Author’s response to reviews

Title: Development of a Blood-based Gene Expression Algorithm for Assessment of Obstructive Coronary Artery Disease in Non-Diabetic Patients

Authors:

Michael R Elashoff (melashoff@cardiodx.com)
James A Wingrove (jwingrove@cardiodx.com)
Philip Beineke (pbeineke@cardiodx.com)
Susan E Daniels (sdaniels@cardiodx.com)
Whittemore R Tingley (tingley.whittemore@gene.com)
Steven Rosenberg (srosenberg@cardiodx.com)
Szilard Voros (szilard.voros@piedmont.org)
William E Kraus (william.kraus@duke.edu)
Geoffrey Ginsburg (geoffrey.ginsburg@duke.edu)
Robert S Schwartz (robert.stockton.schwartz@gmail.com)
Stephen G Ellis (elliss@ccf.org)
Naheem Tahirkheli (ntahirkheli@ocaheart.com)
Ron Waksman (ron.waksman@medstar.net)
John McPherson (john.mcpherson@vanderbilt.edu)
Alexandra Lansky (alexandralansky@gmail.com)
Eric J Topol (etopol@scripps.edu)

Version: 2 Date: 31 January 2011

Author’s response to reviews: see over
To: BMC Medical Genomics Editorial Staff,

Re: Revised Manuscript 5663589274862453

Development of a blood-based gene expression algorithm for assessment of obstructive coronary artery disease in non-diabetic patients


Thank you for the reviewer and editorial comments on our manuscript. We have revised the manuscript as suggested and made sure that it conforms to the journal style and formatting, and uploaded a revised manuscript in “Track changes” mode as well as a “Clean” copy. This covers only substantive and not copy-editing or formatting changes.

Specific responses to the reviewers’ comments are detailed below:

Responses to Reviewer 1:

1. As presented in table 1, “Caucasian” is also a significant risk factor (with p-value = 0.023) for obstructive CAD. This was ignored by the authors. Also note that mean gene expression may be significantly different between European-derived and Asian-derived populations. Therefore, information of the ethnic makeup for the CATHGEN registry and PREDICT study patient cohorts would be very helpful.

1. While it is correct that in univariate analyses, caucasian race was a risk factor for CAD, after adjustment for sex and age, race was no longer significantly associated with CAD in any of the three cohorts. We therefore omitted discussion of race. The reviewer makes a good point regarding differences between European and Asian populations. However, the number of Asians was very low in CATHGEN and PREDICT (approx 1% of the populations) and thus we were not powered to assess these potential differences. We have added a sentence with respect to this point in the Limitations section.

2. To assess the performance of the classifier, the authors utilized only one measure: the cross-validated area under the curve (AUC) in the ROC analysis. Other more straightforward measures such as the estimated prediction accuracy, sensitivity, specificity, and especially the plot of ROC should also be presented.

2. We have added the ROC plot to the manuscript (Figure 7). We focused on AUC since the test has a continuous score outcome, as opposed to a binary result, where sensitivity/specificity might be more appropriate. The more clinically oriented validation study manuscript does contain these numbers in the context of clinical decision making (as opposed to statistical validation).
3. In the section “Algorithm Derivation and Performance,” the authors mention that they applied LASSO and Ridge regression to finalize the classifier. More technical details on how they implemented these two methods should be provided. More importantly, because LASSO does both parameter shrinkage and model selection, why did they apply Ridge regression, which does parameter shrinkage only, after they already performed LASSO?

3. We used LASSO for its variable selection properties, but felt based on comparative cross validation that it did not provide sufficient shrinkage. Hence, after LASSO variable section, ridge regression was used to estimate the final model weights. In retrospect, we could have used a combined L1/L2 shrinkage method such as elastic net to accomplish both steps at once (Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Statist Soc B. 2005;67:301–320. We have added some additional technical explanation of the mode fitting to the methods, and a reference to the above method in the Discussion.

4. Because diabetic patients were excluded from the study, the implications of the derived classifier may be limited.

4. We agree and the results from the final classifier are clearly limited to non-diabetic patients as noted in the limitations section.

5. A significant amount of content in Results and Discussion were highly repetitive.

5. We have endeavored to reduce the redundancy in these sections as reflected in the revised text.

Responses to Reviewer 2:

1. As pointed out by the authors, other studies have identified signatures associated with CAD (e.g. refs 12 and 13); how do these previous findings compare with the signatures associated with the signature described in this manuscript? It would seem logical to factor this information in the selection process.

1. It is often difficult to discern why different approaches yield different gene sets. As has been pointed out by others, alternate sets of genes can represent very similar information, due to the complex correlation structure of gene expression. Reference 13 is our previous work and the gene signature described therein is highly correlated with Term 2 in the present algorithm (especially in males); note that the most significant gene from Reference 13, S100A12, is part of this term; we have added a note to this effect in the Discussion.

With respect to Reference 12, these authors took quite a different analytical approach to ours in focusing on correlation rather than classification, on extreme controls with no angiographic disease, and also by looking at least in gene discovery at matched cohorts of female and male patients within a restricted age range. We also note that these authors used the manual extraction method for RNA isolation which we found to be quite
variable, as compared to the automated, bead-based method we developed and used in algorithm development. We have added a brief section on these results in the Discussion.

2. The approach employed here for the development of a diagnostic signature of CAD is principled but also “myopic”. Given the large number of samples available it would have been straightforward during the training/discovery phase to compare the performance of several competing prediction models. Is the algorithm presented here – that is rather convoluted – truly constitutes the best approach for classifying cases vs controls? It might turn out to be unnecessarily complicated or arcane. The authors need to benchmark other approaches and report their performance.

2. Several alternate prediction methods were evaluated early in the model building process: CART, unpenalized logistic regression, and partial least squares. However, these models were found to have inferior accuracy compared to the penalized methods (LASSO and ridge regression) we ultimately selected. However, we did not carry though these alternate methods though the complete biologically guided process of combining genes into terms, and thus benchmarking alternate statistical approaches is beyond the scope of the current investigation. We would also add that if a single method outperforms others with a given data set, it is often the case that it is over-fitting and will not generalize. For this reason we focused on the penalized models.

3. The manuscript lacks clarity on some key points. The authors should define clearly how this algorithm can be applied, in other words what it can do and cannot do. The language used throughout the manuscript is not always consistent and can easily lead to confusion. …Please state clearly what the intended use for this test is, what it can and cannot do. Basic details are also lacking in the result section, which requires going back and forth between this section and methods (e.g. what type of sample was used, whether multiple testing corrections were applied, the type of array).

3. We have further clarified the use of the test and have endeavored to make the language more consistent as suggested; specifically we have indicated that the use of the test is “to assess obstructive CAD likelihood” throughout. Although it is typical to separate rigorously Results and Methods, it may have detracted from the presentation. Thus, we have incorporated some of the reviewer’s suggestions into the Results section.

4. The choice of controls is also critical and the authors should provide more details on this point. Are all control subjects with complaints of chest pain? Does the choice of the control population limit the application of this test in any way? Would the test be able to distinguish between early stage and late stage disease?

4. For the PREDICT algorithm development cohort the detailed inclusion/exclusion criteria have been published (Rosenberg et al., 2010); all patients were clinically referred for angiography, as this is the population one would like to classify as to the presence of obstructive CAD. As shown in Table 1B, approximately 36% of patients were asymptomatic, including 35% of controls. Although we defined disease in a dichotomous manner at 50% stenosis by QCA for the algorithm derivation, we subsequently showed
that the algorithm score is highly correlated with maximum percent stenosis as well (Rosenberg et al., Figure 3).

We hope that the revisions to the manuscript and specific responses enumerated above, as well as the copy-editing, are sufficient for acceptance of the manuscript for publication in BMC Medical Genomics.

Sincerely,

Steven Rosenberg, PhD
Corresponding Author
Address: CardioDx, Inc., 2500 Faber Place, Palo Alto, CA 94303
Fax: 650 475 2714
Phone: 650 475 2714
E-mail: srosenberg@cardiodx.com