Editorial

Clinical Utility of Genetic Variants for Cardiovascular Risk Prediction
A Futile Exercise or Insufficient Data?

Emanuele Di Angelantonio, MD, MSc, PhD; Adam S. Butterworth, MSc, PhD

Estimation of an individual’s cardiovascular disease (CVD) risk usually involves measurement of risk factors correlated with risk of CVD to identify people who may especially benefit from preventive action, such as lifestyle advice or pharmacologic agents.1 Since the Framingham Risk Score was first developed, several other risk-prediction algorithms have been proposed, each involving a core set of the same established risk factors (ie, age, sex, smoking, blood pressure, and total cholesterol), but differing in their inclusion of various other characteristics (eg, ethnicity or presence of diabetes mellitus).2 The challenge in recent years has been to improve existing CVD risk-prediction models by including additional information to the traditional risk factors generally included in risk scores. Several additional soluble biochemical factors have been advocated for inclusion, but contradictory evidence has been reported on the incremental predictive gain afforded by these markers, and there is divergence of expert opinion about their clinical usefulness.3,4 There is also debate about the value of supplementing conventional risk factors with information on family history of disease, which may reflect shared genetic or environmental effects (or even epigenetic effects) within families.3,5 The challenge in recent years has been to improve existing CVD risk-prediction models by including additional information to the traditional risk factors generally included in risk scores. Several additional soluble biochemical factors have been advocated for inclusion, but contradictory evidence has been reported on the incremental predictive gain afforded by these markers, and there is divergence of expert opinion about their clinical usefulness.3,4 There is also debate about the value of supplementing conventional risk factors with information on family history of disease, which may reflect shared genetic or environmental effects (or even epigenetic effects) within families. Although individuals with a close relative who has CVD are at approximately 2-fold risk of having CVD themselves,4 there is conflicting evidence on the benefit of including family history in risk-prediction algorithms,3,6 and many commonly used risk scores do not therefore include family history.2

More recently, however, there has been increasing interest in adding genetic variants associated with CVD rather than adding family history to risk-prediction models. This is a potentially attractive concept because measurement of genetic variants is now accurate and cheap, plus variants would only need to be measured once as they are fixed at conception and therefore may be relevant for younger individuals before the conventional risk factors, such as hyperlipidemia, hypertension, and diabetes mellitus, become apparent. Until a few years ago, genetic epidemiologic studies of CVD were predominantly candidate gene studies involving focused investigation of relatively few genetic variants based on plausible biological hypotheses. Many of these studies had anticipated identification of variants that are common in populations with moderate-to-large effects on disease risk. However, the combination of the low prior odds of the variants selected for study, inadequate power (ie, small sample size), and overliberal declarations of significance, resulted in the reporting of many seemingly positive findings that remain unreplicated or directly refuted.5 In recent years, genome-wide association studies (GWAS) have demonstrated that so-called hypothesis-free global-testing methods can advance discovery and understanding of genetic variants in relation to chronic complex disease, given appropriate study design and sample size. In particular, the advent of GWAS has greatly enhanced our knowledge of the genetic architecture of vascular disease, yielding >30 variants confirmed to be associated with CVD to date, as well as >200 associated with traditional vascular risk factors (including lipids, blood pressure, body mass index, and type 2 diabetes mellitus).8–11 This recent and continuing success in identifying increasing numbers of robustly associated genetic variants has led to reassessment of whether genetic data can provide clinically useful information by improving risk prediction and reducing disease risk through a more efficient application of primary and secondary prevention strategies.

To date, however, it remains unclear whether genetic variants have the potential to deliver sufficiently accurate predictions to make genetic-targeted intervention a realistic possibility. Despite the weak individual effects of the common genetic variants so far identified (per allele odds ratios typically <1.2 for CVD risk),8 it has been hypothesized that such variants could significantly impact on disease development when their contributions are combined, as well as substantially improve CVD risk prediction. Indeed, several studies have explored whether genetic markers can improve risk prediction using genetic risk scores (GRS) constructed on the basis of the combination of risk alleles inherited.12–20 These studies, however, have reported conflicting or generally modest results (Table): although some have shown that genetic variants are related with risk of CVD independent of conventional risk factors using measures of association (such as relative risks), none were able to demonstrate a clinically significant improvement in predictive ability using both measures of risk reclassification (eg, net reclassification index) and discrimination (eg, C-index), which are more informative for the purpose of risk prediction.3

In this issue of Circulation: Cardiovascular Genetics, Patel et al examine the association of 11 previously identified...
genetic variants with both prevalent and incident myocardial infarction (MI) in 2 cohorts of subjects with coronary artery disease undergoing coronary angiography for diagnosis or treatment purposes. This study extends the range of previous genetic studies by assessing a GRS (based on variants predominantly identified through previous GWAS) in the context of risk prediction for secondary prevention. Patel et al compared MI cases with controls aged ≥70 years but without evidence of MI. The investigators found that these variants were significantly associated with prevalent MI and that the area under the receiver operator characteristic curve for predictive models of prevalent MI increased modestly after addition of the GRS to traditional risk factors. Furthermore, patients were followed up prospectively for incident MI during an average follow-up of ≈3 years. In contrast with the association reported for prevalent MI, the investigators reported that these variants were not associated with short-term incident events and therefore were of limited clinical value for risk prediction in the context of secondary prevention.

Although these findings further reduce the likelihood that currently available genetic information can substantially improve CVD risk prediction, especially in the context of secondary prevention, there are limitations to the evidence so far. First, as acknowledged by the investigators, this study, as with most previous studies, was underpowered to detect significant improvement in risk prediction of incident MI. Analyses of many thousands of incident MI cases will be required to provide enough power to evaluate reliably any association of genetic variants and to assess whether new risk scores that incorporate novel genetic loci could improve risk prediction. Second, only a few (typically around a dozen) selected genetic variants, discovered through GWAS, have been included in the GRS used, limiting the power of the GRS to improve risk prediction. Alternative approaches to increasing power, such as using all variants nominally associated with CVD risk in one particular study are prone to bias and nontransferability. Furthermore, methodologic considerations need to be addressed before inclusion of genetic variants in risk scores. For example, some of the risk alleles commonly incorporated in GRS are known to act via pathways that are already captured by the risk factors included in standard prognostic models (eg, lipids), and are therefore unlikely to improve prediction, unless genetic markers of life-long levels of risk factors are found to be stronger predictors than measurements at a single time point. (Addition of a GRS to a standard prognostic model containing family history may suffer from the same limitation.) Third, genetic effect estimates from GWAS are likely to be inflated owing to the winner’s curse, and several studies have used extreme subjects to identify genetic associations by sampling high genetic-risk cases and low genetic-risk control subjects, which may further inflate effect estimates. Finally, as noted by the investigators of the current study, risk prediction is particularly challenging in the context of individuals with established coronary artery disease, given that different pathophysiologic mechanisms could be involved in the development and progression of atherosclerosis versus complications of established disease such as MI, which are often caused by rupture of plaques and thrombosis.

Several issues should be considered when evaluating the apparent failure of currently available genetic variants to demonstrate clinically meaningful improvements in CVD risk prediction. First, a potentially crucial factor is the small proportion of the total genetic variation in CVD explained by common genetic variants currently identified, as suggested by a recent GWAS to be just 10%. Modeling exercises have estimated that complex traits are likely to be underpinned by very large

Table. Prospective Studies Assessing CVD Risk Prediction Using Multiple Genetic Markers and Risk-Prediction Metrics

<table>
<thead>
<tr>
<th>Authors, Year</th>
<th>Population Source</th>
<th>No.</th>
<th>Type</th>
<th>Outcome Assessed</th>
<th>Selection of SNPs</th>
<th>Association With: CVD/CHD</th>
<th>Risk Factors</th>
<th>Established Risk Factors</th>
<th>Family History</th>
<th>Risk Discrimination</th>
<th>Risk Reclassification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drenos et al, 2007</td>
<td>General population</td>
<td>183</td>
<td>CHD</td>
<td>12</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>NA</td>
</tr>
<tr>
<td>Morrison et al, 2007</td>
<td>General population</td>
<td>1452</td>
<td>CHD</td>
<td>11</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Kathiresan et al, 2008</td>
<td>General population</td>
<td>238</td>
<td>CVD</td>
<td>11</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>NA</td>
</tr>
<tr>
<td>Paynter et al, 2010</td>
<td>General population</td>
<td>777</td>
<td>CVD</td>
<td>101</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Ripatti et al, 2010</td>
<td>General population</td>
<td>1015</td>
<td>CVD</td>
<td>13</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Thanassoulis et al, 2012</td>
<td>General population</td>
<td>539</td>
<td>CVD</td>
<td>102</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Vaarhorst et al, 2012</td>
<td>General population</td>
<td>642</td>
<td>CHD</td>
<td>179</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Lluis-Ganella et al, 2012</td>
<td>General population</td>
<td>536</td>
<td>CHD</td>
<td>153</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Andreassi et al, 2012</td>
<td>Hospital (preexisting CHD)</td>
<td>119</td>
<td>CVD</td>
<td>48</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>NA</td>
</tr>
</tbody>
</table>

NA indicates not available; CHD, coronary heart disease; and CVD, cardiovascular disease.

Established risk factors include age, sex, smoking status, history of diabetes mellitus, plasma lipids, blood pressure, and body mass index; risk discrimination assessed using either area under the ROC curve or C-index; risk reclassification assessed using net reclassification index.
numbers of common variants with even weaker effects,26,27 and simply expanding the number of such variants in panels as they are discovered is unlikely by itself to address this issue. However, with larger-scale studies and new genomic-sequencing technologies, there is also the potential to identify low-frequency variants that may have larger effect sizes. Inclusion of such variants in GRS could potentially provide some improvement to their ability to predict risk, although these would need to be sufficiently common in the population and have sufficiently strong effects to provide any real benefit to risk-prediction models. Second, the selection and combination of variants in GRS is essential to maximize the potential improvement in risk prediction over and above risk factors currently used in risk prediction. Simulation studies have shown that the addition of markers moderately correlated with factors already included in prediction models is likely to provide only minimal improvement.28 Furthermore, although identification of gene–gene or gene–environment interactions may provide important insights regarding the biological mechanism of many common complex diseases, such interactions are likely to only modestly impact on risk-prediction models designed for use in the general population.29 Third, it is not necessarily clear how the inclusion of genetic information would lead to changes in patient management, improvements in risk factor profiles, or reductions in disease outcomes. The gold standard for establishing the effectiveness of such strategies is randomized trials. However, there are few examples of trials including genetic information in cardiology, either in the primary or secondary prevention setting. Recently, a randomized trial has been launched to investigate whether CVD risk factor profiles can be improved by providing participants with knowledge related to their inherited risk of CVD in addition to information on their risk as estimated by conventional risk factors. Genetic risk information will be provided using a GRS incorporating information from 19 loci that have been securely associated with CVD but do not appear to be associated with conventional risk factors, with risk factor profile improvements to be assessed 3 and 6 months later. In aggregate, therefore, the available evidence does not currently support the clinical usefulness of GRS for secondary or primary prevention of CVD, but improvements in power and design of studies combined with discovery of additional robustly associated genetic variants will provide more conclusive insight. Further research is also required to address other clinically relevant questions concerning risk prediction before genetic variants could be incorporated into routine practice, such as cost-benefit analysis, optimal clinical laboratory testing procedures, and ethical issues.

Disclosures

None

References


7. Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Repli-


24. Libby P, Theroux P. Pathophysiology of coronary artery disease. Circula-

analysis identifies 13 new susceptibility loci for coronary artery disease. 


**Key Words:** Editorials ◼ cardiovascular diseases ◼ genetic polymorphisms ◼ prediction
Clinical Utility of Genetic Variants for Cardiovascular Risk Prediction: A Futile Exercise or Insufficient Data?
Emanuele Di Angelantonio and Adam S. Butterworth

_Circ Cardiovasc Genet_. 2012;5:387-390
doi: 10.1161/CIRCGENETICS.112.964148
_Circulation: Cardiovascular Genetics_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2012 American Heart Association, Inc. All rights reserved.
Print ISSN: 1942-325X. Online ISSN: 1942-3268

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circgenetics.ahajournals.org/content/5/4/387

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in _Circulation: Cardiovascular Genetics_ can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to _Circulation: Cardiovascular Genetics_ is online at:
http://circgenetics.ahajournals.org/subscriptions/