca and AstraZeneca has been re-organised and de-prioritised work on this cohort). Nevertheless, we feel that our analyses based on the GWAS of Samani et al. (2007, 2009)[1, 2] provide novel insights and are of great interest for the readership of BMC Medical Genetics.

We have added the Schunkert (Nat. Gen. 2011) reference to the revised document.

2. I would strongly recommend re-arranging the results section and tables in a more reader-friendly way: At first, it is not sufficiently clear which models were analysed, i.e. what does “Cov 4” really means for the specific endpoints. I suggest specifying the clinical models (i.e. without genetics) first and then presenting the results of adjusted and non-adjusted genetic analyses. I also wonder why not all SNPs are reported in the tables. What are the SNP selection criteria for different endpoints? I recommend reporting all SNP results which could be helpful for subsequent meta-analyses.

We thank the reviewer for his suggestions to further improve the readability of the manuscript. We now explain which models we use and were these models originate from. For the reviewers information, the models were a priori defined and are based on the previous report of Corona by Wedel et al.[3]. We considered the model named ‘cov 4’ which indeed is very unclear. Therefore we explained the model in the Statistical Analyses section (see below).

“Cov 4” has been renamed and called “adjusted analyses” in the whole manuscript.

All SNP association data with a P value < 0.05 are represented in the main manuscript. We present data on all variants in a table in the Supplemental Material.

From

“The baseline variables were analysed in linear model with only genotype as predictor and if significant, covariates were added in four subsequent steps; Step I: the covariates used in CORONA default tabulations (age, gender, diabetes mellitus, left ventricular ejection fraction, hypertension, myocardial infarction, NYHA, total cholesterol), Step II-IV: Covariates of steps 2-4 as listed in Additional Table 2 and explained in detail in Wedel et al.[3]”

To

“The baseline variables were analysed in linear model with only genotype as predictor and if significant, the following covariates were added to the adjusted model: age, sex, ejection fraction, NYHA class, systolic blood pressure, heart rate, body mass index, history of myocardial infarction, angina pectoris, diabetes mellitus, hypertension, stroke, intermittent claudication, aortic aneurysm, percutaneous coronary intervention, coronary artery bypass graft surgery, atrial fibrillation, implanted pacemaker, implanted cardiac defibrillator, smoking status, serum creatinine, alanine aminotransferase, creatine kinase, thyroid-stimulating hormone, triglycerides, hsCRP and NT-
proBNP, as explained in detail by Wedel et al. [3]. HF outcome determinants were analysed using Cox regression of outcome versus number of minor alleles. Analyses were conducted unadjusted and after adjusted for the above mentioned co-variates.”

3. Since this is a candidate analysis, more emphasize could be placed on analyzing different genetic models.

The reviewer is correct this is a candidate analysis. However, we focussed on the additive genetic model and have made it more clear in the results section:

“Genotypes were coded additively as 0, 1 or 2 in terms of the number of minor alleles.”

We agree that certain SNPs may be better modelled according to dominant or recessive model if this fits better. Samani et al. (2007) [2] used Cochran-Armitage test for trend (an additive model) for the discovery in a follow up study (Samani (2009) [1], there was no evidence of non-additivity for any of the loci assessed (i.e. better fit using a dominant or recessive model) therefore, we focussed on the additive model for this candidate analysis as well. This also reduces the multiple testing burden.

4. Since the majority of results are negative, I would suggest a power analysis to discuss the size of genetic effects detectable by this study.

Using the Genetic Power Calculator [4], we calculated the power and the number of cases needed for our analyses. We considered that with 381 cases we could detect a representative SNP with a risk allele frequency of 0.35 with an odds ratio of 1.3 at an α threshold of 0.05 with a power of 0.80. The power we have with 500 cases is 0.90. Therefore, our sample size is large enough to replicate CAD cases, although we tested other phenotypes as well.

Minor issues:

1. I think the applied Bonferroni correction does not make much sense. On one hand, much more than seven tests were calculated since multiple endpoints were analysed resulting in further accumulations of type I errors. On the other hand, I understand that this is a candidate study looking for alternative endpoints associated with CAD loci. Therefore, I would recommend dropping any Bonferroni corrections. Found associations should be discussed as suggestive requiring further replications.

We thank the reviewer for this comment and we agree. In the manuscript we have reported the raw (unadjusted) p-values to facilitate the interpretation by the readership. In the method section, we explain why we considered a P-value <0.0071 for the primary endpoint as statistically significant (Bonferroni adjustment of <0.05 for 7 independent loci) and as suggestive for all secondary endpoint analyses, since we tested novel associations not necessarily linked to CAD risk. In the discussion we included the statement as suggested by the reviewer on further replication; “The observed associations are suggestive and require further replication.”

2. Table 3: I have concerns regarding the Jonckheere-Terpstra trend test for the endpoint “number of hospitalizations”. Since the endpoint suggests count data type, Poisson regression appears to be more appropriate. If the endpoint was divided into categories, this should be explained and justified in more detail.

We thank the reviewer for bringing this to our attention and have reconsidered the methods used, also in consideration with our previous work published on this endpoint. We consider the Jonckheere-Terpstra test applicable for testing independent samples of ordered treatments in ordinal categories or continuous data. The Poisson regression test should be applied when testing nominal categories. Here, we tested the association of genotype (AA, AB, BB) with the number of hospitalisations. Since genotype is divided over 3 categories, we consider the Jonckheere-Terpstra test appropriate for our analyses.

3. Discussion: I think discussion of ANRIL was outdated too. There are more recent publications regarding this topic.

The reviewer is correct. We have updated our discussion on ANRIL.
“Furthermore, this locus is related to atherosclerotic disease burden in different vascular beds [11]. Despite these findings, the presence of this genetic variant, and presumably increased atherosclerotic burden, did not translate into increased HF severity or worse outcome in patients with ischemic HF. A possible explanation is that the effect of this locus acting through increased atherosclerotic disease burden is confounded by events defined by the severity of heart failure.”

“This locus is related to atherosclerotic disease burden in different vascular beds [10] and deletion of ANRIL in human aortic smooth muscle cells leads to a increase in proliferative capacity in culture [11]. Furthermore, the rate of proliferation of vascular smooth muscle cells is attenuated by the 9p21 genotype and the CAD risk allele (C allele) increases vascular smooth muscle cell proliferation, thereby likely playing an important role in the development of atherosclerosis [12]. Despite these findings, the presence of this genetic variant, and presumably increased atherosclerotic burden, did not translate into increased HF severity or worse outcome in patients with ischemic HF in the present study. A possible explanation is that the effect of this locus acting through increased atherosclerotic disease burden is confounded by events defined by the severity of heart failure.”
Reviewer: Inke König

Reviewer's report:

In their manuscript, the authors describe an association study on known loci associated with coronary artery disease and the outcome after heart failure. There is a clear rationale for the study, but some issues need to be addressed.

We appreciate the acknowledgement of our rationale, and we address the issues as follows:

Major compulsory revisions

1. Methods: It is unclear why the authors only consider the seven loci identified in one of the first GWA studies performed. Since 2007, many more loci have been reliably identified, and there is no a priori reason why these should not be part of this study.

   The reviewer is correct that we tested the first reported variants only. The analyses are based on the SNPs from Samani et al. (2007, 2009)[1, 2] because they are the most robustly associated SNPs with stronger effect sizes compared to the newly identified SNPs. Budget and time-constraints unfortunately do not allow us to easily continue further genotyping in the CORONA cohort (DNA is due to informed consent requirements based at AstraZeneca and AstraZeneca has been re-organised and de-prioritised work on this cohort). Nevertheless, we feel that our analyses based on the GWAS of Samani et al. (2007, 2009)[1, 2] provide novel insights and are of great interest for the readership of BMC Medical Genetics.

2. Methods: The description of when covariates were included needs clarification: Was the mentioned linear model used for the primary outcome? If this is time to event, this is not adequate. Were covariates only considered in the model if they were significant themselves, or if the genotype was significant on the outcome? Why were these covariates chosen in the first place? For a prognostic model, it would be important to include all prognostic variables that have reliably been reported previously.

   The primary endpoint was tested using Cox regression analyses (outcome versus number of minor alleles).

   We tested the association of genotype with outcomes using two models (unadjusted and adjusted for the following confounding factors: age, sex, ejection fraction, NYHA class, systolic blood pressure, heart rate, body mass index, history of myocardial infarction, angina pectoris, diabetes mellitus, hypertension, stroke, intermittent claudication, aortic aneurysm, percutaneous coronary intervention, coronary artery bypass graft surgery, atrial fibrillation, implanted pacemaker, implanted cardiac defibrillator, smoking status, serum creatinine, alanine aminotransferase, creatine kinase, thyroid-stimulating hormone, triglycerides, hsCRP and NT-proBNP. These co-variates have been distilled out of the publication by Wedel et al.[3], in which predictors of fatal and non-fatal outcomes have been reported.

3. Methods and Results (also in Abstract): The authors need to be more precise on their use of the term significance. For instance, “suggestive evidence for association” should be avoided, since the result from a statistical test can only be significant or not. Similarly, the authors use the term “convincingly associated”, for which there is no definition. Also, “P<0.05” is given in results, which is not sufficient.

   We agree with the reviewer that there are no clear definitions for the terms of suggestive or convincingly, even if it not directly relates to a P-value threshold. We have reviewed our use of the term “significance” and adjusted the use when appropriate.

Abstract

From

“However, the 1p13.3 locus (rs599839) showed suggestive evidence for association with all-cause mortality...”

to
“However, the 1p13.3 locus (rs599839) showed evidence for association with all-cause mortality…”

From
“Genetic variants strongly associated with CAD risk are not convincingly associated with the severity and outcome of ischemic heart failure.”

to
“Genetic variants strongly associated with CAD risk are not associated with the severity and outcome of ischemic heart failure in our cohort.”

Methods
From
“We studied 7 loci (1p13.3, 1q41, 2q36.3, 6q25.1, 9p21.3, 10q11.21, 15q22.33) which have been convincingly linked to CAD risk by previous GWAS”

to
“We studied 7 loci (1p13.3, 1q41, 2q36.3, 6q25.1, 9p21.3, 10q11.21, 15q22.33) which have been linked to CAD risk by previous GWAS”

Results
From
“Although some of the unadjusted association P-values were smaller than 0.05 (Table 2), we considered none of these loci convincingly associated with LVEF or NT-proBNP considering the multiple testing burden of these secondary endpoints.”

to
“Although some of the unadjusted association P-values were smaller than 0.05 (Table 2), we considered none of these loci associated with LVEF or NT-proBNP considering the multiple testing burden of these secondary endpoints.”

Conclusions
From
“Genetic variants associated with CAD and atherosclerotic disease burden are not convincingly associated with the severity and prognosis of patients with ischemic HF in the CORONA trail.”

to
“Genetic variants associated with CAD and atherosclerotic disease burden are not associated with the severity and prognosis of patients with ischemic HF in the CORONA trail.”

4. Results: In describing the associations, it should be stated whether the effects were in the direction expected from the literature.

We have added the directions of the main outcomes in Table 3 and added an asterisk if directions were concordant with previous observations (Samani et al. 2007, Samani et al. 2009) [1,2].

<table>
<thead>
<tr>
<th>Locus</th>
<th>SNP</th>
<th>Model</th>
<th>n (total)</th>
<th>n (events)</th>
<th>Hazard ratio</th>
<th>95% CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All-cause Mortality</td>
<td>1p13.3</td>
<td>rs599839</td>
<td>Unadj.</td>
<td>3,300</td>
<td>527</td>
<td>0.86*</td>
<td>0.74-1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Adj.²</td>
<td>2,218</td>
<td>341</td>
<td>0.74*</td>
<td>0.61-0.90</td>
</tr>
<tr>
<td>Mortality or WHF hospitalization</td>
<td>10q11.21</td>
<td>rs501120</td>
<td>Unadj.</td>
<td>3,300</td>
<td>1046</td>
<td>0.85*</td>
<td>0.75-0.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Adj.²</td>
<td>2,216</td>
<td>670</td>
<td>0.82*</td>
<td>0.70-0.96</td>
</tr>
</tbody>
</table>

5. Discussion: The authors state that their sample is exceptionally large. Still, the power of their study to detect meaningful effects should be estimated. Alternatively, there should be a thorough discussion based on the confidence intervals to describe which effects can safely be excluded.
Using the Genetic Power Calculator [4], we calculated the power and the number of cases needed for our analyses. We considered that with 381 cases we could detect a representative SNP with a risk allele frequency of 0.35 with an odds ratio of 1.3 at an $\alpha$ threshold of 0.05 with a power of 0.80. The power we have with 500 cases is 0.90. Therefore, our sample size is large enough to replicate CAD cases, although we tested other phenotypes as well.

6. Tables: In addition to giving the effect estimates, please show effects for all the genotype groups, for instance, in terms of mean values with variation measures.

We agree with the reviewer that we could provide more data on the reported associations with the variants studied. We now have included Supplementary Tables 3-6 and provide data on all variants.

Minor essential revisions:

1. Abstract – Methods: Please state the outcome variables here.

We have stated the outcome variables in the methods.

We’ve added

“The primary outcome was the composite of time to first event of cardiovascular death, non-fatal myocardial infarction and non-fatal stroke, secondary outcomes included mortality and hospitalisation due to worsening heart failure.”

to Abstract – Methods

2. Background and Abstract: It is not correct to state that the genetic loci affect the risk of CAD or cause CAD, since merely associations have been described.

We agree and we changed the text to reflect this.

Abstract

From

“Recent genome-wide association studies have identified multiple loci that affect the risk of developing coronary artery disease (CAD).”

to

“Recent genome-wide association studies have identified multiple loci that are associated with an increased risk of developing coronary artery disease (CAD).”

Introduction

From

“Whether these variants, with strong prior evidence to cause CAD, are also relevant for ischemic HF progression as reflected by HF severity and prognosis remains to be determined.”

To

“Whether these variants, with strong prior evidence to be associated with increased CAD risk, are also relevant for ischemic HF progression as reflected by HF severity and prognosis remains to be determined.”

3. Methods: A distinction is made between normally and non-normally distributed variables, but it is not stated where this information comes from.

Normality of the data was determined by visual inspection. HsCRP and NT-proBNP identified as being distributed skewed and therefore log-transformed.

We’ve added the following sentence in the Statistical Analysis section:

“Normality of the data was determined by visual inspection. HsCRP and NT-proBNP were non-normally distributed and therefore log-transformed.”

4. Methods: The genotypes were coded for an additive model. Is this the model that seemed to be most plausible for all SNPs in the original publication?
See answer to question 3 of reviewer #1.
Since the original publication was tested using the additive model, we have chosen this model for our analyses as well. We did not test other models in order to reduce the multiple testing burden.

5. Results: Please state whether the MAFs were similar as in the original publication.
MAFs were similar in the original publications:

<table>
<thead>
<tr>
<th>SNP</th>
<th>Minor allele</th>
<th>Major allele</th>
<th>MAF (current analysis)</th>
<th>MAF (Samani 2007 / 2009)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs599839</td>
<td>G</td>
<td>A</td>
<td>0.23</td>
<td>0.23 / 0.22</td>
</tr>
<tr>
<td>rs17465637</td>
<td>A</td>
<td>C</td>
<td>0.26</td>
<td>0.29 / 0.26</td>
</tr>
<tr>
<td>rs2972147</td>
<td>T</td>
<td>C</td>
<td>0.37</td>
<td>0.36 / 0.35</td>
</tr>
<tr>
<td>rs6922269</td>
<td>A</td>
<td>G</td>
<td>0.26</td>
<td>0.29 / 0.30</td>
</tr>
<tr>
<td>rs1333049</td>
<td>C</td>
<td>G</td>
<td>0.5</td>
<td>0.55 / 0.54</td>
</tr>
<tr>
<td>rs501120</td>
<td>C</td>
<td>T</td>
<td>0.15</td>
<td>0.13 / 0.16</td>
</tr>
<tr>
<td>rs17228212</td>
<td>C</td>
<td>T</td>
<td>0.27</td>
<td>0.30 / 0.27</td>
</tr>
</tbody>
</table>

6. Results: Please state how conformation to Hardy-Weinberg equilibrium was established.
We have calculated $P$-HWE according to the chi-square test for deviation.

7. Results: In table 3, please highlight the primary outcome.
Hazard ratio has been highlighted in Table 3.
EDITOR'S COMMENTS:

“Based on the reviewers comments, there are some severe issues pertaining to the selection of variants and statistics used. These need to be fixed before publication can be considered.”

We hope we have addressed these questions appropriately in the revised document.

EDITORIAL REQUIREMENTS:

Please update your ethics statement to include the name of the ethics committees that approved your study.

Since the CORONA study was conducted as a multi-center trial in with more than 20 affiliations, the list of ethical committees approving this study is very large. We feel that this list is too long to represent in the Methods section. Please let us know if this is obligatory, because we will add it if pertinent.
References


