Prediction of Cardiovascular Disease Outcomes and Established Cardiovascular Risk Factors by Genome-Wide Association Markers

John P.A. Ioannidis, MD

Background—Genome-wide association (GWA) platforms have yielded a rapidly increasing number of new genetic markers. The ability of these markers to improve prediction of clinically important outcomes is debated.

Methods and Results—A systematic review was performed of GWA-derived markers associated with cardiovascular outcomes or other phenotypes that represent common established risk factors for cardiovascular outcomes. Sources of information included the National Human Genome Research Institute catalog of published GWA studies, and perusal of the eligible GWA articles, meta-analyses on the respective associations, and articles on the incremental predictive performance of common variants in the GWA era. A total of 95 eligible associations were retrieved from the National Human Genome Research Institute catalogue of published GWA studies as of September 2008. Of those 36 have statistical support of $P < 10^{-7}$. In depth evaluation of the respective articles shows 28 independent associations with such statistical support, pertaining to coronary artery disease, myocardial infarction, atrial fibrillation/flutter, prolongation of QT interval, as well as type 2 diabetes, body mass index, high-density lipoprotein levels, low-density lipoprotein levels, and nicotine dependence. Between-study heterogeneity is not taken into account usually, but it seems common and it would pose a challenge to generalizability across different populations for these markers. Still limited data are available in non-white populations. Effect sizes are small and may be even smaller in subsequent replications and meta-analysis. Population attributable fractions are substantial, given the large frequency of the risk alleles. However, individualized risk measures are typically very small (proportion of variance explained <1% per marker). When used in conjunction with traditional predictors, improvement in overall prediction (eg, area under the curve) or risk reclassification is limited, and subject to methodological caveats.

Conclusions—Despite very promising signals in terms of statistical significance, evidence for improvement in cardiovascular prediction by currently available markers derived from GWA studies is sparse. Clinical use of such markers currently would be premature. (Circ Cardiovasc Genet. 2009;2:7-15.)

Key Words: genome-wide association studies ■ prediction ■ cardiovascular ■ risk factors

In the past 2 years, the discovery of genetic associations for common diseases and complex traits has been accelerated by the advent of high-throughput genotyping platforms and by substantial improvements in the available sample size and quality control measures in genetic epidemiology studies.1 Currently, there is a rapidly increasing number of association signals that attain very strong statistical support and get replicated in additional datasets.

Editorial see p 1
Clinical Perspective see p 15

Clinical Perspective see p 15

A major challenge is whether these new genetic markers can improve outcome prediction. Many companies are already marketing genome testing for this purpose. However, despite the enthusiasm, the rationale and public health implications of such applications remain debatable. I examine the status of the literature on the predictive potential for common genetic variants that have emerged from genome-wide association (GWA) studies as of September 2008. The review focuses on genetic markers for cardiovascular outcomes and also on genetic markers for established risk factors of cardiovascular outcomes.

Review Methods
I consider here all genetic markers shown to be associated with any cardiovascular outcome, or with common established independent risk factors of cardiovascular disease, including type 2 diabetes mellitus, body mass index, smok-
ing, hypercholesterolemia, and hypertension. I excluded risk factors that are contentious on the magnitude and/or independence of their effects (eg, C-reactive protein, triglycerides, uric acid).

The search for studies and eligible markers was based on the online catalog of GWA studies hosted by the National Human Genome Research Institute (last searched September 20, 2008). Details on the National Human Genome Research Institute catalog appear elsewhere. The catalog is regularly updated to include all published studies that have performed genome-wide evaluations for human disease phenotypes and traits and it lists associations with \( P \leq 10^{-5} \). Information was used here on the study, chromosomal region, potentially implicated gene(s), single nucleotide polymorphism with the strongest statistical support, the frequency of the risk allele, and the respective effect size, and 95% confidence interval.

The current review focuses primarily on those associations that have reached a \( P < 10^{-7} \) as analyzed by the primary authors. Genome-wide significance depends also the studied populations and their LD structure as well as the available sample size, but the \( 10^{-7} \) threshold is used for convenience. It is a lenient threshold, if one accounts also for the multiplicity of analyses involving different phenotypes in such studies. Several associations with less strong support may also have been replicated in subsequent, additional GWA and other studies, but I wanted to focus on a sample of associations with strongest support upfront.

The articles of the associations reaching genome-wide significance were further examined to see whether any of associations in the National Human Genome Research Institute catalog were duplicate (pertaining to same locus and outcome) and whether additional eligible associations with \( P < 10^{-7} \) had not been catalogued. For each of the eventually eligible associations with genome-wide significance, I recorded information on the presence and extent of between-study heterogeneity in the genetic effects, whenever this had been assessed. I searched also for identifying meta-analyses and large-scale subsequent collaborative replication studies relevant to these associations. Finally, I searched for articles that have used common genetic variants, including markers derived from GWA platforms, to examine the improvement of predictive ability for the eligible cardiovascular outcomes and related established risk factors. The search was based on screening the citations of the original GWA investigations (Thomson Web of Science).

Statement of Responsibility

The author had full access to and takes full responsibility for the integrity of the data.

Eligible GWA Associations

Of 454 entries in the catalog, 95 pertained to outcomes and traits that would be eligible for this review. Of the 95, 36 (38%) pertained to associations with \( P < 10^{-7} \). Some of the 36 associations had duplicate entries because the same markers or markers with very high LD had been found in 2 or more investigations. Excluding 10 such duplicates, 26 independent associations with genome-wide significance remained and I identified 2 more with eligible genome-wide significance after perusing the respective full articles. The 28 independent associations are shown in Table 1. These associations pertained to cardiovascular outcomes including coronary artery disease or myocardial infarction (n = 4), QT interval duration (n = 1), and atrial fibrillation/flutter (n = 2); and established risk factors, including type 2 diabetes (n = 11), high-density lipoprotein (HDL) (n = 2), low-density lipoprotein (LDL) (n = 5), body mass index (n = 2), and nicotine dependence (n = 1).

The other 59 associations that had less strong statistical support included also additional phenotypes, such as blood pressure/hypertension, coronary spasm, electrocardiographic traits, heart rate variability, coronary artery calcification, subclinical atherosclerosis, echocardiographic traits, exercise treadmill testing, endothelial function tests, and heart failure.

In some cases, associations with the same or highly linked polymorphisms had been identified for different correlated phenotypes. For example, FTO was first found to be associated with type 2 diabetes, but the association was inconsistent, whereas the association with body mass index and obesity was consistent across diverse populations. Another polymorphism (rs599839) was identified to be associated with LDL levels and a subsequent analysis of 2 combined GWA studies documented an association with coronary artery disease, which may be in part or in whole a reflection of the LDL effect.

Between-Study Heterogeneity

An important feature for a marker or set of markers that are considered to be used for predictive purposes is to have consistent effects across diverse populations. If the genetic effects are not consistent, then the applicability of the markers may be limited to those populations where the genetic effect is clearly seen. Between-study heterogeneity can be estimated empirically from the available data when the association between a marker and a disease has been measured in several populations. A measure of heterogeneity of the study-specific effects is the \( I^2 \) metric (“inconsistency”); it takes values from 0% to 100% and it shows the extent of heterogeneity beyond chance. Traditional tests for assessing whether the null hypothesis of homogeneity can be rejected (Q or Breslow-Day tests) are usually grossly underpowered to detect heterogeneity with limited numbers of independent populations.

The main results for all the associations reported in Table 1 did not take the potential of heterogeneity into account, ie, when several populations were tested for replication, the results were combined with simple pooling, simple stratification, or an equivalent fixed effects model (eg, Mantel-Haenszel or inverse variance fixed effects). Estimates of \( I^2 \) were provided in only 8 of the 25 associations where several populations were assessed in the replication phase (Table 2). In 5 of those, the estimated \( I^2 \) was >0% and in 2 the heterogeneity even reached formal statistical significance. Of the 6 associations where the investigators had evaluated between-study heterogeneity and had examined also random effects models in secondary analyses, 4 of the 6 associations no longer had \( P < 10^{-7} \). These were the associations of...
Table 1. Eligible Associations Reaching “Genome-Wide Significance” Levels for Cardiovascular Disease and Related Established Risk Factors

<table>
<thead>
<tr>
<th>References</th>
<th>Disease/Trait</th>
<th>Region</th>
<th>Gene Neighborhood</th>
<th>SNP-Risk Allele</th>
<th>Frequency</th>
<th>Effect</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>CAD</td>
<td>6q25.1</td>
<td>MTHFD1L</td>
<td>rs6922269-A</td>
<td>0.25</td>
<td>3.00E-08</td>
<td>0.123</td>
</tr>
<tr>
<td>11</td>
<td>CAD</td>
<td>10</td>
<td></td>
<td>rs501120-T</td>
<td>0.85</td>
<td>9.00E-08</td>
<td>0.133</td>
</tr>
<tr>
<td>13</td>
<td>CAD</td>
<td>9p21.3</td>
<td>CDKN2A, CDKN2B</td>
<td>rs1333049-C</td>
<td>0.47</td>
<td>1.00E-13</td>
<td>0.147</td>
</tr>
<tr>
<td>14</td>
<td>Myocardial infarction</td>
<td>9p21.3</td>
<td>CDKN2A, CDKN2B</td>
<td>rs107577278-G</td>
<td>0.45</td>
<td>1.00E-20</td>
<td>0.128</td>
</tr>
<tr>
<td>12</td>
<td>Atrial fibrillation/flutter</td>
<td>4q25</td>
<td>PITX2, ENPEP</td>
<td>rs2200733-T</td>
<td>0.11</td>
<td>3.00E-41</td>
<td>0.172</td>
</tr>
<tr>
<td>12</td>
<td>Atrial fibrillation/flutter</td>
<td>4q25</td>
<td>PITX2, ENPEP</td>
<td>rs10033574-G</td>
<td>0.08</td>
<td>7.00E-11</td>
<td>0.139</td>
</tr>
<tr>
<td>18</td>
<td>QT prolongation</td>
<td>1q23.3</td>
<td>NOS1AP</td>
<td>rs10494366-G</td>
<td>0.36</td>
<td>1.00E-10</td>
<td>7.9 ms</td>
</tr>
<tr>
<td>7</td>
<td>Type 2 diabetes</td>
<td>2p21</td>
<td>THADA</td>
<td>rs7578597-T</td>
<td>0.90</td>
<td>1.00E-09</td>
<td>0.115</td>
</tr>
<tr>
<td>7</td>
<td>Type 2 diabetes</td>
<td>7p15.1</td>
<td>JAZF1</td>
<td>rs864745-T</td>
<td>0.50</td>
<td>5.00E-14</td>
<td>0.110</td>
</tr>
<tr>
<td>7</td>
<td>Type 2 diabetes</td>
<td>3q14.1</td>
<td>ADAMTS9</td>
<td>rs4607103-C</td>
<td>0.76</td>
<td>1.00E-08</td>
<td>0.109</td>
</tr>
<tr>
<td>7</td>
<td>Type 2 diabetes</td>
<td>12q21.1</td>
<td>TSPAN6, LGR5</td>
<td>rs7961581-C</td>
<td>0.27</td>
<td>1.00E-09</td>
<td>0.109</td>
</tr>
<tr>
<td>7</td>
<td>Type 2 diabetes</td>
<td>10p13</td>
<td>CDC123, CAMK1D</td>
<td>rs12779790-G</td>
<td>0.18</td>
<td>1.00E-10</td>
<td>0.111</td>
</tr>
<tr>
<td>7</td>
<td>Type 2 diabetes</td>
<td>1</td>
<td>NOTCH2</td>
<td>rs10923931-T</td>
<td>0.11</td>
<td>4.00E-08</td>
<td>0.13</td>
</tr>
<tr>
<td>15</td>
<td>Type 2 diabetes</td>
<td>10q23.33</td>
<td>HHEX</td>
<td>rs5015480-C</td>
<td>0.43</td>
<td>6.00E-10</td>
<td>0.113</td>
</tr>
<tr>
<td>15</td>
<td>Type 2 diabetes</td>
<td>3q27.2</td>
<td>IGFBP2</td>
<td>rs4402960-T</td>
<td>0.32</td>
<td>9.00E-16</td>
<td>0.114</td>
</tr>
<tr>
<td>15</td>
<td>Type 2 diabetes</td>
<td>6p22.3</td>
<td>CDKAL1</td>
<td>rs10946398-C</td>
<td>0.32</td>
<td>4.00E-11</td>
<td>0.112</td>
</tr>
<tr>
<td>15</td>
<td>Type 2 diabetes</td>
<td>9p21.3</td>
<td>CDKN2B</td>
<td>rs10811661-T</td>
<td>0.83</td>
<td>8.00E-15</td>
<td>0.120</td>
</tr>
<tr>
<td>17</td>
<td>Type 2 diabetes</td>
<td>8q24.11</td>
<td>SLC30A4</td>
<td>rs31266634-C</td>
<td>0.30</td>
<td>6.00E-08</td>
<td>0.118</td>
</tr>
<tr>
<td>17</td>
<td>Type 2 diabetes</td>
<td>8q24.11</td>
<td>SLC30A4</td>
<td>rs31266634-C</td>
<td>0.30</td>
<td>6.00E-08</td>
<td>0.118</td>
</tr>
<tr>
<td>18</td>
<td>BMI</td>
<td>1q21.22</td>
<td>FTO*</td>
<td>rs9939609-A</td>
<td>0.39</td>
<td>2.00E-20</td>
<td>0.36 kg/m²</td>
</tr>
<tr>
<td>5</td>
<td>BMI</td>
<td>18q21.32</td>
<td>MC4R</td>
<td>rs17782313-C</td>
<td>0.24</td>
<td>3.00E-15</td>
<td>0.05 log (BMI)</td>
</tr>
<tr>
<td>8</td>
<td>LDL cholesterol</td>
<td>1p13.3</td>
<td>CERSR2, PSRC1†</td>
<td>rs599839-G</td>
<td>0.19</td>
<td>1.00E-33</td>
<td>0.16 mmol/L</td>
</tr>
<tr>
<td>9</td>
<td>HDL cholesterol</td>
<td>1q42.13</td>
<td>GALNT2</td>
<td>rs4846914-A</td>
<td>0.40</td>
<td>2.00E-13</td>
<td>0.07 SD</td>
</tr>
<tr>
<td>9</td>
<td>LDL cholesterol</td>
<td>19p13.11</td>
<td>CLP2, PX4</td>
<td>rs6996148-G</td>
<td>0.90</td>
<td>3.00E-08</td>
<td>0.10 SD</td>
</tr>
<tr>
<td>10</td>
<td>LDL cholesterol</td>
<td>1p13.3</td>
<td>CERSR2, PSRC1, SORT1</td>
<td>rs646776-C</td>
<td>0.24</td>
<td>3.00E-29</td>
<td>0.16 SD</td>
</tr>
<tr>
<td>10</td>
<td>HDL cholesterol</td>
<td>12q24.11</td>
<td>MVK, MMAB</td>
<td>rs2338104-G</td>
<td>0.45</td>
<td>3.00E-08</td>
<td>0.48 mg/dl</td>
</tr>
<tr>
<td>10</td>
<td>HDL cholesterol</td>
<td>6p21.32</td>
<td>B3GALT4</td>
<td>rs2254287-G</td>
<td>0.38</td>
<td>5.00E-08</td>
<td>1.01 mg/dl</td>
</tr>
<tr>
<td>10</td>
<td>LDL cholesterol</td>
<td>19p13.11</td>
<td>NCF1, CILP2</td>
<td>rs6996148-G</td>
<td>0.89</td>
<td>3.00E-09</td>
<td>3.32 mg/dl</td>
</tr>
<tr>
<td>6</td>
<td>Nicotine dependence</td>
<td>15q25.1</td>
<td>CHRNA3, CHRNA5, CHRNA4</td>
<td>rs1051730-T</td>
<td>0.35</td>
<td>6.00E-20</td>
<td>1 cigarette/day</td>
</tr>
</tbody>
</table>

Some of these associations have been found in several GWA studies; in this case the reference and respective data pertain only to the first study or only one of several concurrent studies. Genetic effects are given per allele copy except for Ref. 18 where they pertain to contrast of homozygotes. CAD indicates coronary artery disease; BMI, body mass index; LDL, low-density lipoprotein; HDL, high-density lipoprotein; OR, odds ratio; NR, not reported. It is not fully clarified yet whether the 2 entries in 1p13.3 for LDL cholesterol and the 2 entries for 9p21.3 for coronary artery disease or myocardial infarction have independent effects. SD, standard deviation.

*FTO was first found to be associated with type 2 diabetes, but the association was less consistent than the association for body mass index and obesity.
†The rs599839 polymorphism has also been reported (Ref. 11) to be associated with coronary artery disease (myocardial infarction).

With models that account for between-study heterogeneity, even better, one has to factor in the uncertainty in between-study heterogeneity, as can be done in a fully Bayesian meta-analysis, to derive a predicted interval for what are likely to be the underlying true effects is diverse populations. With these considerations, some 95% predicted intervals of the genetic effects of these associations may fall on both sides on the null, ie, in a non-negligible proportion of populations the genetic effect may be null or in the opposite direction than what has been observed. Most of the initially reported data on GWA signals is derived from white populations that share largely similar LD structures. Only 1 of the 28 associations was found to be significant in both white and Asian descent populations in the original GWA publication (Table 2). GWA-derived signals are almost always indirect markers in LD with the real with models that account for between-study heterogeneity.
culprits, because it is unlikely that the latter are directly hit. Therefore, it may not be uncommon that genetic effects for the tagging markers may be different in populations with different LD structure. Additional replication studies have started appearing for several of these variants on populations of different ancestry.22–25 Although most tend to replicate the proposed associations, differences in the effect sizes and frequency of the implicated alleles may not be uncommon. Such replication studies should be conducted as rigorously as the original GWA studies. Given the very strong probability values of the original findings, researchers, editors, and reviewers may be intimidated against report contradictory results. In the GWA era, such conformity26 would be disastrous for the credibility of replication efforts.

**Effect Sizes**

The effect sizes for the identified genome-wide markers are modest or small (Table 1), even smaller compared with those discovered in the pre-GWA era. Type 2 diabetes is a typical example. Documented genetic effects in the pre-GWA era represented per allele odds ratios (ORs) of 1.2 to 1.4. Genetic effects that emerged from the combination of data from the first 3 GWA studies15 represent ORs of 1.12 to 1.2. Finally, genetic effects that emerged from the replication of suggestive signals in meta-analysis of several GWA investigations7 represent ORs of 1.05 to 1.15. For lipid levels, identified polymorphisms have effects in the range of 0.05 to 0.16 standard deviations; for nicotine dependence the effect corresponds to 1 cigarette per day; and for body weight the effects corresponds to ~0.5 kg.

GWA-discovered effects are even likely to be inflated because of the winner’s curse in underpowered studies.27,28 It is possible that even sizeable studies and consortia are still underpowered for what are apparently subtle effects. The magnitude of the inflation can be appreciated only if we perform many additional replication studies. This has been done for the rs1333049 polymorphism in 9p21.3 that has one of the strongest discovered genetic effects in Table 1 (OR

<table>
<thead>
<tr>
<th>Reference</th>
<th>Disease/Trait</th>
<th>SNP-Risk Allele</th>
<th>( P )</th>
<th>( P ) for Heterogeneity</th>
<th>Ancestry*</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>CAD</td>
<td>rs6922269-A</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White</td>
</tr>
<tr>
<td>11</td>
<td>CAD</td>
<td>rs501120-T</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White</td>
</tr>
<tr>
<td>13</td>
<td>CAD</td>
<td>rs1333049-C</td>
<td>Single population†</td>
<td>Not applicable</td>
<td>White</td>
</tr>
<tr>
<td>14</td>
<td>Myocardial infarction</td>
<td>rs10757278-G</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White</td>
</tr>
<tr>
<td>12</td>
<td>Atrial fibrillation/flutter</td>
<td>rs1010033464-T</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White†‡</td>
</tr>
<tr>
<td>18</td>
<td>QT prolongation</td>
<td>rs10494366-G</td>
<td>Single population†</td>
<td>Not applicable</td>
<td>White</td>
</tr>
<tr>
<td>7</td>
<td>Type 2 diabetes</td>
<td>rs7578597-T</td>
<td>53%</td>
<td>0.008</td>
<td>White</td>
</tr>
<tr>
<td>7</td>
<td>Type 2 diabetes</td>
<td>rs864745-T</td>
<td>0%</td>
<td>0.70</td>
<td>White</td>
</tr>
<tr>
<td>7</td>
<td>Type 2 diabetes</td>
<td>rs4607103-C</td>
<td>25%</td>
<td>0.17</td>
<td>White</td>
</tr>
<tr>
<td>7</td>
<td>Type 2 diabetes</td>
<td>rs7961581-C</td>
<td>23%</td>
<td>0.20</td>
<td>White</td>
</tr>
<tr>
<td>7</td>
<td>Type 2 diabetes</td>
<td>rs12779790-G</td>
<td>0%</td>
<td>0.67</td>
<td>White</td>
</tr>
<tr>
<td>7</td>
<td>Type 2 diabetes</td>
<td>rs10923931-T</td>
<td>58%</td>
<td>0.004</td>
<td>White</td>
</tr>
<tr>
<td>15</td>
<td>Type 2 diabetes</td>
<td>rs5015480-C</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White</td>
</tr>
<tr>
<td>15</td>
<td>Type 2 diabetes</td>
<td>rs4402960-T</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White</td>
</tr>
<tr>
<td>15</td>
<td>Type 2 diabetes</td>
<td>rs10946398-C</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White</td>
</tr>
<tr>
<td>15</td>
<td>Type 2 diabetes</td>
<td>rs10811661-T</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White</td>
</tr>
<tr>
<td>17</td>
<td>Type 2 diabetes</td>
<td>rs13266634-C</td>
<td>Single population†</td>
<td>Not applicable</td>
<td>White</td>
</tr>
<tr>
<td>16</td>
<td>BMI</td>
<td>rs9939609-A</td>
<td>0%</td>
<td>Not significant</td>
<td>White</td>
</tr>
<tr>
<td>5</td>
<td>BMI</td>
<td>rs17782313-C</td>
<td>15%</td>
<td>0.24</td>
<td>White</td>
</tr>
<tr>
<td>8</td>
<td>LDL cholesterol</td>
<td>rs599839-G</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White</td>
</tr>
<tr>
<td>9</td>
<td>HDL cholesterol</td>
<td>rs4846914-G</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White</td>
</tr>
<tr>
<td>9</td>
<td>LDL cholesterol</td>
<td>rs16996148-G</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White</td>
</tr>
<tr>
<td>9</td>
<td>LDL cholesterol</td>
<td>rs846776-C</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White</td>
</tr>
<tr>
<td>10</td>
<td>HDL cholesterol</td>
<td>rs2338104-G</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White</td>
</tr>
<tr>
<td>10</td>
<td>LDL cholesterol</td>
<td>rs2254287-G</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White</td>
</tr>
<tr>
<td>10</td>
<td>LDL cholesterol</td>
<td>rs16996148-G</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White</td>
</tr>
<tr>
<td>10</td>
<td>Nicotine dependence</td>
<td>rs1333049-C</td>
<td>Not reported</td>
<td>Unknown</td>
<td>White</td>
</tr>
</tbody>
</table>

CAD indicates coronary artery disease; BMI, body mass index; LDL, low-density cholesterol; HDL, high-density cholesterol.

*For populations in the original article; some of these associations that were presented originally only with data on white populations have been subsequently replicated also in populations of other ancestry.

†Multistage GWA study, where the final replication study includes a single population (or samples that are treated as a single population).

‡No replication in Asian ancestry participants.
1.47 in the original GWA publication). A subsequent meta-
alysis\(^29\) in 12,004 cases and 28,949 controls resulted in an
OR of 1.24 with 95% CI of 1.20 to 1.29. Thus, the effect size
is likely to be only half compared with what was proposed
originally. The only other ORs exceeding 1.33 in Table 1 are
the 2 polymorphisms associated with the risk of atrial
fibrillation/flutter. Only one of them was found to be associ-
ated also in Asian descent populations (with smaller OR 1.42,
compared with 1.72 in Europeans). Surprisingly, no addi-
tional replication studies on these 2 polymorphisms have
been published in the first 15 months after the original
discovery published in Nature.

Small effect sizes have important corollaries. One may
infer that for diseases with a large heritability, a very large
number of such markers may exist, each with small or very
small effect size. This inference presumes that heritability
estimates are correct and that common variants are indeed
implied in explaining a substantial part of the genetic com-
ponent of these diseases. Second, one would need to conduct
extremely large studies and meta-analyses thereof, beyond
the current research capacity of existing epidemiological
studies, to unearth subtle effects, let alone reach accurate
effect estimates. With between-study heterogeneity, the
power is further eroded, and above a given threshold of
heterogeneity, it even becomes impossible to reach adequate
power.\(^30\) Some potentially genuine markers may thus remain
irreproducible.

**Population Attributable Fractions and Measures of
Individualized Prediction**

The population attributable fraction (PAF) is a popular metric
in the epidemiological literature for over 50 years. For a
nonconfounded, unadjusted single risk factor with relative
risk (RR), PAF is given by:

\[
\text{PAF} = \frac{\text{PR}(\text{RR} - 1)}{1 + \text{PR}(\text{RR} - 1)}
\]

where PR is the prevalence of the risk factor. For a case-
control study, the OR approximates the population RR.

There are many misunderstandings and misapplications.\(^31\)
PAF represents the proportion of the disease/phenotype of
interest that would have been avoided if the risk factor could
be eliminated from the population. PAF is thus attractive to
report and it is even highlighted in the abstracts of articles that
want to claim that they have discovered an important risk
factor. The abstract of Helgadottir et al\(^14\) in Science thus
claims that the discovered variant at chromosome 9 that is
associated with coronary artery disease has a PAF of 21% and
it is even 31% for early-onset cases. The PAF is largely
dependent on the frequency of the risk variant. Even with
small ORs, common variants will have large PAF estimates.
PAF is even higher in populations with higher risk allele
frequencies. For example, in the Nature article presenting 2
variants associated with atrial fibrillation/flutter,\(^13\) their joint
PAF is estimated to be 21% in Europeans and 35% in the
Chinese. In fact, in the Chinese, only 1 of the 2 variants is
replicated, and the other one has a much smaller OR than in
European populations, but the risk allele frequency is very
high (61%), 3 times higher than in Europeans.

A common error is to simply add the PAF values across
many markers. The PAF of 2 or several markers is not the
sum of the PAF of each. Second, even if the, correctly
calculated, overall PAF is high, this does not guarantee that
we can tell who will get the disease and who will not. If the
PAF is 100%, it simply means that eliminating the people
who carry any risk marker would eliminate the disease.
However, 90% of the population may carry at least 1 risk
marker. The disease may occur only in 2% of those, whereas
the other 88% are categorized as carrying a risk marker but
never develop disease.

In contrast to PAF, the additional penetrance due to a
variant may be a more direct measure of the magnitude of the
risk. For a dominant model, it can be shown that this is
approximated by the formula \(f = y(\text{OR} - 1)/(1 + y(\text{OR} - 1))\),
where \(y\) is the proportion of people who get the disease in the
absence of the genetic variant of interest.\(^32\) Typically, we can
assume that for small effects, \(y = \text{PR}\) approximately. There-
fore, for a marker with OR = 1.15 for a disease with preva-
ience of 1%, we get \(f = 0.15\%\) only. For a disease with
prevalence of 5%, an OR = 1.15 corresponds to \(f = 0.75\%\),
which is still quite small. In general, for the GWA signals
identified to-date, the \(f\) estimates are very small.

Similar inferences are obtained, when one calculates the
proportion of variance explained: a very small portion of
between-individuals’ variability of risk is explained by the
identified genetic variants. Typically each one of them
explains <1%, and for prediction of coronary artery disease,
type 2 diabetes or lipid levels; currently, we can explain
anywhere between <1% and 5% of the variance.

A useful way to show the small contribution of these
markers to individualized prediction is through the area under
the curve. Candidate predictors that have no predictive ability
have area under the curve (AUC) = 0.50, whereas perfect
prediction has AUC = 1.00. In type 2 diabetes, where we have
already 18 genetic variants with strong statistical support as
of September 2008, cumulatively all these variants can
achieve an AUC = 0.60, whereas simple knowledge of age,
gender, and body mass index can achieve AUC = 0.78,\(^33\) even
without counting family history.

Another way to look at this is the proportion of the sibling
relative risk explained: whereas the sibling relative risk is
\(\approx 3.0\) for type 2 diabetes, the 18 identified variants account
only for a sibling relative risk of \(\approx 1.07\). Overall, these estimates suggest that we capture only the tip of the iceberg
in explaining genetic risk for multifactorial diseases.

**Incremental Predictive Ability**

New candidate predictors would be useful only if they can be
shown to offer incremental information against other predic-
tors that are already available and easily measured. For
cardiovascular disease, the literature on predictive models is
very rich. Several widely validated scores are already avail-
able, such as the Framingham Risk Score for coronary artery
disease.\(^34\) Depending on the population tested, these predic-
tive scores usually already achieve AUC around 0.75. Pre-
diction of other common diseases such as type 2 diabetes
based on traditional risk factors also has AUC values in the
same range. The question is how much better we can do with genetic markers emerging from GWA investigations.

In the aforementioned evaluation of the 18 type 2 diabetes variants, the AUC improved from 0.78 to 0.80 with the consideration of the 18 markers on top of the age, gender, and body mass index. The improvement is highly statistically significant \((P=3 \times 10^{-12})\), but very modest in absolute terms. The same probably applies to the performance of 3 identified independent effects are smaller than what is usually pre-

sessed. Fourth, the model probably suffers from considerable error in the study population. In the presence of large nondifferential measurement error, classic risk factors such as LDL may have their predictive ability diluted. Conversely, genotypes are typically measured with high accuracy and the genetic effects do not shrink from measurement error. If so, then genotyping these variants would be an expensive way of correcting problems with measurement of lipids or other common risk factors that are already in everyday use.

**Predictive Reclassification**

Another important question in predictive modeling is whether a model including the genetic information can classify individuals in more or less appropriate risk categories and this may influence therapeutic or preventive decision making. This is different from the overall discriminating ability of the prediction as conferred for example by an AUC; conversely, this is an issue of reclassification. Reclassification analyses examine how many patients shift into a different risk category based on a predictive model compared with a standard one; how many of these are appropriately reclassified at higher or lower risk than before versus for how many the reclassification is in the wrong direction; and finally, whether there are changes in recommended decision-making based on this new information.

In the study of Kathiresan et al discussed earlier, the AUC did not change materially, but the authors noted that among patients who were originally classified in the intermediate-risk category, 26% were reclassified into a higher or lower risk category in the risk model containing the genotype score, as compared with the model without the genotype score. The net reclassification index that captures the extent of correct movements in predicted risk categories (higher risk for subjects in whom cardiovascular disease subsequently developed and lower risk for subjects free of incident cardiovascular disease) suggested a significant improvement in risk classification \((P=0.01)\). A significant improvement was also seen in the related integrated discrimination index \((P=0.02)\).

These are promising findings, but caveats exist. First, they are derived from the comparison of the performance of models that have all the limitations alluded to in the previous section. Second, only 9% of the study sample was originally classified in the III intermediate risk category. For most patients, reclassification, even if correct, would not have materially affected optimal decision making. More studies are needed to decipher if there is a reproducible window of risk
where genotype information from these or also additional markers may improve decision making.

In another study, 37 consideration of rs10757274 did not significantly improve the AUC for coronary heart disease versus using conventional risk factors alone (0.64 versus 0.62, P = 0.14), but the authors claimed improved reclassification, with 21.9% of the participants reclassified in different risk categories (<5%, 5% to 10%, 10% to 20%, >20%) and with 63% of these reclassifications being correct (P = 0.01). The very low observed discriminating performance (AUC = 0.62 only) of the selected conventional risk factors in this study is spurious. Moreover, the observed genetic effect for this marker was quite large in that study (OR 1.38 for heterozygotes). Other data suggest that this single marker alone does not improve reclassification 38; this is likely to be the typical case for single common variants. 39

The final proof would then lie in the demonstration that genotypic information indeed results in better clinical outcomes. Ideally, this would require randomized trials to be performed. Given that the genomic information is still in its infancy, randomized trials should be considered with caution. A “negative” result for clinical outcomes is very likely, but then it is difficult to tell whether this is proof that genotypic information indeed results in better clinical outcomes. Nevertheless, in the absence of robust randomized evidence, genomic tests for risk prediction should clearly be labeled as exploratory.

Conclusions
Cardiovascular genetics is undergoing a rapid transformation and robust findings have emerged, 5 -18,40 after 2 decades of mostly irreproducible findings. 41 -43 However, we have a very long road before applying meaningfully this information for predictive purposes at a population level. We are still dealing with markers of risk with subtle effects, and usually we do not know the culprit genetic variation that these markers are mirroring. Between-population heterogeneity is substantial, underappreciated or even inappropriately ignored in the gold rush for genome-wide statistical significance. Improvements in prediction based on the current markers are small, if at all present. Clinical portent is not yet sufficiently established. Although one can be excited about the new possibilities for more discoveries, incorporation of these markers in every day routine clinical practice and public health cannot be justified currently.

Sources of Funding
No funding was received for this review.

Disclosures
None.

References


CLINICAL PERSPECTIVE

A large number of common genetic variants associated with diverse phenotypes have been identified recently. A systematic review was performed of genome-wide association-derived markers associated with cardiovascular outcomes or other phenotypes that represent common established risk factors for cardiovascular outcomes. The review examined 28 independent associations with the very strongest statistical support, pertaining to coronary artery disease, myocardial infarction, atrial fibrillation/flutter, prolongation of QT interval, as well as type 2 diabetes, body mass index, high-density lipoprotein levels, low-density lipoprotein levels, and nicotine dependence. Between-study heterogeneity in the genetic effects seems common and it would pose a challenge to generalizability across different populations for these markers. Effect sizes are small and may be even smaller in subsequent replications and meta-analysis. Individualized risk measures are typically very small (proportion of variance explained $<1\%$ per marker). When used in conjunction with traditional predictors, improvement in overall prediction (e.g., area under the curve) or risk reclassification is limited, and subject to methodological caveats. Despite the excitement about these new discoveries, use for predictive purposes in clinical practice would probably be premature.
Prediction of Cardiovascular Disease Outcomes and Established Cardiovascular Risk Factors by Genome-Wide Association Markers

John P.A. Ioannidis

_Circ Cardiovasc Genet._ 2009;2:7-15; originally published online January 23, 2009; doi: 10.1161/CIRCGENETICS.108.833392

_Circulation: Cardiovascular Genetics_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2009 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/2/1/7

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/