Author's response to reviews

**Title:** Investigating the Complex Genetic Architecture of Ankle-Brachial Index, a Measure of Peripheral Arterial Disease, in non-Hispanic Whites

**Authors:**

Sharon LR Kardia (skardia@umich.edu)
M. Todd Greene (mtgreene@umich.edu)
Eric Boerwinkle (Eric.Boerwinkle@uth.tmc.edu)
Stephen T Turner (sturner@mayo.edu)
Iftikhar J Kullo (kullo.iftikhar@mayo.edu)

**Version:** 2  **Date:** 19 February 2008

**Author's response to reviews:** see over
February 19, 2008

Dear Editor-in-Chief,

Please find enclosed our revised manuscript that we have re-titled ‘Investigating the Complex Genetic Architecture of Ankle-Brachial Index, a Measure of Peripheral Arterial Disease, in non-Hispanic Whites’ which we would like to re-submit for publication in *BMC Medical Genomics* as an original research article. As part of our re-submission process, we have gone through the reviewer comments and provide our point-by-point responses below.

All authors have read and approved re-submission of the manuscript and the manuscript has not been published nor is not being considered for publication elsewhere.

None of the authors have any conflict of interest to declare.

Thank you very much for your consideration of the manuscript.

Sincerely,

Sharon L. Kardia, PhD

**Point-by-point response to the comments provided by Reviewer 1 (Marylyn Ritchie).**

**Major Compulsory Revisions**

How many patients in the study had evidence of PAD? In the section describing ankle-brachial index, the definition is provided; however, there is never an indication of how many patients in this study would be defined as PAD. In the discussion it says than abnormal ABI was low in this study. Again, how low?

*In the Discussion section, we have revised the section that describes PAD in our cohort to include a prevalence estimate and additional details that provide further justification for investigating ABI variation in our cohort despite the low prevalence estimate. As indicated in our revisions, the prevalence estimate of PAD in our cohort was 6.8% using a common clinical definition of an ABI value ≤ 0.90. This estimate increased to 10.2% when an alternate definition (also commonly used in the clinical setting) of an ABI value ≤ 0.95 was used.*

Is the manuscript in the proper BMC format? It seems that most BMC journals have the methods section last. Perhaps it is different for this one or it has changed.

*The manuscript has been checked to ensure that it complies with the journal’s formatting requirements. While many articles published in BMC journals list the Methods section last,*
Minor Essential Revisions

How does the internal replication differ from doing a 2-fold cross-validation? It would be nice to describe the benefit to doing the analysis this way. It would also be helpful to add more detail on the homogeneity of effects test.

As indicated in the paper, we used internal replication and 4-fold cross-validation (rather than 2-fold cross-validation) strategies as methods to reduce false positives. While both strategies were used to minimize type I error, internal replication is rooted in hypothesis testing, whereas cross-validation is more focused on predictive ability. Internal replication divides the sample into two distinct unrelated groups and regression estimates are derived for each of the groups separately. Results for each group are then compared to identify significant, replicable effects. Our internal replication strategy also permits the investigation of homogeneity of effects. Since there has been a huge emphasis on replication in genetic studies, we felt that the potential increase in type II error (as a result of small sample sizes in our internal replication groups) was worth it to reduce type I errors. With 4-fold cross-validation, the sample is divided into four equal parts. Unlike the holdout method (or 2-fold cross validation) where evaluation will depend largely on which data points end up in the training set versus the testing set, the evaluation with 4-fold cross validation depends less on how the sample is divided, as every data point will end up in the test set once and in a training set three times. Model predictive ability is then assessed by applying the model estimated using the training sets to the test set. In other words, cross-validation assesses how well a model generalizes to new data. We have revised the Statistical Analysis sub-section of the Methods section to include additional details that clarify how the homogeneity of effects was assessed.

When the cross-validation subsets are created, is familial relationship considered? In other words, are all members of a family kept in the same subset?

While we acknowledge the potential impact of familial correlation, we did not consider familial relationships when creating our cross-validation subsets for a couple of reasons. First, the cross-validation statistical analysis does not rely on family structure as other types of analyses (eg. FBAT or linkage studies) do. Secondly, generalized estimating equations (GEE) results, which account for error introduced by correlations in the sample, are unstable for cross-validation purposes. We are currently trying to augment our association methods to account for correlations when using cross-validation techniques.

How many statistically significant results would you expect by chance in Table 3? Are you seeing more than expectation?

Since Table 3 summarizes our results after we have applied all 3 methods to reduce type I errors, we wonder if the reviewer is really referring to Table 2 where the raw results are presented in terms of the number of tests and the number of significant associations based on various criteria. Prior to applying any of our methods to reduce type I error, we exceeded expectations (using a p-value cutpoint of p<0.10) when testing for each type of effect (i.e.- SNP
main effects, SNP-Covariate interactions and SNP-SNP interactions). For example, as highlighted in Table 2, 77 SNP main effects were identified, where 43 were expected. While we are seeing more than expectation in this sense, many of the tests of association are correlated. As such, exceeding expectation here is not a great indicator of success.

**Point-by-point response to the comments provided by Reviewer 2 (Andreas Ritsch).**

**Major**

1) Selection of SNPs: The title of this article suggests a comprehensive analysis of the genetic background of the clinical parameter ABI (The Genetic Architecture of ABI ...). For this ambitious claim one might expect the inclusion of all genetic variants known to be associated with this parameter. However, this is clearly not the case and it is not clear how the SNPs were selected for this study.

   We agree that the nature of the title is relatively grandiose and have tried to modify it slightly to be more representative of our multigenic approach and less ostentatious. We have revised the SNP Selection sub-section of the Methods section to include additional details on how SNPs were selected for this particular study.

2) Selection of patients: there is an obvious bias in this study, as patients were recruited from a hypertensive study population (GENOA). Data are at least non-representative for the total population.

   We agree, and point out in the Discussion section, that our findings are not generalizable to individuals that are younger, normotensive, or of other ethnicities. Although not representative of the total population, our study cohort was selected because it represents a high-risk population for diseases that have a large burden on public health. According to recent data from the National Health and Nutrition Examination Survey (NHANES), approximately 65 million adults have hypertension in the US (Fields et al., The burden of adult hypertension in the United States 1999 to 2000: a rising tide. Hypertension 2004; 44(4):389.). This is significant, as hypertensives are a high-risk sub-group for PAD. Furthermore, studies suggest that both the prevalence and incidence of PAD is higher in hypertensive subjects compared to non-hypertensive counterparts (DeBuyzere ML and Clement DL. Management of Hypertension in Peripheral Arterial Disease. Progress in Cardiovascular Diseases 2008; 50(4):238-263.). As such, we feel that investigating ABI variation among a high-risk, hypertensive cohort is both important and warranted, despite generalizability limitations.

3) The most important results of this paper have been presented in an earlier publication (Kullo et al., Association of polymorphisms in NOS3 with the ankle-brachial index in hypertensive adults, Atherosclerosis 2007)

   We do acknowledge, and report in our Discussion section, that the main effect associations reported here between the NOS3 SNPs and ABI were reported in the above referenced
publication. While we did not find main effects in polymorphisms from other genes, the current study investigated potential main effects from 111 other candidate genes. Furthermore, we report on a number of SNP-Covariate and SNP-SNP interactions in the present study, which we feel highlight the multifactorial and polygenic nature of complex, common diseases such as PAD. As reported in the Results section, the two NOS3 polymorphic main effects only explained 0.65% of the inter-individual variation in ABI, while the top SNP-Covariate and top SNP-SNP interactions additionally explained 2.25% and 4.5% of the ABI variation, respectively. While our results may be preliminary, we believe they extend beyond the previously reported main effect understanding of this complex phenotype.

**Minor**

1) Basic clinical parameters:
- Waste to hip ratio should be presented in addition to BMI.
- Hba1c values should be presented
- TG and CRP values should not be presented in log values used for statistic analysis.

We have revised Table 1 to include waist to hip ratio values in addition to BMI, and present the triglyceride and C-reactive protein values that have not been log transformed. While we acknowledge the usefulness of hemoglobin A1C values, these measurements were not taken as a part of this study and cannot be presented here.