Association of a Genetic Risk Score with Prevalent and Incident Myocardial Infarction in Subjects Undergoing Coronary Angiography

Riyaz S. Patel, MD1,2, Yan V. Sun, PhD1,3, Jaana Hartiala, MS4, Emir Veledar, PhD1, Shaoyong Su, PhD1, Salman Sher, MD1, Ying Liu, BS1, Ayaz Rahman, MD1, Ronak Patel, MD1, S. Tanveer Rab, MD1, Viola Vaccarino, MD, PhD1,3, A. Maziar Zafari, MD, PhD1,5, Habib Samady, MD1, W.H. Wilson Tang, MD6, Hooman Allayee, PhD4, Stanley L. Hazen, MD, PhD6, and Arshed A. Quyyumi, MD1

1Dept of Medicine, Emory University School of Medicine, Atlanta, GA
2Dept of Medicine, Cardiff University, Cardiff, Wales, United Kingdom
3Dept of Epidemiology, Emory University Rollins School of Public Health, Atlanta, GA
4Dept of Preventive Medicine and Institute for Genetic Medicine, Keck School of Medicine, University of Southern California, Los Angeles CA
5Dept of Medicine, Atlanta Veterans Affairs Medical Center, Decatur, GA
6Dept of Cardiovascular Medicine, Heart and Vascular Institute, Cleveland Clinic, Cleveland, OH

Abstract

**Background**—Genome-wide association studies have identified multiple variants associating with coronary artery disease (CAD) and myocardial infarction (MI). Whether a combined genetic risk score (GRS) is associated with prevalent and incident MI in high risk subjects remains to be established.

**Methods and Results**—In subjects undergoing cardiac catheterization (n=2597) we identified cases with a history of MI onset at age <70 years and controls ≥70 years old without prior MI, and followed them for incident MI and death. Genotyping was performed for 11 established CAD/MI variants and a GRS calculated based on average number of risk alleles carried at each locus weighted by effect size. Replication of association findings was sought in an independent angiographic cohort (n=2702). The GRS was significantly associated with prevalent MI, occurring before age 70 years, compared to older controls (≥70 years) with no history of MI (p<0.001). This association was successfully replicated in a second cohort yielding a pooled p value of <0.001. The GRS modestly improved the c-statistic in models of prevalent MI with traditional risk factors.
However, the association was not statistically significant when elderly controls without MI but with stable angiographic CAD were examined (pooled p=0.11). Finally, during a median 2.5 year follow-up, only a non-significant trend was noted between the GRS and incident events, which was also not significant in the replication cohort.

**Conclusions**—A GRS of 11 CAD/MI variants is associated with prevalent MI but not near term incident adverse events in two independent angiographic cohorts. These findings have implications for understanding the clinical utility of genetic risk scores for secondary as opposed to primary risk prediction.

**Keywords**
cardiovascular disease risk factors; coronary artery disease; myocardial infarction; Genetic Risk Score

**Introduction**

Genetic susceptibility for myocardial infarction (MI), the potentially fatal consequence of coronary artery disease (CAD), is due at least in part, to the cumulative effect of common polymorphisms with each risk variant conferring a small additive effect.\(^1,2\) Recent genome wide association studies (GWAS) have identified multiple single nucleotide polymorphisms (SNPS) associating with both MI and CAD.\(^3-9\) Combination of these variants into a risk score may thus permit examination of the cumulative effect of risk SNP burden on prevalent and incident MI.

Kathiresan et al. demonstrated that a genetic risk score (GRS) based on the 9 variants identified in their GWAS, was linearly associated with their primary phenotype of early onset MI.\(^9\) However, as the GRS was based on the same discovery cohort the effect size may have been overestimated. Replication in independent cohorts is therefore required to demonstrate that (1) a GRS derived from large GWAS can be applied to other independent populations, (2) such a GRS is also associated with non premature MI and (3) that a GRS that predicts prevalent MI or CAD is also able to predict future incident major adverse events.

Recent large prospective studies in younger low risk community cohorts have either confirmed or refuted the value of GWAS derived GRS in prediction of incident MI over 10-12 years.\(^10-12\) This discrepancy may be because these variants associate primarily with CAD and mechanisms underlying development of atherosclerosis, but not with MI per se. Attempts to dissect these two phenotypes require study of cohorts with detailed phenotyping for both MI and CAD.\(^13,14\) However, to date, no studies have examined the association between a GRS and either prevalent or incident MI in the context of detailed CAD phenotyping. Furthermore, little data exists on use of such a score to predict future prospective risk of adverse events in a CAD cohort in whom secondary risk prediction algorithms are currently inadequate.\(^15\)

We therefore sought to examine the association between a genetic risk score, based on 11 CAD/MI variants discovered through GWAS, with (1) prevalent MI with onset <70 years of age and (2) incident MI and death in a cohort of subjects undergoing coronary angiography for diagnosis or treatment of CAD.
Methods

Study Design

All primary and subgroup association analyses were performed in the Emory Cardiovascular Biobank (Atlanta, Georgia, USA) with replication of key findings sought in a second independent cohort from the Cleveland Clinic Genebank (Cleveland, Ohio, USA). Significant association results were then pooled through meta-analysis.

Study Samples

The Emory Cardiovascular Biobank consists of over 5000 consecutively recruited patients undergoing cardiac catheterization at Emory University and affiliated hospital sites since 2003 with documented demographic characteristics, medical histories, behavioral factors, and fasting blood samples. Specific details regarding risk factor definitions and coronary angiographic phenotyping have been described previously. Family history (FH) of CAD was defined as having any first degree relative (parents, siblings) with MI or coronary revascularization prior to the age of 60 years. All subjects were prospectively followed by telephone interview, chart abstraction and through state vital records data to document major cardiovascular events and death. Follow-up was performed at time periods ranging from 1-5 years. Events were considered as outcomes if they occurred greater than 72 hours after date of discharge. Patients with heart transplantation and missing or incomplete phenotype data were also excluded. As most risk variants have not yet been validated in non-Caucasian populations, we excluded those with self reported non-Caucasian ancestry. The study was approved by the Institutional Review Board and all subjects provided written informed consent.

The Cleveland Clinic GeneBank study is a single site sample repository generated from consecutive patients undergoing elective diagnostic coronary angiography or elective cardiac computed tomographic angiography with extensive clinical and laboratory characterization and longitudinal observation. Subject recruitment occurred between 2001 and 2006. The GeneBank cohort has been used previously for discovery and replication of novel genes and risk factors for atherosclerotic disease. Ethnicity was self-reported and information regarding demographics, medical history, and medication use was obtained by patient interviews and confirmed by chart reviews. Angiographic CAD was determined as evidence of ≥50% stenosis of one or more major epicardial vessel. Prospective cardiovascular risk was assessed by the incidence of major adverse cardiac events (MACE) during three years of follow up from the time of enrollment, which included MI, stroke, and all-cause mortality. Adjudicated outcomes ascertained over the ensuing 3 years for all subjects following enrollment were confirmed using source documentation. All patients provided written informed consent prior to being enrolled in GeneBank and the study was approved by the Institutional Review Board of the Cleveland Clinic.

Prevalent MI

MI was consistently defined according to standard criteria from medical records as documented history of acute (at enrollment) or remote MI. Age at first MI was recorded to the closest year. Subjects with MI onset at age <70 years were defined as cases. We also specified a-priori subgroup analyses to explore association with more premature MI defined as cases with MI <60 and <50 years of age. To mitigate case misclassification bias, whereby younger control subjects could become cases by developing MI in later life, we selected subjects ≥70 years at enrollment without a history of prior MI (or during the study period), as controls. These controls included those without ever experiencing an MI, but who may have had significant angiographic CAD. Given that our statistical test of MI versus no MI would also test the hypothesis of CAD versus no CAD, we utilized previously described
methods to examine the effect of the risk score on odds of MI in the presence of significant CAD determined as angiographic CAD ≥50% luminal stenosis in any major epicardial vessel (MI+/CAD+ versus MI-/CAD+).13, 14

Incident MI

All patients in both the Emory and Cleveland Clinic cohorts were prospectively followed for adverse events, as described above. The primary endpoint for incident event analysis in this study was defined as incident MI, using standard criteria as described above. The composite outcome of all cause death and MI was defined as a secondary endpoint.

SNP Selection and Genotyping

We selected CAD/MI risk variants discovered through GWAS and published before June 2009, whose primary phenotype was MI or a combination of CAD/MI and which satisfied the genome wide association threshold of 5×10^{−8}, (reflecting a million tests of association). The selected SNPs are listed in Table 1 along with risk allele frequencies for each cohort with published effect sizes. The LPA locus was not included for logistical reasons despite publication in 2009.24, 25

Genotyping for Emory Cardiovascular Biobank Samples was performed using the Centaurus (Nanogen) platform at deCODE genetics, Reykjavik, Iceland or with the SNPstream (Beckman Coulter) platform at Emory University, Atlanta, GA.26, 27 Genotyping for Cleveland Clinic GeneBank samples was performed using the Affymetrix Genome-Wide Human SNP Array 6.0 platform at the University of Ottawa with imputation for untyped SNPs using MACH 1.0 software. Full details are provided in the Supplementary Methods.

Genetic Risk Score (GRS)

A weighted GRS was calculated based on the average number of risk alleles (0, 1 or 2) in the SNPs genotyped weighted by the effect size published in larger scale studies (natural log of the OR x no. of risk alleles for each SNP), using the “Profile Scoring” option in PLINK.28 Thus, 2 copies of the risk allele 9p21 (published OR per risk allele 1.3) contributed proportionately more to the overall GRS for an individual than two copies of the risk allele 6p24 (OR 1.1). For the Emory cohort, missing genotype data for individual SNPs was accounted for by mean imputation based on overall risk allele frequency. Details of missingness are provided in Supplementary Table 1. We performed validation analysis using only data from subjects with the full complement of genotyped variants to ensure imputation did not adversely influence our results. Cleveland Clinic genotype data was derived from imputation of a GWAS dataset (Supplementary Methods).

Statistical Analysis

Continuous variables are presented as means (SD) and categorical variables as proportions (%) with one way ANOVA and chi-squared tests used to determine differences between groups.

Prevalent MI association analysis—Logistic regression models were constructed to first test the effect of the individual SNPs (under an additive model, with risk allele coding 0,1,2) and then the weighted GRS, as a continuous variable on prevalent MI risk (case= MI<70yrs; controls = no MI ≥70yrs). We further examined the association with the GRS in younger cases, defined as MI<60 years and MI<50 years against the same control group. Replication analyses were performed in the Cleveland Clinic Genebank, with a random effects meta-analysis to combine results in order to derive a final pooled estimate. For illustrative purposes, we then categorized into quintiles of GRS to derive estimates for effect
size (OR) for the top quintile compared to the first quintile as the referent category. We also tested for association between the continuous GRS and MI in all patients with CAD (MI+/CAD+ versus MI-/CAD+), again with attempted replication in Cleveland Clinic Genebank. This was performed using the same cases (MI<70yrs) but selecting controls (≥70 yrs) without prior MI but with significant stable angiographic CAD (defined as ≥50% stenosis).

Analyses were performed before and after adjusting for gender, body mass index, diabetes, hypertension, hyperlipidemia, smoking and family history (FH) of premature MI/CAD. Interaction terms were tested for association between the risk score and risk factors including FH.

Incremental predictive value of the GRS for prevalent MI occurring before the age of 70 in the Emory cohort was tested using ROC analysis and estimated through the c-statistic. Models were constructed with and without the risk score using (1) traditional risk factors (gender, BMI, diabetes, hypertension, hyperlipidemia, smoking and FH), and (2) with gender and FH only, which are the only variables likely to be known in early life, when risk prediction could be employed.

Incident MI risk analysis—Cox regression analyses were applied to identify associations between the individual SNPs (under an additive model, with risk allele coding 0,1,2) and then the GRS with risk of incident MI and combined death and MI during follow-up. Models were constructed using both a continuous and a categorized GRS, with and without traditional risk factors and presence of baseline angiographic CAD (≥50%). Hazard Ratios (HR) were estimated for the top quintile of the GRS referenced to the first quintile as before. Subgroup analyses included stratifying by history of MI at enrollment to examine risk of recurrent and first MI during follow up and examining incident risk in younger subjects <60 years. The proportional hazards assumption was evaluated graphically using log-log plots.

A 2-tailed P value <0.05 was considered significant. PLINK statistical software was used to estimate Hardy-Weinberg equilibrium and minor allele frequencies for all SNPs. All remaining statistical analyses were performed using SAS (Cary, NC) and SPSS 17.0 (Chicago, IL).

Results

For the 11 SNPs genotyped in the Emory and Cleveland Clinic cohorts, the observed frequencies for each risk allele were similar to published and Hap-Map data. All SNPs were in Hardy Weinberg equilibrium (HWE) with p-value threshold of 0.01 considering multiple testing. Only one SNP, rs3184504, in the Cleveland Clinic Genebank had a nominal p < 0.05 (Table 1). A total of 2597 self-reported Caucasians (mean age 63.9 ± 11.1 years and 67% male) were genotyped from the Emory Cardiovascular Biobank and 2702 (mean age 63.0 ± 11.2 years and 72% male) from the Cleveland Clinic Genebank. Although the two cohorts were generally similar, differences were noted including for example a greater prevalence of treated hyperlipidaemia and smoking in the Cleveland Clinic Genebank compared to those in the Emory Cardiovascular Biobank (Table 2).

Prevalent MI association with a genetic risk score

For prevalent MI (cross sectional) analysis, we examined 889 cases of MI<70 years and 449 controls ≥70 years without prior MI from the Emory Cardiovascular Biobank (Table 2). The replication analysis included 1154 cases and 495 controls from Cleveland Clinic Genebank. Cases were significantly younger than controls by design, and expected significant differences in other risk factors were also observed (Table 2). Individual SNP associations with prevalent MI are presented in Supplementary Table 2. While trends were noted for association with most of the SNPS, statistically significant associations were noted only for
SNPs at loci 1q41 (rs17465637), 6p24 (rs12526453), 9p21 (rs10757278) and 21q22 (rs9982601) in the Emory cohort and only for 9p21 (rs10757278) in the Cleveland Clinic Genebank.

In our primary analysis there was a significant association between the overall weighted GRS and prevalent MI, when modeled as a continuous variable (p=9.74×10^{-5}) (Table 3). A modest but statistically significant association was identified with prevalent MI using the same definitions and analysis methods (p=0.04). When combined in the pooled analysis, there was a significant overall association for the GRS with prevalent MI (p=4.5×10^{-5}) (Table 3). These findings persisted after adjustment for risk factors (pooled p value=1.4×10^{-4}).

To illustrate this association further, we grouped the Emory cohort by quintile of GRS. In comparison to the lowest quintile, the adjusted OR for MI for an individual in the highest quintile of GRS was 1.81 (95% CI 1.21-2.70) (Figure 1). Furthermore, the effect sizes were greater in cases with MI <60 years or <50 years of age compared to older controls. Thus, the adjusted OR for MI for the highest quintile of GRS compared to the lowest was 1.90 (95% CI 1.24-2.92) for MI<60 years and 2.37 (95% CI 1.44-3.91) for MI<50 years (Figure 1). Interestingly, it appears that a significant increase in odds of MI was only observed in the highest quintiles, with no apparent increase between the first 3 quintiles (Figure 1).

We did not observe significant interactions between the GRS and covariates in the association with MI <70 years. Most pertinent was the lack of interaction for family history (FH) of CAD (p for interaction 0.63). Overall, the mean weighted GRS was marginally higher in subjects with a positive FH compared to those with a negative or unknown FH (0.085±0.01 v 0.083±0.01 respectively, p=0.05).

Given the significant association with prevalent MI, we examined the additive value of the GRS when modeled with existing risk factors for prevalent MI risk at <70 years using the area under the curve statistic. There was a statistically significant increase in the c-statistic when the GRS was added to a model containing gender, BMI, traditional risk factors and FH (0.681 (0.015) v 0.693 (0.015), p=0.012) (Table 4). Similarly, when modeled for MI at younger ages, there was again a significant increase in the c-statistic (<60yrs 0.699 (0.016) v 0.712 (0.016) p=0.03; <50yrs 0.716 (0.018) v 0.733 (0.018), p=0.04) (Table 4). The difference between the AUC was even greater when the GRS was modeled with the non modifiable risk factors of FH and gender (difference in AUC = 0.027; 0.030; 036 for MI <70; <60 and <50yrs respectively, all p<0.005) (Table 4).

Finally, to determine if these variants confer risk of MI (by plaque rupture or thrombosis) over and above the risk of developing CAD, we repeated the main analysis using older CAD controls (methods). We examined 889 cases (MI<70yrs) and 370 controls (no MI ≥70 years with CAD+; Supplementary Table 3). Overall the results were significantly attenuated with a modest association identified only in the Emory cohort (p=0.02). This finding however was not replicated in the Cleveland Clinic cohort (1154 cases and 241 controls; Supplementary Table 3), with a final non significant association in the pooled adjusted analysis (p=0.11) (Table 3).

**Incident MI: Association with a genetic risk score**

For *incident* MI (prospective) analysis we followed the entire Emory cohort of 2597 subjects for adverse outcomes. Demographic data by quintile of risk score for the Emory cohort are presented in Supplementary Table 4, for which there was a statistically significant difference in age with increasing risk score, whereby those with lower scores were older than those
with higher scores (p trend =0.001). Also, those with the highest scores tended to have higher incidence of hyperlipidemia.

After a median of 2.5 years follow-up, 101 subjects experienced an MI and 358 the composite endpoint of death or MI. For the primary endpoint of incident MI there was no significant association for either the individual SNPS (Supplementary Table 2) or the cumulative GRS (adjusted p=0.68). When grouped into quintiles of GRS, compared to the lowest quintile, subjects in the highest quintile trended towards more events with an adjusted HR of 1.08 (95% CI 0.57-1.99) for incident MI (Table 5). In subgroup analysis, there was no association between the GRS and future recurrent or first MI during follow up (p= 0.67 and p=0.21 respectively).

For the composite endpoint of Death/MI however, there was a significant, paradoxically, protective association with GRS (adjusted β= −11.4(3.98), p=0.004), with an adjusted HR of 0.68 (95% CI 0.49-0.94) for top v bottom quintile. Other significant multivariate predictors in the latter model included age (HR 1.03 (1.02-1.04)), diabetes (HR 1.51 (1.20-1.90)), hypertension (HR 1.28 (1.00-1.65)), smoking (HR 1.39 (1.10-1.75)) and baseline significant angiographic CAD (HR 1.31 (1.0-1.73)). A history of hyperlipidaemia was also noted to be protective (HR 0.78 (0.61-0.99)).

In the Cleveland Clinic cohort, we also observed a trend towards decreasing age with increasing quintile of risk score (p=0.01) (Supplementary Table 5). After 3 years of follow up in 2702 subjects there were 124 incident MIs and 284 composite end points of Death/MI. There was again no significant association with individual SNPs (Supplementary Table 2) or the cumulative GRS (adjusted p=0.40) with incident MI, although there was again a non significant trend with more events in the top quintile yielding an adjusted HR of 1.28 (95% CI 0.85-1.93) compared to the lowest quintile. For the composite endpoint of Death/MI there was also no significant association (p=0.91) with an adjusted HR of 1.07 (0.80-1.44).

Analysis of those <60 years of age at enrollment revealed no significant associations in either cohort for incident MI or Death/MI (Table 5).

Finally, we repeated the analyses for prevalent and incident MI using only subjects with full genotype data for all 11 variants in the Emory cohort (Supplementary Table 1) and identified similar findings, thereby confirming that mean imputation did not appreciably influence our results (data not shown).

**Discussion**

In a study of high risk angiographically phenotyped patients, we demonstrate that a cumulative risk score generated from published GWAS variants is: (1) significantly associated with MI before the age of 70 years in two Caucasian cohorts, independent from the discovery samples; (2) not associated with MI when both cases and controls have established CAD; and (3) not a strong predictor of incident near term adverse events.

Although genetic risk scores using variants identified through genome wide linkage scans and candidate gene studies have been previously studied, the first attempt to devise a cumulative risk score from GWAS variants was by Kathiresan et al, using 9 risk variants for premature MI. They found that subjects within the highest quintile of a weighted risk score had two-fold higher odds of premature MI than those in the lowest quintile. Our study validates the use of such a score in two independent cohorts after addition of two extra variants that have since been identified. In comparison, our cohorts were significantly older, but we still showed that subjects within the highest quintile of risk score also had an almost two fold greater odds of MI prior to the age of 70 years than those in the lowest
quintile, with an even greater risk in subjects with more premature disease (MI <60 and <50 years), thereby confirming that hereditary effects are greatest in younger age. Interestingly the association observed for prevalent MI in our study was non-linear, with only the highest quintile of risk score significantly associating with MI. These findings may reflect possible gene-gene interactions or simply more extensive CAD prone to complications, such as a greater lipid core or thin cap fibro-atheroma arising as a consequence of inheriting a greater than average number of these atherosclerosis predisposing risk variants.

Use of the broad phenotype of “coronary heart disease” by early large scale GWAS has raised concerns as to whether these risk variants confer risk of MI or just CAD, since the processes involved in atherosclerosis may differ from those that lead to MI. We and others showed that 9p21, for example, associates primarily with CAD and atherosclerosis. Moreover in subjects with angiographically significant CAD, the 12 risk variants individually do not associate with MI. Thus, at the pathophysiologic level, these variants are unlikely to be associated with thrombosis or plaque rupture, and MI risk observed in prior studies is likely secondary to development of greater CAD burden. This study extends these observations by demonstrating that the association between the cumulative risk score and MI is considerably attenuated and not replicated in a second cohort. In contrast to the study of Reilly et al, where controls were younger and thus subject to possible case-misclassification, use of older controls is a major strength of this study, given the critical time sensitive nature of the MI phenotype. Although this finding is of major interest for mechanistic understanding, MI is a more tangible clinical endpoint and as such association of a risk score with MI, even by proxy may still be clinically relevant.

Although subjects reporting a positive FH of CAD/MI had a marginally higher risk score, the cross sectional association between the risk score and MI was independent of FH, and furthermore, there was no interaction with FH, suggesting that genetic risk is mediated both in those with and without a FH of premature CAD. Despite the challenges related to accurate documentation of FH, this serves to highlight the complexity of heritability of MI in relation to common variants. Identification of rarer variants with more Mendelian characteristics through next generation sequencing may yield further insights into the role of FH in relation to complex diseases.

Genetic risk score and incident MI

Three large studies have examined the value of genetic risk scores in predicting first MI in asymptomatic populations. In a prospective study of 19,000 community based women, no association was observed between a panel of 12 MI risk variants and future adverse CV events after 12 year follow-up. However, Ripatti et al found that a risk score of 13 variants was associated with prevalent and incident CV events in a community based population of 31,000 men and women; subjects in the highest quintile were at 1.66-fold greater risk of adverse events during a 10-year follow-up period compared to those in the lowest quintile. An association between incident MI risk and 13 SNPs was also observed in the Framingham cohort. In contrast to our population who had established CAD or were at high risk of CAD where future risk prediction is particularly challenging, all of these studies examined primary risk of MI.

In this study, although we observed a very marginal trend towards more incident MI in those with the highest quintile of risk score, the final association results were not statistically significant in both cohorts. This is likely to reflect the fact that the primary phenotype for most of these variants is atherosclerosis, which had already developed by the time of enrollment into a cardiac catheterization cohort. Interestingly, we noted that the distribution of risk scores in both the Emory and Cleveland Clinic cohorts was related to age with older subjects trending towards lower scores, potentially indicating that those with high scores
were at greater risk for premature fatal events. This finding may help to explain the paradoxically protective association for those with the highest risk scores with the composite outcome of Death/MI in the Emory cohort. Given that a history of hyperlipidemia was also protective in the same model, it is plausible that unmeasured treatment effects such as intensive statin use may also be responsible. Overall, the observation of a lack of association between genetic variants and incident events in a diseased cohort has also been described in other studies. For example, the 9p21 risk is associated with incident risk prediction in primary prevention populations but not in those with diagnosed CAD.32

Finally, we also tested the discriminatory value of the genetic risk score for identifying cases of prevalent MI, which may be greater than for single SNPS as advocated by Davies et al.33 Indeed, after addition of the genetic risk score to traditional risk factors, we found a modest, but statistically significant improvement in the c-statistic for prevalent MI <70 yrs as well as more premature MI. However, genetic data is unique and provides risk information at birth, unlike the majority of acquired risk factors that impact later in life. When the genetic risk score was added to a model containing the non modifiable risk factors of gender and FH (the only ones likely to be known in adolescence), we found that it markedly enhanced the c-statistic, with the greatest discriminatory value for more premature MI. Use of genomic data in this manner would represent a shift in genetic biomarker assessments and requires further exploration.

Limitations & Strengths

There are some limitations of our study. Firstly, both our discovery and replication cohorts are selected populations with CAD and thus our findings may be susceptible to unmeasured bias. Secondly, our controls were not healthy, although risk allele frequency would likely be even lower in healthier controls and would strengthen the observed statistical associations. Thirdly, we used imputation to estimate risk scores for those with missing data. However, an analysis using non imputed data revealed similar overall results. Fourthly, while we may have been underpowered for analysis of incident events, our study was powered sufficiently to be able to detect the effects of age/10 yr to a HR of 1.16. Finally, we only included 11 MI SNPs in our score and have not included variants in relevant metabolic pathways or newer and other recently discovered CAD/MI variants.24, 25, 34 However, the recent study by Thannassoulis et al demonstrates that addition of 16 newer variants to a 13 SNP model did not impact on their association findings.12 Moreover, some of the SNPs examined may not to be true risk SNPs in a particular locus, possibly attenuating the results.35

Major strengths of our study include (1) study of independent discovery and replication cohorts (2) the use of elderly controls in the cross sectional analysis to mitigate the important problem of case-misclassification bias (3) examination of the association with the risk score in the context of proven CAD to determine if there was MI risk over and above association with atherosclerosis, which to our knowledge has not been performed previously and is important given the overlap in phenotypes and (4) study of both prevalent and incident MI in the same cohorts to avoid concerns including those of population stratification.

Conclusion

In conclusion, in two independent CAD cohorts, we demonstrate that a high burden of CAD/MI risk variants is associated with prevalent MI but not with near term incident MI or death. We propose that cumulatively these variants confer risk of MI through development of CAD, but once disease has developed the association with future risk of adverse events is attenuated. As such the clinical utility of a GRS is of limited value in secondary risk
prediction. These findings have implications for understanding the clinical utility of genetic risk scores for secondary as opposed to primary risk prediction.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Acknowledgments**

We would like to thank all our research coordinators, nursing staff and volunteers for help with patient recruitment and sample collection for the Emory Cardiovascular Biobank. We would also like to thank the attendings and fellows from the Cardiology Division of Emory University as well as the cardiac catheterization laboratory staff for their support and assistance in recruiting patients for this study.

**Funding Sources:** Emory: This work was supported by the American Heart Association (Postdoctoral Fellowship for RSP), National Institutes of Health R01 HL89650-01, Robert W. Woodruff Health Sciences Center Fund, Emory Heart and Vascular Center Funds and supported in part by NIH Grant UL1 RR025008 from the Clinical and Translational Science Award program and NIH grant R24HL085343. Cleveland Clinic: GeneBank has been supported by National Institutes of Health grants P01HL087018-020001, P01HL076491-055328, R01HL103866, RO1HL103931, and the Cleveland Clinic Clinical Research Unit of the Case Western Reserve University CTSA (1UL1RR024989). SLH has also been supported in part by a grant from the LeDucq Foundation, and a gift from the Leonard Kreiger Fund.

**References**


Over the last 5 years, genome wide association studies or GWAS have identified multiple novel variants associating with coronary artery disease and myocardial infarction (MI). Given the small effect sizes individually, there is interest in combining these multiple variants into genetic risk scores for predicting events and determining an individual’s personalized risk. With increasing direct-to-consumer testing and growing physician requested testing for genetic variants; this approach is likely to be of increasing relevance. In this paper, we sought to explore the association between a cumulative genetic risk score composed of 11 of the earliest and most replicated variants with both prevalent and incident MI risk in two independent populations of subjects at high risk for or with established CAD. We demonstrate robust association with prevalent MI occurring before the age of 70, confirming the cumulative impact of genetic burden over a lifetime. However we only demonstrate non significant trends for association with near term MI, which is in contrast to some population studies which have shown significant association with incident MI. We propose that cumulatively these CAD/MI variants confer risk of MI through development of CAD, but once disease has developed the association with future risk of adverse events is likely attenuated. These findings therefore have implications for understanding the clinical utility of genetic risk scores for secondary as opposed to primary risk prediction.
Figure 1.
Association between quintiles of genetic risk score and MI in different age groups within the Emory Cardiovascular Biobank. Adjusted Odds Ratios and 95% confidence intervals are presented. Covariates included gender, BMI, diabetes, hypertension, hyperlipidemia, smoking and family history of CAD.
Table 1

Details of the selected myocardial infarction loci and their risk variants in each cohort

<table>
<thead>
<tr>
<th>Locus</th>
<th>SNP ID</th>
<th>Nearest Gene</th>
<th>Risk allele</th>
<th>Emory CV (Published risk allele frequency, Observed frequency, HWE p value)</th>
<th>Cleveland Clinic (Published risk allele frequency, Observed frequency, HWE p value)</th>
<th>Reported allelic MI risk (OR)</th>
<th>Ref*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1p13</td>
<td>rs646776</td>
<td>CELSR2-PSRC1-SORT1</td>
<td>T</td>
<td>0.81, 0.80, 0.18</td>
<td>0.80, 0.80, 0.80</td>
<td>1.19 (1.13-1.26)</td>
<td>9</td>
</tr>
<tr>
<td>1p32</td>
<td>rs11206510</td>
<td>PCSK9</td>
<td>T</td>
<td>0.81, 0.82, 0.19</td>
<td>0.83, 0.12, 1.15</td>
<td>1.15 (1.10-1.21)</td>
<td>9</td>
</tr>
<tr>
<td>1q41</td>
<td>rs17465637</td>
<td>MIA3</td>
<td>C</td>
<td>0.72, 0.73, 0.89</td>
<td>0.73, 0.26, 1.14</td>
<td>1.14 (1.10-1.19)</td>
<td>9</td>
</tr>
<tr>
<td>2q33</td>
<td>rs6725887</td>
<td>WDR12</td>
<td>C</td>
<td>0.14, 0.14, 0.33</td>
<td>0.15, 0.83, 1.17</td>
<td>1.17 (1.11-1.23)</td>
<td>9</td>
</tr>
<tr>
<td>3q22.3</td>
<td>rs9818870</td>
<td>MRAS</td>
<td>T</td>
<td>0.15, 0.16, 0.89</td>
<td>0.16, 0.53, 1.15</td>
<td>1.15 (1.11-1.19)</td>
<td>7</td>
</tr>
<tr>
<td>6p24</td>
<td>rs12526453</td>
<td>PHACTR1</td>
<td>C</td>
<td>0.65, 0.67, 0.09</td>
<td>0.67, 0.21, 1.12</td>
<td>1.12 (1.08-1.17)</td>
<td>9</td>
</tr>
<tr>
<td>9p21</td>
<td>rs10757278</td>
<td>CDKN2A/B</td>
<td>G</td>
<td>0.50, 0.50, 0.53</td>
<td>1.00, 1.29, 1.22</td>
<td>1.29 (1.22-1.35)</td>
<td>3</td>
</tr>
<tr>
<td>10q11</td>
<td>rs1746048</td>
<td>CXCL12</td>
<td>C</td>
<td>0.84, 0.87, 0.66</td>
<td>0.86, 0.64, 1.17</td>
<td>1.17 (1.11-1.24)</td>
<td>9</td>
</tr>
<tr>
<td>12q24</td>
<td>rs3184504</td>
<td>SH2B3</td>
<td>T</td>
<td>0.47, 0.49, 0.30</td>
<td>0.52, 0.03, 1.13</td>
<td>1.13 (1.08-1.18)</td>
<td>8</td>
</tr>
<tr>
<td>19p13</td>
<td>rs1122608</td>
<td>LDLR</td>
<td>G</td>
<td>0.75, 0.75, 1.0</td>
<td>0.76, 0.22, 1.15</td>
<td>1.15 (1.10-1.20)</td>
<td>9</td>
</tr>
<tr>
<td>21q22</td>
<td>rs9982601</td>
<td>SLC5A3-MRPS6-KCNE2</td>
<td>T</td>
<td>0.13, 0.13, 0.25</td>
<td>0.15, 0.66, 1.20</td>
<td>1.20 (1.14-1.27)</td>
<td>9</td>
</tr>
</tbody>
</table>

*References relate to studies utilizing myocardial infarction as the primary phenotype, from which values in the table are derived, except 3q22.3 for which a combined phenotype was used.
Table 2
Details of demographics for each cohort and by case-control status for prevalent MI analysis

<table>
<thead>
<tr>
<th></th>
<th>Emory Cardiovascular Biobank</th>
<th>Cleveland Clinic Genebank</th>
<th>Overall Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall N=2597</td>
<td>Cases (MI&lt;70 yrs) N=889</td>
<td>Controls (no MI ≥70 yrs) N=449</td>
</tr>
<tr>
<td>Age, yrs</td>
<td>63.9 (11.3)</td>
<td>62.9 (10.0)</td>
<td>75.7 (4.20)</td>
</tr>
<tr>
<td>Male Gender</td>
<td>67.6</td>
<td>76.2</td>
<td>59.0</td>
</tr>
<tr>
<td>Body Mass Index, kg/m2</td>
<td>28.9 (5.7)</td>
<td>29.7 (5.8)</td>
<td>27.6 (5.3)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>30.2</td>
<td>35.9</td>
<td>27.2</td>
</tr>
<tr>
<td>Hypertension</td>
<td>70.6</td>
<td>70</td>
<td>73.9</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>68.2</td>
<td>77.8</td>
<td>69.0</td>
</tr>
<tr>
<td>Ever Smoked</td>
<td>61.2</td>
<td>67.8</td>
<td>55.6</td>
</tr>
<tr>
<td>Family History CAD &lt;60yrs</td>
<td>15.3</td>
<td>20.5</td>
<td>8.7</td>
</tr>
</tbody>
</table>

Mean and standard deviations or % presented unless otherwise indicated. Overall comparison for Emory versus Cleveland Clinic cohorts.
Table 3
Association between weighted risk scores and prevalent MI <70 years with and without consideration of CAD

<table>
<thead>
<tr>
<th></th>
<th>Emory Cardiovascular Biobank</th>
<th>Cleveland Clinic Genebank</th>
<th>Pooled Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta</td>
<td>SE</td>
<td>P</td>
</tr>
<tr>
<td>MI &lt;70 yrs v no MI ≥70 yrs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(All)</td>
<td>17.05</td>
<td>4.37</td>
<td>9.74×10⁻⁵</td>
</tr>
<tr>
<td>MI&lt;70 yrs v no MI ≥70 yrs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(CAD+)</td>
<td>11.73</td>
<td>4.96</td>
<td>0.02</td>
</tr>
</tbody>
</table>

*See methods. Adjusted in both groups for gender, BMI, diabetes, hypertension, hyperlipidaemia, smoking
Table 4

Discriminatory analysis using AUC for prevalent MI in the Emory Cardiovascular Biobank

<table>
<thead>
<tr>
<th></th>
<th>Traditional RFs</th>
<th>Traditional RFs + GRS</th>
<th>Improvement in C statistic</th>
<th>P value</th>
<th>Early RFs</th>
<th>Early RFs + GRS</th>
<th>Improvement in C statistic</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MI&lt;70</td>
<td>0.681(0.015)</td>
<td>0.693(0.015)</td>
<td>0.012</td>
<td>0.012</td>
<td>0.612</td>
<td>0.639</td>
<td>0.027</td>
<td>0.002</td>
</tr>
<tr>
<td>MI&lt;60</td>
<td>0.699(0.016)</td>
<td>0.712(0.016)</td>
<td>0.013</td>
<td>0.013</td>
<td>0.625</td>
<td>0.655</td>
<td>0.030</td>
<td>0.002</td>
</tr>
<tr>
<td>MI&lt;50</td>
<td>0.716(0.018)</td>
<td>0.733(0.018)</td>
<td>0.017</td>
<td>0.017</td>
<td>0.639</td>
<td>0.675</td>
<td>0.036</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Traditional Risk Factors (RFs) = gender, BMI, diabetes, hypertension, hyperlipidaemia, smoking and family history (FH) CAD; Early RFs = Gender and FH CAD only; GRS = Genetic Risk Score; MI<70/60/50 = myocardial infarction occurring before respective age compared to older controls (≥70) without prior MI; P values for comparison of AUC.
Table 5

Prospective follow up events and association with weighted risk score by quintiles

<table>
<thead>
<tr>
<th></th>
<th>Emory Cardiovascular Biobank</th>
<th></th>
<th>Cleveland Clinic Genebank</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events by quintile</td>
<td>Top v Bottom Quintile</td>
<td>Events by quintile</td>
<td>Top v Bottom Quintile</td>
</tr>
<tr>
<td></td>
<td>Q1</td>
<td>Q2</td>
<td>Q3</td>
<td>Q4</td>
</tr>
<tr>
<td>All Subjects</td>
<td>N=2597</td>
<td></td>
<td></td>
<td>N=2702</td>
</tr>
<tr>
<td>Non Fatal MI</td>
<td>18</td>
<td>20</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td>Death/MI</td>
<td>82</td>
<td>74</td>
<td>73</td>
<td>63</td>
</tr>
<tr>
<td>&lt; 60 yrs</td>
<td>N=970</td>
<td></td>
<td></td>
<td>N=1108</td>
</tr>
<tr>
<td>Non fatal MI</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Death/MI</td>
<td>18</td>
<td>12</td>
<td>20</td>
<td>19</td>
</tr>
</tbody>
</table>

* All Subjects includes all those enrolled and prospectively followed irrespective of age at enrollment

** Adjusted for age at enroll, gender, BMI, hyperlipidaemia, hypertension, diabetes, ever smoked, baseline angiographic CAD ≥50%