Use of a genetic risk score to predict adverse events post myocardial infarction and age of first myocardial infarction

by

Christopher Labos MD CM, FRCPC

Master's of Science

Department of Epidemiology, Biostatistics and Occupational Health

McGill University

Montreal, Quebec, Canada

April 2014

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master's of Science

DEDICATION

To my family, for believing in me.

ACKNOWLEDGMENTS

I would like to begin by thanking Dr. George Thanassoulis for mentoring and supervising me during my master's degree. His support and advice have been invaluable as I tried to map out my future research career. Without our frequent conversations, I would surely be adrift and this master's thesis would not have been possible.

Also, I would like to thank Dr. Jay Brophy for taking me on as a student during this year. I am continually impressed by the breadth of his knowledge, his ethical standards, and his willingness to help and advise at any time. His frank and honest manner are an ideal I strive to emulate.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This thesis has been prepared as a manuscript-based thesis in accordance with the McGill University guidelines.

Chapter 1 presents an introduction and outline of the thesis. Chapter 2 presents a literature review of various risk scores in cardiology, the development of genetics in cardiology and the use of these new genetic markers in refining risk prediction. Chapter 3 describes the method logical and statistical testing used in the analysis of the thesis. Chapter 4 includes the first manuscript entitled "Utility of a Genetic Risk Score to predict recurrent cardiovascular events after Acute Coronary Syndrome". Chapter 5 provides an examination of the concept of index event bias and how it applies to our study. Chapter 6 includes the second manuscript entitled, "Use of genetic risk score to predict age of first myocardial infarction." Chapter 7 provides the global conclusion to the thesis. Formatting of all manuscripts was performed according to the requirements of the journals to which they have been submitted or accepted. Tables and figures for each manuscript are located at the end of each manuscript. References are located at the end of the thesis and reflect references may throughout the entire document.

STATEMENT OF ORIGINALITY

Both manuscripts contained in this thesis represent original scholarship and contribute to the advancement of their fields. The first manuscript is the first study to use a genetic risk score to predict adverse events post myocardial infarction. The second manuscript is also the first study to use a genetic risk score to predict age of first myocardial infarction.

ABSTRACT

Background: Risk prediction in cardiology is of critical importance because the risk stratification of patients into low, medium, and high risk categories is necessary for matching resources to patient needs.

Methods: Using data from two prospective cardiology study cohorts, we examined whether a genetic risk score (GRS) composed of 30 single nucleotide polymorphisms (SNP) was associated with recurrent events (either death, myocardial infarction (MI), or re-hospitalization) post MI and whether it could be used to predict the age of first acute coronary syndrome (ACS). Results were considered in each cohort individually and then pooled across both cohorts.

Results: The GRS was not associated with recurrent events, pooled hazard ratio (HR) 1.00 (95%CI 0.94-1.06). When the individual SNPs were evaluated none were consistently significantly associated with recurrent events in both cohorts. However, the GRS was associated with a younger age of first ACS. Each one standard deviation change (equivalent to 3.4 point change) in GRS was associated with 1.0 years (95%CI 0.1 - 2.0) decrease in the age of first ACS. Although other traditional risk factors seemed to have a stronger association with age of first ACS, such as smoking 8.1 years (95%CI 6.1 - 10.0), male sex 6.9 years (95%CI 4.1 - 9.7), and obesity 5.2 years (95%CI 2.6 – 7.9).

Conclusion: Thus, while a GRS may not be useful for predicting recurrent events post ACS, it does seem to be associated with the age of first ACS. However, traditional risk factors are also strongly associated with the age of first ACS. Therefore, attempts to understand the cause of premature ACS that occurs at a young age needs to account for traditional risk factor burden as well as genetic factors present in the individual under investigation.

ABRÉGÉ

Mise en conteste : En cardiologie, la prédiction du risque est importante parce que la stratification de patients en groupe de risque faible, moyen et élevé est nécessaire pour alloquer les ressources appropriées pour les besoins du patient.

Méthodes : En utilisant, les donnes de deux études prospectives en cardiologie, on a examiné si un score de risque génétique (SGR) basé sur 30 polymorphismes de nucléotide unique (PNU) est associé avec des événements récurrents (soit la mortalité, les infarctus du myocarde, ou la réhospitalization) suite a une infarctus et si il peut être utilise pour prédire l'âge d'un premier infarctus. Les résultats étaient considérés individuellement dans chaque études et après mis en commun.

Résultats : Le SGR n'est pas associé avec des évènement récurrents (HR 1.00, (95%CI 0.94-1.06). Quand les PNU étaient considérés individuellement, aucuns étaient associés avec des événements récurrents dans les deux banque de données. Cependant, le SGR est associé avec un âge du premier infarctus plus jeune. Chaque augmentation d'un écart-type du SGR est associé avec un diminution de l'âge d'un premier infarctus de 1.0 (95%CI 0.1 - 2.0) années. Mais d'autres facteurs de risque traditionnels étaient aussi associés avec l'âge d'une premier infarctus comme le tabagisme 8.1 années (95%CI 6.1 - 10.0), être male 6.9 années (95%CI 4.1 - 9.7), et l'obésité 5.2 years (95%CI 2.6 – 7.9).

Conclusions : Même si un SGR ne peut pas prédire les évènements récurrents, un SGR peut prédire l'âge d'un premier infarctus. Cependant les facteurs de risque traditionnels jouent un rôle important aussi. Donc l'investigation d'infarctus à un jeune âge nécessite une compréhension des facteurs de risque traditionnels en plus des facteurs génétiques.

Chapter 1 Introduction

Genetics in cardiology has been a rapidly expanding field ever since the first genome wide association studies (GWAS) in coronary artery disease were published in the past decade. However, while genetics has been very useful in specific conditions with a clear familiar inheritance pattern, such as hypertrophic cardiomyopathy or long QT syndrome, its full role in coronary artery disease (by the far the most common cardiac condition in the world) remains to be fully elucidated. Furthermore, how genetic factors interact with and compare to modifiable lifestyle risk factors also remains to be explored.

What is desperately needed and what has been sought for many years is a method to predict coronary disease. While a number of different clinical risk prediction models already exist[1-3], the hope has remained that newly discovered genetic factors would help refine that risk and predict not simply if an individual would develop an acute cardiac event but also when they might develop that event.

In this thesis, we undertake to explore this potential utility of newly discovered genetic markers. While the focus of research in cardiovascular genetics has been in identifying predictors of incident or prevalent cardiovascular disease, little work has been done in patients with established cardiovascular disease or those with early onset cardiovascular disease. Genetically-predisposed individuals for acute coronary syndromes (ACS) may, in fact, be at high risk for early recurrent events given that current medical treatment may be ineffective in reducing the genetic risk post-ACS which led to the initial CV event. Furthermore, if a GRS can further improve the prediction of risk this may have therapeutic implications in the management of ACS. Furthermore, individuals with an

early onset of cardiovascular disease would also likely represent a higher risk group than the general population and would be more likely to have a significant genetic contribution to their disease.

After this introductory chapter, the thesis continues with a literature review of the relevant subject matter and the prior work done on this subject in Chapter 2. The review encompasses the literature in prediction of incident and recurrent cardiovascular events followed by the recent work in cardiovascular genetics and the use of genetic markers for risk prediction.

We then present two original manuscripts that reflect our recent work on the subject. Our scholarship looks at whether a genetic risk score comprised of 30 single nucleotide polymorphisms offers any predictive value cardiovascular risk prediction. The overall methodological plan used throughout this thesis is presented in Chapter 3 with more specific details included within each manuscript. In the first manuscript, presented in Chapter 4, we investigate the role of a genetic risk score in predicting recurrent cardiovascular events in a population of patients presenting with an acute coronary syndrome.

Chapter 5 expands upon a potential source of bias encountered in the analysis of data from the first manuscript, and details how index event bias (sometimes called recurrent event bias) might be affecting the results.

Finally, in Chapter 6, we present our second manuscript that investigates whether our genetic risk score can predict the timing of a first cardiovascular event in terms of predicting the age of a first

acute coronary syndrome. In this manuscript we also examine how a genetic risk score compares to other traditional cardiac risk factors in determining the age of a first cardiac event.

Therefore we intend to explore whether a genetic risk score composed of single nucleotide polymorphisms (SNPs) identified through GWAS studies and associated with incident and prevalent cardiovascular disease can be used in the prediction of recurrent coronary events and in the prediction of the age of first coronary event.

Chapter 2: Literature Review

2.1 Prediction of coronary heart disease

Coronary heart disease (CHD) is extremely prevalent in the general population. In 2012 it accounted for 17.3 million deaths worldwide.[4] The lifetime risk of developing CHD was first estimated in the Framingham Heart Study. Amongst 7733 patients recruited between 1971 and 1975, the lifetime risk of CHD for those free of the disease at age 40 was 49% in men and 32% in women.[5] A more recent analysis of 18 different cohorts demonstrated similar results. [6]

The traditional cardiac risk factors (age, gender, hypertension, diabetes, cholesterol, and smoking) were first recognized in the original Framingham Heart Study cohort. [7,8] Multiple reports since then have confirmed the association between these risk factors and the presence of CHD. [9-11] Using this data, the Framingham Risk Score (FRS) was developed [12] and validated [13] to predict cardiovascular risk in the general population. Despite refinements to the FRS in 2002 by the ATP III panel and again in 2008[3], a major criticism of the FRS was that it did not incorporate some of the newer non-traditional risk factors. Traditional risk factors explain much of the risk for the CHD[14]. However, these risk factors do not account for the totality of cardiovascular risk and 10-15% of patients lack any of the 4 traditional modifiable risk factors (cigarette smoking, diabetes, hyperlipidemia, and hypertension).[11,14]

One of the most important omissions was a family history of CHD. A family history of CHD has been shown to increase the risk of CHD in several cohorts, such as the Physician's Health Study, Women's Health Study, Framingham Offspring Study, and INTERHEART study. [14-22] A significant family history was shown to double the predicted risk as calculated by the FRS although the re-classification of patients from intermediate to low or high risk categories was minimal.[22] Though other risk scores, such as SCORE[23], QRISK[24], and Reynold's risk score[2], have been developed and

include variables (such as ethnicity, socioeconomic status and family history) that were not included in the FRS, they tend not to enjoy widespread clinical use in North America.

2.2 Prediciton of recurrent events after myocardial infarction

Predicting cardiovascular risk in the post myocardial infarction (MI) setting is an important aspect of clinical care, as proper risk stratification is critical for deciding on optimal treatment strategies for patients with MI. [25] The Thrombolysis in Myocardial Infarction (TIMI) score was initially validated to predict adverse events (all-cause mortality, new or recurrent MI, or severe recurrent ischemia requiring revascularization) up to 14 days post non ST elevation MI.[26] The model included seven variables: age greater than 65 years, at least 3 risk factors for CHD, coronary stenosis greater than 50%, ST segment deviation on the electrocardiogram, at least two anginal episodes in the past 24 hours, elevated serum cardiac biomarkers, and the use of aspirin in prior seven days. It was also shown to be associated with more severe angiographic coronary artery disease [27] and was subsequently shown to retain its predictive ability up to 1 year post MI.[28,29] However, its main drawback lay in the fact that a separate and more complex algorithm needed to be used for ST elevation MI.[30]

The Global Registry for Acute Coronary Events (GRACE) registry devised a risk score that overcame this limitation and was validated in a global patient population from 14 countries, including Canada, with both ST elevation and non ST elevation MI.[31] Eight risk factors were determined to have prognostic importance: age, Killip class, systolic blood pressure, presence of ST segment deviation, cardiac arrest during presentation, serum creatinine concentration, presence of elevated serum cardiac biomarkers, and heart rate. The GRACE risk score was initially validated to predict in-hospital mortality, however was later shown to also predict recurrent events at 6 months

and 1 year.[29,32-34] Its versatility in both forms of acute coronary syndromes gives it an advantage compared to the TIMI risk score[29], however its greater complexity and more cumbersome risk normograms[35] make it more difficult to use clinically. Other risk scores, such as the CRUSADE risk score[36] and the CHADS risk score[37], have also been developed for risk prediction post MI but are not generally used clinically.

2.3 Genetic markers in CHD

Although a family history of CHD is considered in most of the risk scores used clinically for prediction of primary cardiac events, it is a crude metric of genetic susceptibility to cardiovascular disease for a number of reasons. First, considering only family history assigns equal risk to all family members regardless of which allele they may have inherited, while genetic testing holds the promise of an improved individualized risk profile by avoiding the misclassification associated with family history alone. Therefore, not surprisingly, the identification of specific genetic markers associated with cardiovascular risk held promise as a way to refine the risk prediction associated with heritable MI risk. Gene-environment studies have suggested that the heritability of MI ranges from 40-60%, which points to a strong role for genetic factors in estimating MI risk.[38-40] Secondly, what constitutes a family history of CHD has varied in different studies.[41] Current Canadian guidelines define it as an age of first MI before the age of 60 in either sex.[42] Finally, reports suggest that self reported family CHD has limited accuracy and reliability. An analysis of the Framingham Offspring study[43] demonstrated limited positive predictive value of 66% and 47% for fathers and mothers, respectively. Negative predictive value though was greater than 90%. Thus, there has been great interest in finding gene candidates that correlate with and could potentially replace a family history of CHD.[44]

A number of single nucleotide polymorphisms (SNP) have been identified that are associated with CHD.[45-52] Some SNPs are associated with an increased risk of developing CHD through their interaction with traditional risk factors such as obesity,[53-55] hypertension,[56] or cholesterol levels[49,51,57-60]. Others are genetic determinants of inflammatory biomarkers such as C-reactive protein (CRP) or interleukin-6[61,62] although the genetic link between CRP and MI risk has been disputed by Mendelian randomization studies.[63] Blood group O has been associated myocardial infarction (MI) risk since the 1960's.[64] The identification of the genetic locus 9q34.2 that codes for the pepetidoglycan responsible for ABO blood type, was shown in a CARDIoGRAM study to be associated with risk of MI. [65] Finally, some such as 9p21 appear to mediate their effects independently of any known risk factor. [66]

However, despite the number of SNPs that have been identified in recent years, only about 10% of CHD heritability can be explained through these genetic variants.[1] Some of this unexplained heritability could be accounted for by the interaction of genes with the environment. For example, the effect of 9p21 variants may be modified by dietary intake[67] and APOE genotypes by smoking.[68,69] Epigenetic mechanisms (i.e. changes in gene activity or expression not due to changes in the DNA sequence itself) such as DNA methylation, post-translational modification of histone proteins, and RNA based mechanisms also serve to explain some of this missing heritability.[70] Incomplete penetrance and linkage disequilibrium likely also play a role. Nevertheless, there likely remain a large number of common variants with small effect and rare variants of strong effect that will be described in the near future and help explain much of the missing heritability.[71]

2.4 Using SNPs and Genetic Risk Scores to refine cardiovascular risk prediction

Using newly identified SNPs to predict cardiovascular risk has had mixed benefits and genetic risk scores explain less than 5% of the variance in individual risk.[72] Firstly, it is important to consider that some polymorphisms though strongly associated with coronary artery disease (usually defined as >50% coronary obstruction determined by angiography) are not associated with myocardial infarction. The prime example would be one of the first SNPs to be defined, 9p21. A number of studies have shown that 9p21 was associated with coronary artery disease (CAD) and CAD severity but not with myocardial infarction.[73-76] Also, although many SNPs have been identified as being associated with risk of incident cardiovascular disease, not all have been replicated in independent studies.[77,78]

Single SNPs at 9p21.3 although strongly associated with CAD (relative risk of 1.27 95%CI(1.23-1.31) in one recent meta-analysis)[76], did not improve risk prediction.[79,80] However, both these studies had some limitations with the study of Paynter et al.[79] only including women and Talmud et al.[80] not including diabetes in their analytical model. A 9 SNP genetic risk score showed a 2-fold increased odds ratio for MI but did not examine the incremental value of this risk score to traditional risk factors.[49] Another 9 SNP model associated with LDL or HDL cholesterol also showed a 15% increased risk of CAD and while it showed an improvement in risk classification, it did not improve clinical risk prediction, as assessed by the C statistic. [60]

Attempts at incorporating more SNPs into the genetic risk score (GRS) have met with similar mixed success. A 101 SNP model using data from the Women's Genome Health study did not show any improvement in discrimination or reclassification of incident cardiovascular disease.[81] However, inherent in its design, the study can only be extrapolated to women. Similarly, Ripatti et al. using a 13 SNP model in a prospective cohort from Finland and Sweden and found that the genetic risk

score was associated with incident CHD but did not improve net reclassification.[82] However, a 12 SNP model from the Ottawa Heart Genomics study performed better than 9p21.3 alone whether traditional risk factors were considered or not. [83]

The addition of 9p21 to the traditional risk factors of the Framingham Risk Score did result in an improvement in re-classification in an Atherosclerosis Risk in Communitites (ARIC) cohort. [84] However, family history was not included in the risk factors. A 116 SNP model also from the ARIC cohort, showed improvement in re-classificaiton but only African Americans.[85] Other studies showed no clear benefit to the use of a genetic risk score.[86-89]

A more recent study from the Framingham group looked especially at two genetic risk scores.[90] The first was composed of a 13 SNP model using previously described variants associated with CHD and the seconded added 16 recently identified SNPs from genome wide association studies.[57,91] The 13 SNP GRS improved risk reclassification (net reclassification index (NRI) = 0.17 95%CI 0.01-0.33) for incident CHD. The NRI measures the proportion of subjects with events whose predicted risk moved up a risk category minus those whose predicted risk moved down a risk category with application of the algorithm plus the proportion of subjects without events whose predicted risk moved down a risk category minus those whose predicted risk moved up a risk category with application of the algorithm [NRI=P(up|event)−P(down|event)+P(down|nonevent)−P(up|nonevent)].[92]

The results remained consistent when parental history of cardiovascular disease was included in the model. Furthermore, the addition of the 16 new SNPs did not significantly modify the results.

2.5 Using SNPs and a GRS to Predict Recurrent Events post MI

While most studies have looked at prevalent or incident cardiovascular disease, there has been limited investigation into the prediction of recurrent cardiovascular events. In fact, only two studies have looked at the predictive abilities of various SNPs in terms of recurrent events. In the first study, Buysschaert et al.[93] examined the association between the rs1333049 variant on chromosome 9p21 and the recurrent events post MI in 3247 individuals from the GRACE registry. They found a statistically significant association [Hazard ratio (HR) 1.58, 95%CI 1.00-2.48] for the recurrent MI and cardiac death, but not for the outcome of recurrent MI alone [adjusted HR 1.47, CI = 0.99-2.18] in the first 6 months after MI although the study may have been underpowered for this endpoint. The improvement in risk prediction was assessed using the integrated discrimination index (IDI)[94], calculated as the probability difference in discrimination slopes (i.e. the probability of an event minus the probability of a non-event). The addition of rs1333049 to the GRACE risk variables resulted an improved IDI for the combined outcome of recurrent MI or cardiac death (p=0.040) but again not for recurrent MI alone (p=0.073). Overall, an additional 5.93% of patients had an increased probability of being correctly classified when rs1333049 was included in the model.
In another study, also using the GRACE registry, Wauters et al.[95] examined 23 SNPs in terms of recurrent events post MI. Only one of the 23 SNPS was found to be significantly associated with their outcome. The SNP, rs579459, located upstream of the ABO gene is associated with blood type A. After multivariate adjustment the SNP was associated with recurrent MI and cardiac death (HR 1.80, 95%CI 1.09–2.95) and recurrent MI alone (HR 2.25, 95%CI 1.37–3.71). Improvement in risk prediction was marginal. There was no improvement in the C-statistic. The relative IDI showed an improvement of 6.0% but only for the combined outcome of recurrent MI cardiac death, and not for recurrent MI alone. Also the improvement in IDI was significant only when rs579459 was added to

the multivariate model of the study, not to the GRACE risk score. Furthermore, it was only significant once the discovery and validation cohorts were pooled. Importantly, neither of the above two studies looked at a combined genetic risk score and looked only at individual SNPs.

2.6 Limitations of using Genetic Risk Scores

Why current genetic risk scores have failed to provide better risk discrimination and risk re-classification may be due to several factors. As mentioned above, traditional risk factors and family history already account for a sizable portion of cardiovascular risk and any genetic risk score would have to provide incremental benefit beyond what these risk factors already provide.[96] Furthermore, genetic risk scores validated in one population may not be applicable to other ethnically distinct populations and meta-analyses combining different studies in heterogeneous populations may not accurately reflect the predictive value of a risk score.[72,97] Since the risk with any isolated SNP is small, current risk scores may not include enough SNPs to achieve a significant effect in terms of reclassification. Theoretical estimates suggest that roughly 100 uncorrelated genetic variants with relative risk of 1.5 would be needed to achieve a theoretically ideal degree of risk reclassification, as defined as an area under the curve (AUC) of 0.90.[98] However, relative risks of 1.5 may not be observable since most identified genetic variants showed relative risks closer to 1.1-1.25. Taking this estimate to be the observed relative risk, and conservatively assuming the risk allele frequency at 5%, would require >400 genetic variants would be needed.[99] Whether such a large number of variants will ever be identified, or even exists, is obviously unknown.

Chapter 3: Methodological plan of the thesis

3.1 Patient cohorts used in the analysis

The statistical analyses performed in this thesis were performed on data gathered from two large study cohorts: RISCA[100] and GENESIS-PRAXY[101]. The RISCA cohort consisted of 1210 patients recruited from four tertiary and four Canadian community hospitals (seven in Quebec and one in New Brunswick) who presented with an acute coronary syndrome ACS. The GENESIS-PRAXY study was a prospective multicenter study of patients aged 18-55 recruited from 23 from centers across Canada, as well as one center in the United States and one in Switzerland, admitted to hospital with ACS. Specific details about patient demographics and inclusion criteria for each study are presented with the methods section of both manuscripts.

3.2 Genetic testing and development of the genetic risk score

Genetic testing was performed in the laboratory of Dr. James Engert who is a co-author on both manuscripts. DNA extraction from buffy coats as well as quantification and plating of samples was performed using standard techniques. Extracted DNA was stored at -80°C. Genotyping for 28 SNPs was performed using iPLEX technology on a MassARRAY Compact Analyser (Sequenom Inc., San Diego, CA, USA). Two SNPs, rs4977574 and rs1412444, were genotyped using a Custom TaqMan SNP Genotyping Assay (Applied Biosystems, Foster City, CA, USA). For both the Sequenom and Taqman data, all SNP call rates were >97.5%. All samples had >80% successful genotypes and all SNPs were in Hardy-Weinberg equilibrium (p>0.002).

Two genetic risk scores (GRS) were constructed and used in the analysis. The first GRS was determined *a priori* using genotypes from 30 uncorrelated SNPs ($R^2 < 0.3$) in Hardy-Weinberg

equilibrium ($p>0.002$) that were robustly associated and replicated in published genomewide association studies (GWAS) of myocardial infarction or coronary artery disease. We also constructed a second 8 SNP GRS limited only to SNPs known to be associated with LDL-C in prior genome wide association studies of myocardial infarction/coronary artery disease.

As performed in prior work, a score for each individual was calculated as the unweighted sum of each risk allele across all 30 SNPs (i.e score of 2 for those homozygous for the risk allele, a score of 1 for heterozygotes, and a score of zero for the absence of the risk allele). Missing genotypes (<0.35% of all genotypes) were assumed to be missing at random (i.e. non-informative missingness) and were imputed as two times the risk allele frequency, using the risk allele frequencies from the entire data set. Thus, every individual could have a genetic risk score ranging from 0 to 60 for the 30 SNP MI GRS or 0 to 16 for 8 SNP LDL GRS. The list of SNPs used in the analysis is included in Appendix A.

3.3 Statistical analysis of covariates and model development

Firstly, all potential variables of interest that were available in the datasets were examined individually, their normality was assessed using standard histograms, and then the number of missing values was determined. Covariates that had significant (>10%) missing data were not included in the analysis. Other variables were potentially included as part a complete case analysis based on the assumption that data was missing completely at random[102-104]. A sensitivity analysis was later conducted to assess if imputing missing values in the included covariates would materially change the results.

A time to event analysis was used to minimize the effect of censoring, although one year follow-up was complete for all subjects. The primary outcome was a recurrent event post ACS which was defined as a composite outcome of all-cause mortality, recurrent ACS, and re-hospitalizations. A Cox proportional hazard model was used to assess the association between a GRS and recurrent events.[105] The proportional hazards assumption was verified to be valid with the STATA command –estat phtest– before proceeding (details provided in Appendix B). Other covariates of interest were also investigated in a univariate Cox model. Covariates were included in the final multivariate model either because they were significantly associated with recurrent events in univariate analysis or because prior research has shown them to be important predictors of recurrent events, such as age, gender, and other traditional risk factors. Covariates were added one at a time to assess their impact on the hazard ratio of the GRS and determine which covariates would significantly change the results. This task was then repeated using the –epiconf- command in STATA.[106] Each individual SNP was then also investigated in a univariate Cox model and then adjusted for the same multivariate model.

The predictive value of the GRS was also compared to the GRACE[31] risk score. Although the TIMI risk score[26] is widely used in cardiology, there are two different scoring algorithms for ST elevation myocardial infarctions (STEMI) and non-ST elevation myocardial infarctions (NSTEMI). The GRACE score offers the advantage of a single scoring algorithm for all acute coronary syndromes, both STEMI and NSTEMI. For the GRACE risk score, the percent risk was calculated using the normogram for 6 month outcomes of death or MI.[35]

3.4 Assessing the predictive value of the GRS

The goodness of fit of a model was assessed using the likelihood ratio test by comparing the model with the GRS and the GRACE risk score to the model with the GRACE risk score alone.[107] The

additive predictive value of the GRS to the GRACE risk score was evaluated using Harrel's C statistic,[108] the integrated discrimination improvement (IDI)[109], and the continous net re-classification improvement (NRI)[92] that does not require pre-specified cutoffs like the standard net re-classification improvement.[94] The results of this testing was not reported in the final version of the manuscript submitted for publication and thus appear in Appendix D on this thesis.

3.5 Pooling of results and meta-analytic methods

All testing was performed independently in the RISCA and GENESIS-PRAXY cohort. The results of the multivariate Cox modelling for the GRS and each SNP were then pooled across the results of the two study cohorts. Pooling of the log hazard ratios was done with an inverse variance weighting method and then exponentiated to obtain the hazard ratio.

3.6 Linear regression of age of first acute coronary syndrome (ACS)

For the association between the GRS and the age of first ACS presented in manuscript #2 in Chapter 6, the GRS was regressed against the age of the participants at the time of presentation for their first MI. Patients with any prior history of cardiovascular disease were excluded from the analysis. The appropriateness of the linear model was assessed by 2 methods. Firstly, by plotting the model residuals against the covariates included in the model to demonstrate that variance was constant and not dependent on the covariates. Secondly, by assessing the normality of the distribution of residuals via a histogram. These graphics were not included in the final manuscript and are presented in Appendix C. Selection of covariates for the final multivariate model was performed as described above in Section 3.3.

3.7 Consideration for multiple hypothesis testing

Due to the multiplicity of statistical tests, the Bonferroni correction was applied to all p-values in these analyses such that we only considered SNPs to be statistically significant when they met a p-value threshold of < 0.00167 (0.05/30). If any result was nominally significant at p-value threshold of <0.05 but >0.00167, an a priori decision was made to attempt to replicate that result in an external dataset prior to manuscript submission. Although the Bonferroni has been criticized for being too conservative,[110] since it inflates Type II error at the expense of reducing Type I error, it remains commonplace in the field of genetics and was utilized in this thesis.

Chapter 4: Manuscript #1

4.1 Preface to Manuscript #1

Using genetic information to predict cardiovascular has had mixed results with respect to predicting the risk of a myocardial infarction. However, predicting recurrent events in patients presenting to hospital with a myocardial infarction is potentially possible. Two studies from the GRACE group have looked at individual SNPs and the only study to look at a genetic risk score was very underpowered to detect recurrent events.

The objective of our study was to evaluate whether a genetic risk score composed of 30 SNPs, known to be associated with MI through genome wide association studies, was associated with recurrent events. The GRS was investigated in both the RISCA study cohort that was composed of individuals of all ages admitted to hospital with an acute coronary syndrome and the GENESIS-PRAXY cohort that was composed only of individuals admitted to hospital with an acute coronary syndrome under the age of 55. Two separate cohorts were used firstly to replicate any nominally statistically significant results and secondly to investigate the use of a GRS in both a general population cohort and also in a cohort that was limited to those of younger age.

The results of this study have been presented at both the Canadian Cardiovascular Congress in October 2013 and at the American Heart Association annual meeting in November 2013.

4.2 Contribution of Authors

Dr. Labos, Dr. Engert and Dr. Thanassoulis conceived of the study design and the analysis plan. Dr. Brophy and Dr. Bogaty created the RISCA study cohort. Dr. Pilote created the GENESIS-PRAXY study cohort. Genotyping of individuals was performed by Leo Wang and Dr. Engert. All statistical analyses were performed by Dr. Labos. The first draft of the manuscript was written by Dr. Labos. All authors read and approved the final version of the manuscript.

4.3 Utility of a Genetic Risk Score to predict recurrent cardiovascular events after Acute Coronary Syndrome

Christopher Labos MD CM*, Rui Hao Leo Wang BSc†, Louise Pilote MD MPH PhD†, Peter Bogaty MD‡, James M. Brophy MD PhD*†, James C. Engert PhD§, George Thanassoulis MD MSc†

* Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, QC, Canada

† Departments of Medicine, McGill University, Montreal, QC, Canada

‡ Institut Universitaire de Cardiologie et de Pneumologie, Laval University, Quebec City, QC, Canada

§ Department of Human Genetics, McGill University, Montreal, QC, Canada

Address for correspondence:
George Thanassoulis MD MSc FRCPC
Director, Preventive and Genomic Cardiology
McGill University Health Center
687 Pine Ave W. H4.55
Montreal, QC
H3A 1A1
Telephone: 514-934-1934 x 35465
E-mail: george.thanassoulis@mcgill.ca

Abstract

Introduction: Little evidence exists regarding the utility of genetic risk scores (GRS) in predicting recurrent cardiovascular events after acute coronary syndrome (ACS). We sought to determine whether a GRS would predict recurrent cardiovascular events in individuals after an ACS.

Methods and Results: Participants admitted with acute coronary syndromes from the Recurrence et Inflammation dans Syndromes Coronarien Aigus cohort (RISCA; n=1054, mean age 62 ± 11 years, 24 % female) and the GENESIS-PRAXY cohort (n = 691, mean age 48±6 years, 28% female) underwent genotyping for 30 single nucleotide polymorphisms (SNPs) associated with myocardial infarction (MI) in genome wide association studies. A 30 SNP MI GRS and in secondary analyses a 8 SNP LDL GRS were constructed. The primary endpoint was defined as all-cause mortality or recurrent ACS or re-hospitalization within 1 year of admission. Results across both cohorts were pooled using a random-effects model. In unadjusted and fully adjusted analyses, a 30 SNP GRS was not significantly associated with recurrent events (HR per allele 1.00 95%CI 0.94-1.06). The 8 SNP LDL GRS was also not associated with recurrent events (HR 1.03 95%CI 0.94-1.14). Addition of these GRS to GRACE risk model did not significantly improve risk prediction. In sensitivity analyses, the 8 SNP LDL GRS was significant only in individuals of European descent (HR 1.18 95%CI 1.03-1.36).

Conclusion: Neither the GRS of MI SNPs or LDL SNPs was associated with recurrent events 1-year post ACS in the population overall. However the 8-SNP LDL GRS was associated with recurrent events in individuals of European ancestry.

Introduction

Despite optimal medical therapy, early recurrent cardiovascular (CV) events within the first year of a myocardial infarction (MI) remain common and are associated with significant morbidity and mortality[111]. A family history of premature MI is a risk factor for recurrent CV events which suggests that genetic factors may play a role[112]. Recent large-scale genetic studies have identified several common genetic variants robustly associated with MI[50,113-120]; however it remains unknown whether such variants predispose to early recurrent CV events. Although the exact biological mechanisms underlying these genetic associations remain unknown, these variants appear to act via non-traditional pathways of atherothrombosis which may not be affected by contemporary secondary prevention therapy. Genetically-predisposed individuals for acute coronary syndromes (ACS) may, in fact, be at high risk for early recurrent events given that current medical treatment may be ineffective in reducing the genetic risk post-ACS which led to the initial CV event. Furthermore, if a GRS can further improve the prediction of risk this may have therapeutic implications in the management of ACS.

While a number of studies have looked at the use of single SNPs or a genetic risk score (GRS) to predict prevalent or incident cardiovascular disease [60,79-81,84-89], few have looked at recurrent events post ACS. The GRACE registry evaluated the predictive value of single risk alleles for MI [93,95] with recurrent events post-ACS, but did not report results for a GRS. Therefore the utility of GRS in post-ACS risk stratification remains unknown. Accordingly, we sought to determine whether a GRS composed of 30 SNPs associated with MI and 8 SNPs associated with LDL could predict recurrent events in patients presenting in two hospital-based ACS cohorts.

Methods

Participants

Participants from the Recurrence and Inflammation in the Acute Coronary Syndromes (RISCA) cohort and the Gender and Sex determinants of cardiovascular disease: From bench to beyond-Premature Acute Coronary Syndrome (GENESIS-PRAXY) cohorts contributed data to this analysis.

The RISCA cohort[100] consists of 1210 consecutive patients recruited from four tertiary and four Canadian community hospitals (seven in Quebec and one in New Brunswick). To be eligible, patients had to have an urgent admission to hospital with a diagnosis of either acute MI or unstable angina. All basic demographic and clinical data including traditional risk factors, results of all diagnostic and therapeutic procedures performed in hospital (including biochemical tests) and all medications taken prior to admission and prescribed at discharge were recorded in a comprehensive case report form and stored in electronic format on a secure server. This information was independently verified for consistency and then systematically assessed by on-site visits. Of the 1210 patients enrolled, 1054 provided consent for genetic testing and 14 patients were excluded because of missing covariates for a final sample of 1040 for analysis.

The GENESIS-PRAXY[108] study was a prospective multicenter study of patients aged 18-55 recruited from 23 from centers across Canada, as well as one center in the United States and one in Switzerland, admitted to hospital with ACS.[108] Data on patients was collected with the use of a self administered questionnaire supplemented by a medical chart review performed by a research nurse. Follow-up occurred at 1, 6 and 12 months via telephone interview and repeat questionnaire. There were 1123 individuals with follow-up data available in the GENSIS-PRAXY cohort. Genotyping

was performed on 705 individuals who were of French Canadian or European ancestry. Fourteen patients were excluded because of missing covariates, leaving a final sample 691 participants.

Outcome and Covariate Definitions

The primary outcome was a composite of all-cause mortality, recurrent ACS, and re-hospitalizations. In the RISCA study, all events were verified on site with the use of supporting documentation. All prospective and potential outcomes were centrally adjudicated independently by 2 cardiologists. Recurrent ACS included both myocardial infarction and unstable angina. Myocardial infarction was defined as a history of characteristic chest discomfort or pain with an elevation of creatinine kinase – myocardial band to greater than 1.5 times the upper limit of normal. The definition was subsequently amended to also include a history of characteristic chest pain and elevation in cardiac troponin levels above the upper limit of normal. A diagnosis of unstable angina required either one episode of characteristic discomfort or pain at rest or with minimal exertion lasting more than 10 minutes or two episodes lasting more than 5 minutes with negative cardiac biomarkers. To increase specificity, UA patients had to have electrocardiogram changes consisting of 0.5 mm ST-segment depression or transient ST-segment elevation or 2 mm T-wave inversion in 2 contiguous leads.

In GENSIS-PRAXY an acute coronary syndrome was defined as symptoms of chest discomfort with either electrocardiographic changes suggestive of ischemia (such as ST elevation or depression $\geq$ 1 mm , new T wave inversions of $\geq$ 1 mm, pseudo-normalization of previously inverted T waves, new Q waves, new R> S wave in V1, or new left bundle branch block) or an increase in cardiac enzymes which was defined as: creatine kinase-MB (CK-MB) > 2 times the upper limit of

the hospital's normal range or if no CK-MB available, then total creatine phosphokinase > 2 times the upper limit of the hospital's normal range, positive troponin I, positive troponin T.

Covariate definitions were standard across both cohorts for analysis. In each cohort, patients were defined as having a prior history of cardiovascular disease if they had a history of myocardial infarction prior to the index hospitalization, prior angioplasty, prior coronary artery bypass grafting, prior angina, prior congestive heart failure, prior stroke, or any admission for a cardiac related condition. Patients were defined as having diabetes if they reported a history of diabetes, but not simply glucose intolerance, whether treated with medications or by diet. Similarly, patients were defined as having hypertension or hypercholesterolemia if they had a history of hypertension or hypercholesterolemia documented in their medical record, whether treated or untreated. Current smokers were defined as patients who continued to smoke (>1 cigarette per day) at the time of enrolment or who had quit within the past 30 days. Body mass index (BMI) was calculated using height and weight as measured at time of admission. Medication classes were determined by the medications prescribed at the time of discharge.

Genetic testing and development of the genetic risk score

Blood samples for all patients were taken within 24 hours of admission. DNA extraction from buffy coats as well as quantification and plating of samples was performed using standard techniques. Extracted DNA was stored at -80°C. Genotyping for 28 SNPs was performed using iPLEX technology on a MassARRAY Compact Analyser (Sequenom Inc., San Diego, CA, USA). Two SNPs, rs4977574 and rs1412444, were genotyped using a Custom TaqMan SNP Genotyping

Assay (Applied Biosystems, Foster City, CA, USA). For both the Sequenom and Taqman data, all SNP call rates were >97.5%. All samples had >80% successful genotypes and all SNPs were in Hardy-Weinberg equilibrium (p>0.002).

The genetic risk score (GRS) was constructed using genotypes from 30 SNPs robustly associated and replicated in published genomewide association studies (GWAS) of myocardial infarction or coronary artery disease.[45-52,57,109] (Supplementary Table 1) As performed in prior work[90], a score for each individual was calculated as the unweighted sum of each risk allele across all 30 SNPs (i.e score of 2 for those homozygous for the risk allele, a score of 1 for heterozygotes, and a score of zero for the absence of the risk allele). Missing genotypes (<0.35% of all genotypes) were assumed to be missing at random (i.e. non-informative missingness) and were imputed as two times the risk allele frequency, using the risk allele frequencies from the entire data set. Thus, every individual could have a genetic risk score ranging from 0 to 60; the actual range observed was 18 to 40. We also constructed a second 8 SNP GRS limited only to SNPs known to be associated with LDL levels. This 8 SNP GRS was constructed in the same manner described above with the following 8 SNPs: rs11206510, rs1122608, rs1412444, rs2259816, rs3798220, rs579459, rs646776, and rs964184.

Statistical analyses

Continuous variables were reported as means with standard deviations. Categorical variables were reported as counts with proportions. The association between each GRS and recurrent event was assessed using Cox proportional hazards models. Several Cox regression models with GRS were constructed including (1) univariate model with only the GRS; (2) adjustment for age and sex; and (3) a multivariate model adjusting for age, sex, prior cardiovascular disease, hypertension,

diabetes, hyperlipidemia, body mass index (BMI), smoking status, use of aspirin, use of Clopidogrel or other thienopyridne, use of beta-blocker, use of statin, and use of angiotensin converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB). Log hazad ratios for each GRS were then pooled across both study samples using a random effects DerSimonian & Laird model to account for the heterogeneity between the two study cohorts and the differences in the effect measure being estimated by each cohort.

The predictive value of the GRS was also compared to the GRACE risk score. For the GRACE risk score, the percent risk was calculated using the normogram for 6 month outcomes of death or MI.[35] The goodness of fit of the models including the risk score plus the GRS was evaluated using the likelihood ratio test. The same statistical modelling was repeated for the 8-SNP LDL GRS.

Sensitvity analyses were conducted for individuals presenting to hospital with their first ACS and, with respect to the RISCA cohort, for individuals less than 55 years old which was the maximum age of the GENESIS-PRAXY cohort. All statistical testing was performed using STATA version 12 (StataCorp, College Station, Texas).

Results

Baseline characteristics

There were 1040 individuals from RISCA available for analysis (mean age 61.8+/- 11.4, 24.4% female). Over half (55.5%) had a prior history of cardiovascular disease and 28.5% had a prior revascularization with either CABG or PCI. The vast majority (87.4%) had at least one traditional cardiac risk factor: hypertension, diabetes, cholesterol, or smoking. The mean GRS for the entire population was 31.5 (standard deviation (sd) = 3.4). When divided into tertiles based on

GRS, the baseline characteristics of the study population were not meaningfully different. Full details of the baseline characteristics of the study population can be found in Table 4.1. In the RISCA cohort, there were 72 occurrences of the primary composite endpoint: death, recurrent MI, or hospitalization for a cardiac cause. One year follow-up was complete for all subjects.

There were 691 individuals from GENESIS-PRAXY available for analysis (mean age 48.3 +/- 5.6, 27.8% female). The mean GRS for this sample was 30.0 +/- 3.5. Most (85.1%) had at least one traditional cardiac risk factor while 39.8% had a prior history of cardiovascular disease and 14.6% had a prior history of revascularization. Most of the individuals were of European origin (74.3%). When divided into a tertiles based on GRS, the baseline characteristics of the study population were not meaningfully different. Full details of the baseline characteristics of the study population can be found in Table 4.1. In the GENESIS-PRAXY cohort, there were 93 occurrences of the primary composite endpoint. One year follow-up was complete for all subjects.

Associations between GRACE score and recurrent events in RISCA and GENESIS-PRAXY

The GRACE risk score was significantly associated with recurrent events, HR 1.07 per unit change (95%CI 1.05-1.09 $p<0.001$) in the pooled analysis across both datasets.

Associations between clinical covariates, 30 MI SNP GRS and 8 SNP LDL GRS with recurrent events in RISCA

The mean GRS did not differ significantly between those who did and did not experience the primary endpoint (31.5 vs. 31.05, two sample t-test p-value =0.27). There was no significant association between the GRS and recurrent events Hazard ratio (HR) 0.96 [95% confidence interval

(CI) 0.90-1.02]. Adjustment for age and sex alone, adjustment with the multivariate model or adjustment for GRACE risk score did not materially alter the results HR 0.97 (95%CI 0.91-1.03), HR 0.97 (95%CI 0.90-1.03), and HR 0.96 (0.89-1.03) respectively. (Table 4.2 and Table 4.3) An analysis of the GRS divided into tertiles demonstrated no difference in the survivor function between those with a low, intermediate or high GRS. (Figure 4.1)

After multivariate adjustment, three covariates were significantly associated with recurrent events. A prior history of cardiovascular disease HR 1.95 (95%CI 1.06-3.61, p=0.032), a history of diabetes HR 2.10 (95%CI 1.23-3.58, p=0.006), and statin use at discharge HR 0.47 (95%CI 0.27-0.84, p=0.010). (Table 4.2)

Sensitivity analyses looking at individuals less than 55 years of age or individuals presenting with their first ACS as the index event did not materially change the results, HR 0.86 (95%CI 0.71-1.04) and HR 0.96 (95%CI 0.88-1.05) respectively. When the GRS was added to the GRACE risk model the likelihood ratio tests was not significant, p values 0.23.

With respect to the 8-SNP LDL GRS, there was no association between the GRS and recurrent events with either the unadjusted, HR 0.88 (95%CI 0.77-1.01) p=0.06, or fully adjusted model, HR 0.89 (95%CI 0.77-1.04) p=0.13. (Table 4.4) When added to the GRACE risk model, the likelihood ratio test was not significant, p-value 0.16.

Associations between clinical covariates, 30 SNP GRS and 8 SNP LDL GRS with recurrent events in GENESIS-PRAXY

The mean GRS did not differ significantly between those who did and did not experience the primary endpoint (30.4 vs. 29.9, two sample t-test p-value =0.24). The unadjusted Cox model demonstrated a HR of 1.04 (95%CI 0.98-1.10). Adjustment for age and sex alone, adjustment with the multivariate model, and adjustment for GRACE risk score did not materially alter the results HR 1.04 (95%CI 0.98-1.10), HR of 1.03 (95%CI 0.97-1.09), and HR 1.04 (95%CI 0.98-1.10) respectively. (Table 2 and Table 3) Once again, an analysis of the GRS divided into tertiles demonstrated no difference in the survivor function between those with a low, intermediate or high GRS.

After multivariate adjustment, no covariates were significantly associated with recurrent events. Sensitivity analysis looking at individuals presenting with their first ACS did not change the results, HR 1.03 (95%CI 0.96-1.10). When the GRS was added to both risk TIMI or GRACE models the likelihood ratio tests were not significant, p values of 0.26 and 0.23 for TIMI and GRACE respectively.

With respect to the 8-SNP LDL GRS, there was an association between the GRS and recurrent events, with an unadjusted HR 1.15 (95%CI 1.02-1.30) p=0.02. After adjustment for the multivariable model the HR was unchanged 1.14 (95%CI 1.01-1.29) p=0.04. (Table 4.4) When added to the GRACE risk model, the likelihood ratio test showed an improvement in the goodness of fit with the addition of the GRS, p-value 0.03. The association was significant amongst individuals of European ancestry (HR 1.18 95%CI 1.03-1.36) but not amongst those of French Canadian ancestry HR (1.08 95%CI 0.81-1.45). However, this difference was not statistically significant when tested for interaction p=0.47.

Pooled association results for the 30 SNP MI GRS and 8 SNP LDL GRS with recurrent events

Pooling the results from the RISCA and GENESIS-PRAXY database using the DerSimonian & Laird random effects model did not significantly change the results for the 30-SNP MI GRS, HR 1.00 (95%CI 0.94-1.06). (Table 4.3) When the analysis was limited to individuals under the age of 55, the pooled effect measure was 0.96 (95%CI 0.81-1.14). The 8-SNP LDL GRS was also not significant when pooled across both cohorts, HR 1.03 (95%CI 0.94-1.14) (Table 4.4).

Discussion

In our analysis of the RISCA and GENESIS-PRAXY study population, a GRS composed of 30 SNPs was not predictive of adverse events. This is in keeping with prior studies that have failed to show that a GRS increases the predictive capability of a risk model. A 101 SNP model using data from the Women's Genome Health study did not show any improvement in discrimination or reclassification of incident cardiovascular disease.[81] A 13 SNP model in a prospective cohort from Finland and Sweden found that the genetic risk score was associated with incident CHD but did not improve net reclassification. [82] A more recent study from the Framingham group looked especially at two genetic risk scores, a 13 SNP GRS and a more extensive model using an additional 16 recently identified SNPs. [90] The 13 SNP GRS improved risk reclassification (the ability of a risk model to reclassify individuals from one risk stratum into another risk stratum) but did not improve discrimination (the ability of a risk model to rank order individuals' risks) for incident CHD. The addition of the 16 new SNPs did not modify the results.

However, while most of the prior studies looked at either prevalent or incident cardiovascular disease, only two studies from the GRACE group looked at recurrent events after ACS. In one study by Buysschaert et al., single SNP at the 9p21 gene was associated with the composite outcome of recurrent MI and mortality at 6 months post ACS and improved classification when added to the GRACE score.[93] However, it was not associated with the primary outcome of recurrent MI after multivariable adjustment and three other SNPs at 9p21 locus also showed no association. Another study by Wauters et al. of 23 SNPs found that only one SNP, rs579459, upstream from the ABO gene, was associated with recurrent MI. It also improved risk prediction according to the integrated discrimination statistic, but not according to the C-statistic.[95] Interestingly, Wauters et al. used another 9p21 SNP that differed from the one used by Bussychaert et al. and showed no association with recurrent MI. Though the 23 SNPs were not considered in aggregate as a GRS, the majority of the SNPs analyzed were not predictive of recurrent events. Thus it is unlikely that a GRS comprised of these 23 SNPs would have improved risk prediction.

The only other study to use a GRS to examine recurrent events was the study by Tagrante et al. [92] It showed an association between GRS and myocardial infarction (MI) but not cardiovascular events overall. Also, although the GRS was associated with MI in the unadjusted model and the model adjusted for age and sex, it was not statistically significant in the multivariate model HR 1.13 95%(CI 1.00-1.28) p=0.071. Finally, the analysis was made comparing the top and bottom quartiles of the GRS rather than looking at the entire distribution and was based solely on 30 MIs in the upper quartile. By contrast, our study looked at the entire distribution of the GRS and was based on 165 events (72 events in RISCA and 93 events in GENESIS-PRAXY). This analysis included a much

larger number of recurrent events and would have been more likely to identify a significant association if it had existed.

This study had a number of limitations. Firstly, it included a heterogeneous patient population with a broad range for the age of presentation. Conceivably, younger individuals are more likely to have a genetic contribution to their risk of recurrent cardiovascular events than older individuals. However, no statistically significant association could be found in the younger GENESIS-PRAXY population nor in sensitivity analyses that limited RISCA to those less than 55 years old. Furthermore, many individuals had a long standing history of CVD with multiple prior interventions and MIs. These individuals are likely to have a different set of predictors determining their risk of recurrent events than an individual presenting with their first ACS. However, once again sensitivity analyses limited to individuals presenting with first ACS did not yield significantly different results.

Finally, the 8 SNP LDL GRS was associated with recurrent events in the GENESIS-PRAXY but not in the RISCA. Differences in ethnic origins may partially explain this result. While the RISCA population was composed almost exclusively of French Canadians, GENESIS-PRAXY contained a predominantly European population. Sensitivity analyses demonstrated the positive result in the GENESIS-PRAXY database was driven predominantly by Europeans. However, since the RISCA dataset did not contain individuals of non-French Canadian origin it was not possible to verify and replicate this result. Thus it is possible, that the 8 SNP LDL GRS may offer improvements in risk prediction only in selected populations but requires validation in an external cohort.

Conclusion

A GRS composed of 30 SNPs was not statistically significantly associated with the primary composite outcome of all-cause mortality, recurrent cardiovascular events, and hospitalizations. The 8-SNP LDL GRS was associated with recurrent events in the GENESIS-PRAXY database but not in RISCA, and not when the results were pooled across both datasets. Sensitivity analyses suggest that the 8-SNP LDL GRS may be associated with recurrent events only in individuals of European ancestry.

Table 4.1: Baseline characteristics of patients by tertile of GRS in RISCA

	RISCA								GENESIS-PRAXY							
	Bottom tertile N=408		**Middle tertile N=343**		**Top tertile N=289**		**All patients N=1040**		**Bottom tertile N=238**		**Middle tertile N=233**		**Top tertile N=220**		**Total N=691**	
Age (yrs), mean(sd)	62.5	(11.6)	61.3	(11.1)	61.3	(11.4)	61.8	(11.4)	48.7	(5.2)	48.5	(5.4)	47.8	(6.1)	48	(5.6)
Male, n (%)	104	(25.5)	89	(25.9)	61	(21.1)	254	(24.4)	65.0	(27.3)	71.0	(30.5)	56.0	(25.5)	192	(27.8)
Prior CVD, n (%)	221	(54.2)	195	(56.9)	161	(55.7)	577	(55.5)	88.0	(37.0)	104.0	(44.6)	83.0	(37.7)	275	(39.8)
Prior revascularization, n (%)	111	(27.2)	97	(28.3)	88	(30.4)	296	(28.5)	28.0	(11.8)	41.0	(17.6)	32.0	(14.5)	101	(14.6)
Hypertension, n (%)	209	(51.2)	175	(51.0)	141	(48.8)	525	(50.5)	106.0	(44.5)	107.0	(45.9)	101.0	(45.9)	314	(45.4)
Diabetes, n (%)	81	(19.9)	69	(20.1)	54	(18.7)	204	(19.6)	46.0	(19.3)	32.0	(13.7)	30.0	(13.6)	108	(15.6)
Cholesterol, n (%)	242	(59.3)	212	(61.8)	190	(65.7)	644	(61.9)	121.0	(50.8)	132.0	(56.7)	127.0	(57.7)	380	(55.0)
BMI kg/m2, mean(sd)	27.1	(4.4)	27.0	(4.6)	27.6	(4.2)	27.2	(4.4)	29.7	(7.0)	29.6	(6.3)	29.0	(5.4)	29.4	(6.3)
Current Smoker, n (%)	115.0	(28.2)	116	(33.8)	82	(45.2)	313	(30.1)	97.0	(40.8)	103.0	(44.2)	102.0	(46.4)	302	(43.7)
Angiogram, n (%)	229	(56.1)	191	(55.7)	168	(49.4)	588	(56.5)	193	(81.1)	192	(82.4)	187	(85.0)	572	(82.8)
Medications at discharge																
ASA, n (%)	378	(92.6)	313	(91.3)	261	(28.4)	952	(91.5)	234.0	(98.3)	229.0	(98.3)	216.0	(98.2)	679	(98.3)
Other antiplatelet, n (%)	160	(39.2)	139	(40.5)	132	(58.1)	431	(41.4)	209.0	(87.8)	199.0	(85.4)	198.0	(90.0)	606	(87.7)
Beta-blocker, n (%)	317	(77.7)	273	(79.6)	237	(90.3)	827	(79.5)	208.0	(87.4)	204.0	(87.6)	186.0	(84.5)	598	(86.5)
Statin, n (%)	296	(72.5)	274	(79.9)	236	(45.7)	806	(77.5)	227.0	(95.4)	218.0	(93.6)	206.0	(93.6)	651	(94.2)
ACEi or ARB	235	(57.6)	196	(57.1)	178	(82.0)	609	(58.6)	236.0	(99.2)	233.0	(100.0)	219.0	(99.5)	688	(99.6)

GRS:genetic risk score, sd:standard deviation, CVD cardiovascular disease, BMI body mass index, ASA acetyl salicylic acid, ACE angiotensin converting enzyme, ARB angiotensin receptor blocker,

Table 4.2: Univariate and Multivariate models of covariates

	RISCA						GENESIS - PRAXY					
	Univariate			Multivariate			Univariate			Multivariate		
Covariate	HR	95% CI	P-value	HR	95% CI	P-value	HR	95% CI	P-value	HR	95% CI	P-value
GRS	0.96	0.90-1.02	0.194	0.97	0.90-1.03	0.316	1.04	0.98-1.10	0.237	1.03	0.97-1.09	0.360
Age	1.04	1.02-1.06	<0.001	1.02	0.99-1.04	0.183	0.98	0.95-1.02	0.325	0.98	0.94-1.01	0.183
Female gender	1.31	0.87-1.98	0.190	1.05	0.61-1.80	0.860	1.16	0.75-1.79	0.513	1.15	0.73-1.82	0.534
Prior CVD	2.56	1.54-4.23	<0.001	1.95	1.02-3.41	0.043	1.16	0.77-1.74	0.487	0.97	0.60-1.56	0.886
Prior revascularization	1.40	0.94-2.08	0.095	1.06	0.64-1.86	0.762	1.39	0.83-2.32	0.209	1.15	0.62-2.14	0.655
Hypertension	1.74	1.17-2.57	0.006	0.97	0.57-1.67	0.929	1.32	0.88-1.98	0.178	1.25	0.80-1.96	0.319
Diabetes	2.36	1.60-3.50	<0.001	2.10	1.22-3.52	0.007	1.06	0.61-1.84	0.836	0.93	0.51-1.69	0.812
Cholesterol	1.52	1.00-2.30	0.50	1.75	0.92-3.19	0.092	1.38	0.91-2.10	0.125	1.25	0.79-1.98	0.342
BMI	1.01	0.98-1.05	0.395	0.97	0.92-1.03	0.330	1.00	0.97-1.03	0.935	0.99	0.96-1.03	0.774
Current smoker	0.84	0.49-1.44	0.529	1.60	0.91-2.86	0.100	1.13	0.76-1.70	0.548	1.17	0.77-1.77	0.468
Medications												
ASA	1.10	0.51-2.40	0.798	0.97	0.44-2.14	0.939	1.62	0.23-11.62	0.632	1.52	0.20-11.45	0.685
Clopidogrel or other thienopyridine	0.81	0.52-1.26	0.348	0.83	0.43-1.17	0.178	0.94	0.51-1.72	0.843	0.84	0.45-1.57	0.593
B-blockers	1.14	0.66-1.95	0.646	1.33	0.72-2.49	0.352	1.21	0.64-2.26	0.558	1.21	0.63-2.32	0.570
Statins	0.67	0.43-1.06	0.088	0.47	0.27-0.85	0.012	1.16	0.47-2.85	0.748	1.04	0.41-2.67	0.928
ACEi or ARB	1.32	0.85-2.06	0.216	1.11	0.66-1.83	0.710	0.37	0.05-2.63	0.319	0.38	0.05-2.88	0.348

HR: Hazard ratio, CI: Confidence interval, GRS:genetic risk score, BMI: body mass index, ASA acetyl salicylic acid, ACEi angiotensin converting enzyme inhibitor, ARB angiotensin receptor blocker.

Table 4.3: Hazard ratios for various Cox proportional hazard models of GRS

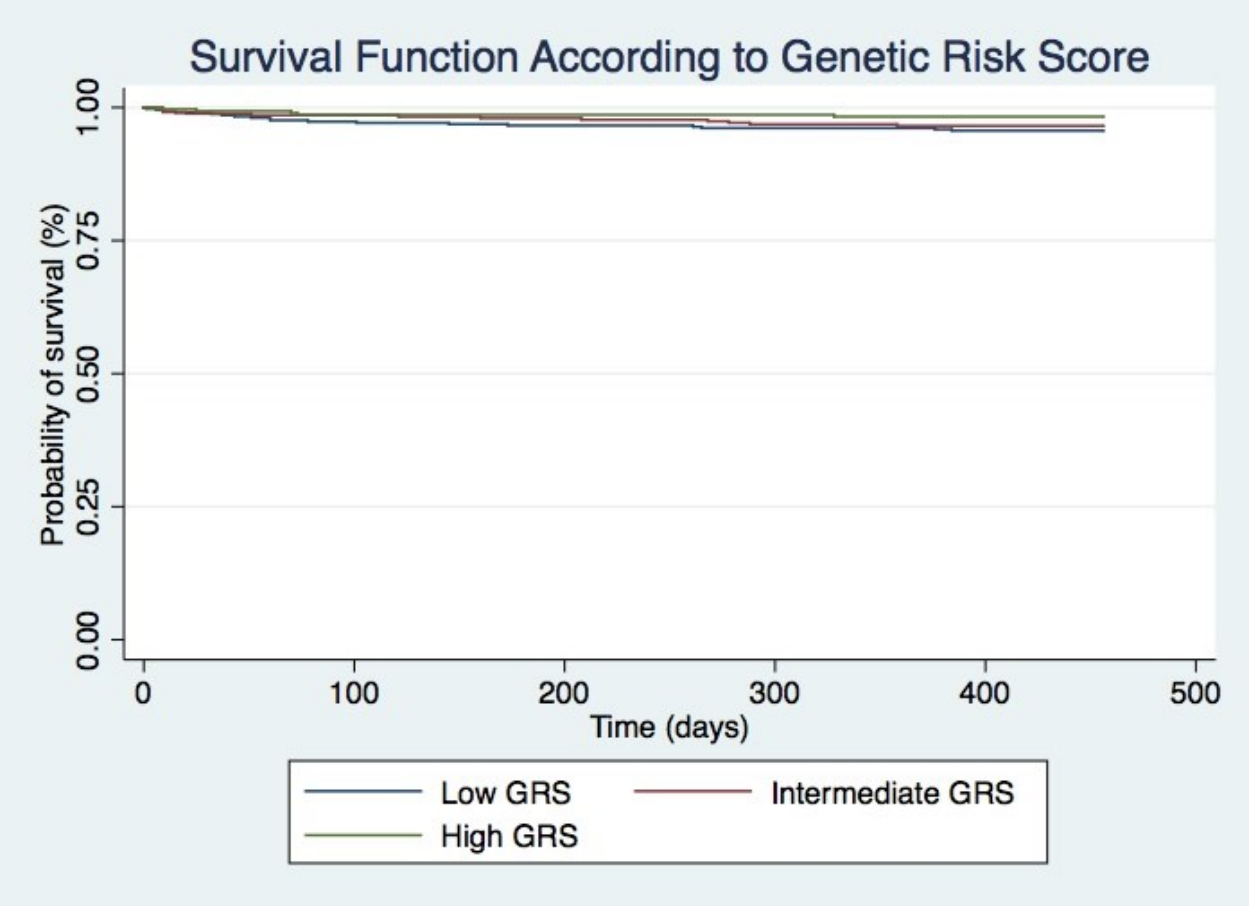
	RISCA		***GENESIS-PRAXY***		***Pooled Results***	
	HR	***95% Confidence Interval***	***HR***	***95% Confidence Interval***	***HR***	***95% Confidence Interval***
Combined Endpoint						
GRS - unadjusted	0.96	0.90-1.02	1.04	0.98-1.10	1.00	0.93-1.08
GRS - adjusted for age & sex	0.97	0.91-1.03	1.04	0.98-1.10	1.01	0.94-1.08
GRS - multivariate model	0.97	0.90-1.03	1.03	0.97-1.09	1.00	0.94-1.06
GRS adjusted for GRACE score	0.96	0.89-1.03	1.04	0.98-1.10	1.00	0.93-1.08

GRS: genetic risk score

Table 4.4: Cox model of 8 SNP GRS of LDL SNPs

	RISCA		***GENESIS-PRAXY***		***Pooled Results***	
	HR	***95% Confidence Interval***	***HR***	***95% Confidence Interval***	***HR***	***95% Confidence Interval***
Combined Endpoint						
GRS - unadjusted	0.88	0.77-1.01	1.15	1.02-1.30	1.01	0.78-1.31
GRS - adjusted for age & sex	0.88	0.77-1.00	1.15	1.02-1.29	1.01	0.78-1.31
GRS - multivariate model	0.89	0.77-1.04	1.14	1.01-1.29	1.01	0.80-1.30
GRS adjusted for GRACE score	0.90	0.77-1.05	1.15	1.02-1.30	1.02	0.80-1.30

Figure 4.1 Suvival Function According to Tertiles of Genetic Risk Score



4.4 Results for each individual SNP

In the preceding manuscript, neither the 30 SNP MI GRS nor the 8 SNP LDL GRS were consistently associated with recurrent events. Each SNP was then investigated individually and as part of the univaraite and multivariate Cox model. The results were then pooled across both cohorts and are presented in Table 4.5 below.

In our study, the predictive capability of rs579459 that was seen in the GRACE registry analysis was not replicated. While the GRACE cohort was comprised of cohorts from the UK, Belgium and Poland ours was a North American population in Quebec and New Brunswick in RISCA and a largely pan-Canadian population in GENESIS-PRAXY . Since ABO blood types can differ across regions and nations, it is possible that this difference may explain the discrepancy. In RISCA we identified two other SNPs that were associated with adverse events rs11206510 and rs2259816 which coded for variants in the PCSK9 and HNF1A gene regions. In GENESIS-PRAXY we identified a third SNP, rs964184 which codes for a variant in the ApoA gene. Interestingly, HNF1A is a critical activator of PCSK9 [121-123], suggesting a common mechanistic pathway may be involved in this association. PCSK9 inhibition remains an active field of research in the treatment of cardiovascular disease [124]. An analysis of diabetic patients from the Nurses Health study and the Health Professional Follow-up Study identified 5 SNPs associated with coronary heart disease. PCSK9 and HNF1A were both significantly associated with coronary heart disease with an OR of 1.26 (95%CI 1.09-1.47) and 1.17 (95%CI 1.04-1.32) respectively.

Unexpectedly, both SNPs showed a negative protective association with our outcome of recurrent adverse events. Although at least one other study in an Asian population [125] suggested

that, similar to our results, the C allele was the risk allele rather than the T allele as was stated in the original GWAS study.[49] However, these results were not replicated in the GENESIS-PRAXY cohort which suggests that the observed association may have been due solely to chance. Another possibility is that the observed association was due to recurrent event bias, which remains a frequent problem in the recurrent events literature.[126] The failure to replicate the observed associations underscores the importance of using a replication dataset in genetic analyses or whenever multiple testing will be used.

Ultimately, pooling the results across both datasets did not yield any statistically significant results. However one SNP, rs12526453, was very nearly statistically significant: HR 1.23 (95%CI 0.999-1.520). It is likely that in a larger sample with greater power this result would have been statistically significant but likely would still not have met the threshold for significance after adjustment for the Bonferroni correction.

Table 4.5: Hazard ratios for each SNP in each cohort and pooled across both populations

	RISCA		GENESIS PRAXY		POOLED		
	HR	95% CI	HR	95% CI	HR	95%	CI
rs10953541	1.19	0.82-1.73	1.11	0.80-1.56	1.145	0.895	1.465
rs11206510	0.64	0.45-0.91	1.46	0.97-2.22	0.905	0.694	1.180
rs1122608	1.12	0.78-1.62	1.23	0.87-1.73	1.177	0.919	1.508
rs11556924	1.04	0.82-1.32	1.19	0.96-1.48	1.120	0.955	1.312
rs12190287	0.80	0.58-1.10	0.85	0.63-1.15	0.826	0.665	1.026
rs12413409	0.90	0.48-1.66	1.08	0.61-1.89	0.994	0.657	1.504
rs12526453	1.20	0.88-1.65	1.26	0.95-1.69	1.232	0.999	1.520
rs12936587	1.26	0.93-1.73	0.83	0.62-1.11	1.009	0.818	1.245
rs1412444	0.98	0.71-1.36	0.99	0.73-1.35	0.985	0.790	1.229
rs17114036	0.82	0.49-1.36	1.12	0.65-1.94	0.948	0.655	1.372
rs1746048	1.01	0.60-1.69	1.13	0.66-1.93	1.066	0.737	1.542
rs17465637	1.05	0.76-1.45	1.03	0.75-1.42	1.040	0.831	1.302
rs17609940	0.95	0.65-1.37	0.84	0.59-1.17	0.889	0.692	1.141
rs216172	1.03	0.76-1.39	0.96	0.72-1.28	0.993	0.808	1.220
rs2259816	0.59	0.42-0.83	1.25	0.94-1.68	0.911	0.732	1.134
rs2505083	0.79	0.57-1.07	0.98	0.74-1.30	0.890	0.723	1.096
rs2895811	0.88	0.64-1.20	0.96	0.72-1.28	0.923	0.748	1.138
rs3184504	1.04	0.77-1.41	0.88	0.67-1.16	0.949	0.776	1.160
rs3798220	0.72	0.26-1.95	0.33	0.08-1.32	0.552	0.246	1.241
rs3825807	1.15	0.85-1.57	1.25	0.93-1.69	1.200	0.971	1.484
rs46522	1.01	0.75-1.37	0.83	0.62-1.11	0.913	0.742	1.123
rs4773144	0.97	0.71-1.32	0.85	0.64-1.13	0.903	0.734	1.111
rs4977574	1.15	0.85-1.57	0.78	0.59-1.04	0.933	0.759	1.146
rs579459	0.98	0.69-1.39	0.91	0.65-1.27	0.943	0.742	1.198
rs646776	1.24	0.82-1.88	1.27	0.87-1.85	1.256	0.953	1.656
rs6725887	1.18	0.79-1.77	1.19	0.80-1.76	1.185	0.897	1.567
rs964184	0.90	0.57-1.43	1.47	1.03-2.11	1.221	0.923	1.616
rs974819	0.95	0.69-1.31	1.03	0.77-1.39	0.992	0.800	1.231
rs9818870	0.66	0.40-1.08	1.16	0.80-1.67	0.950	0.709	1.273

*Statistically significant results are highlighted in yellow

4.5 Afterward to Manuscript #1

Although our study showed no significant association between the GRS and recurrent events in either RISCA or GENESIS-PRAXY, the analysis of individual SNPs demonstrated three SNPs (two in RISCA and one in GENESIS-PRAXY) that were associated with the outcome. However, none of these SNPs was replicated in the other cohort thus failing to provide definitive proof of an association. In the RISCA cohort, there was a non-statistically significant protective association with the GRS. This result was being drive by two individual SNPs that were strongly associated with the primary outcome but paradoxically in the opposite direction from what the prior literature and medical knowledge would suggest. Although, there are many possible explanations for this observation, including the play of chance, a form of collider stratification bias known as index event bias remains a plausible explanation. A formal discussion of index event bias and how it may have affected the results of our study is discussed in the following chapter.

Chapter 5: Index event bias

5.1 Apparent paradoxes in recurrent risk prediction

There are a number of examples in the medical literature where a risk factor for a condition is found to be paradoxically protective for recurrent events. The most well known example is the birth weight paradox where low birth weight infants where found to have better outcomes if they were the offspring of smoking mothers versus non-smoking mothers.[127,128] However, numerous other examples exist as well. There is the thrombophilia paradox where thrombophilias are associated with the occurrence of deep vein thrombosis but not the recurrence of thrombosis.[129] There is also the aspirin paradox where aspirin which is normally protective against MI is associated with worse outcomes and recurrent events after a patient's first MI[130] and the obesity paradox where higher BMI appears to be protective in cardiovascular disease.[131]

5.2 Explaining the recurrent risk paradox

There have been a number of potential explanations for these apparent paradoxes. These apparent paradoxes all have one element in common, namely that inclusion into the study or research protocol will condition upon the presence of a first or index event. Some authors have suggested that this represents a form of over-adjustment, namely the adjustment for an intermediary along the causal pathway. [132] This would be a reasonable explanation given that, from a temporal point of view, risk factors would lead to an index event. However, most authors acknowledge these paradoxes are due to a form of selection bias.[133] In a scenario where multiple risk factors are associated with the development of the index event, then the index event becomes a collider. A collider is defined as a variable that is the outcome (or descendant) of two other

variables or as a variable where two arrowheads intersect on a directed acyclic graph.[134] The relationship can be easily appreciated by graphically representing elements using directed acyclic graphs.[135,136] In Figure 5.1, the Index Event is a collider.

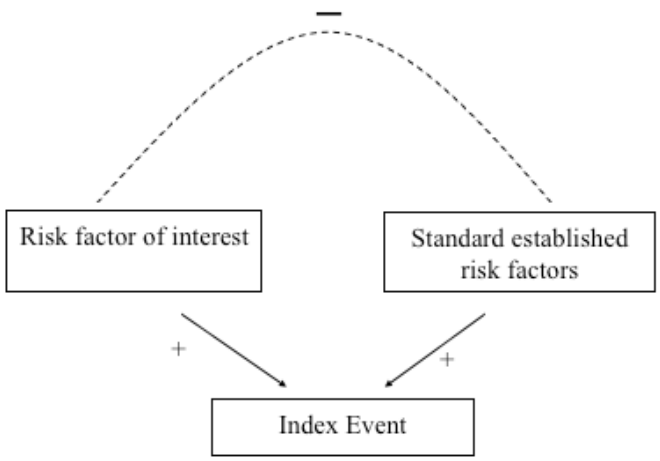


Figure 5.1: Induction of negative association in collider stratification bias

Since all studies looking at risk of recurrent events must by necessity condition upon the index event for inclusion into the study, then one is effectively conditioning upon a collider. [137] The net effect of conditioning on a collider will be to induce a negative association between the risk factors under investigation and subsequently a negative association between the risk factor of interest and the. (Figure 5.2)

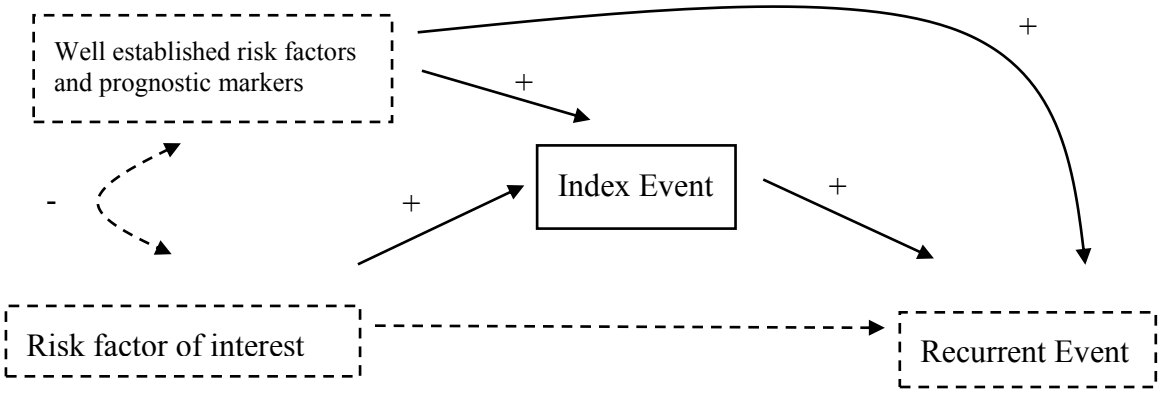


Figure 5.2: Induction of index event bias by conditioning on the index event

*Dashed curved arrow reflects an induced association. Dashed straight arrow represents the association under investigation.

Thus, in such a situation, the true positive association between the risk factor and the recurrent event is being biased towards the null. Smits et al.[138] provide a hypothetical example where such confounding could bias the true causal risk ration (RR) from a RR of 9 to a RR of 1. However, in the clinical examples provided above, the causal estimates are not merely biased towards the null but are qualitatively different and suggest a protective rather than harmful effect. Whitcomb et al.[139] provide another simulation to demonstrate that the degree of confounding between the unmeasured confounder and exposure of interest may explain such extreme results and result in such qualitative bias. Dehabreh and Kent [126] make a case that such biases explain many, if not all, of the paradoxes seen in the clinical literature.

5.3 Potential index event bias with the PCSK9 gene

With respect to our study, we observed a paradox such that the risk allele for rs11206510 (PCSK9) was seen to be protective for recurrent events. The PCSK9 gene encodes a protein that causes the degradation of the LDL receptor. Thus gain of function mutations in PCSK9 lead to decreases in the number of LDL receptors and therefore higher serum cholesterol levels.[140] Antibodies that inhibit the PCSK9 protein have subsequently been shown to lower serum cholesterol levels, [141] and loss of function mutations of PCSK9 have been associated with a lower incidence of coronary heart disease. [142] Therefore, the relationship between PCSK9, LDL cholesterol level and the index MI can be graphically represented by the DAG in Figure 3.

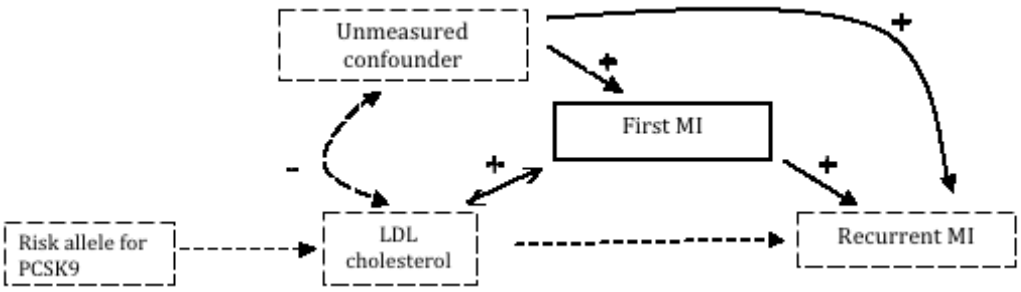


Figure 5.3: Index event bias in the association between PCSK9 and recurrent MI

The presence of a potential unmeasured confounder that is associated with both the index event of a first MI and the outcome of recurrent MI would result in collider stratification bias and the index event bias that has been described above. Consequently, because of this collider stratification bias, a negative association has been induced between LDL cholesterol and the unmeasured confounder whereas no such association might exist in the general population. Most of the other traditional risk factors for coronary heart disease, such as hypertension, diabetes, or smoking would be able to fit within the same causal diagram as they are associated with both first MI and recurrent events. However, by virtue of the fact that they are measured covariates, adjusting for these variables would effectively the block the back-door path responsible for the induced negative association. Only if there were residual confounding could one of the known covariates still provide the means for the back - door association to exist.[143]

Residual confounding could be possible in our study for a number of reasons. Firstly, although many risk factors and confounders for CAD are identified, there always remains the possibility that unknown confounders exist and may be contributing to residual confounding to some degree. Secondly, many of the traditional risk factors in our datasets were coded as dichotomous variables, being either present or absent. There was unfortunately no data available on actual blood pressure measurements, serum cholesterol levels, or glycated haemoglobin levels. As such,

all individuals with a given risk factor were treated the same in our analysis regardless of disease severity which may have contributed to residual confounding.

The seemingly and paradoxically protective effect of the two risk alleles rs11205610 (PCSK9) and rs2259816 (HNF1A) can potentially be explained by index event bias, caused by conditioning on the index event and inducing collider stratification bias between the SNP in question (or its descendant) and a confounder. Since we adjusted for all known confounder in our study, the confounder in question would have to be an unmeasured confounder or the product of residual confounding for incomplete adjustment.

Chapter 6: Manuscript #2

6.1 Preface to Manuscript #2

In the first manuscript, although there was no association between the GRS and recurrent events, there was more evidence for an association, a HR of 1.03 (95%CI 0.97-1.09) in the GENESIS-PRAXY cohort versus a HR of 0.97 (95%CI 0.90-1.03) in the RISCA cohort, although these differences were not statistically significant. This may have been due in part to the fact that the GENESIS-PRAXY cohort was composed of younger individuals than in the RISCA cohort since the inclusion criteria of GENESIS-PRAXY included a restriction that the age of presentation had to be under 55 years of age. It is possible that individuals with an earlier onset of their first ACS would be more likely to have a genetic contribution to their disease. As such, there could potentially be an association between the GRS and the age of the individual at presentation. The following manuscript, explores the association between a GRS and the age of presentation for a first ACS and compares the magnitude of the association of a GRS as compared to that of other traditional cardiac risk factors.

For this manuscript, only the RISCA dataset was used. The GENESIS-PRAXY cohort, being composed of individuals of less than 55 years of age, has a skewed age distribution and by definition includes individuals who presented with ACS at a younger age. As such, the analysis was conducted only the RISCA dataset which had a broader and more representative sample of ages of presentation.

6.2 Contribution of Authors

Dr. Labos, Dr. Engert and Dr. Thanassoulis conceived of the study design and the analysis plan. Dr. Brophy and Dr. Bogaty created the RISCA study cohort. Dr. Pilote created the GENESIS-PRAXY study cohort. Genotyping of individuals was performed by Leo Wang and Dr. Engert. All statistical analyses were performed by Dr. Labos. The first draft of the manuscript was written by Dr. Labos. All authors read and approved the final version of the manuscript. The manuscript presented here represents the version submitted for publication in Heart. Minor edits were made with respect to the numbering of figures to maintain consistency throughout the thesis and the references were incorporated into the global reference list at the end of the thesis. The abstract of this manuscript was presented at the American Heart Association EPI/NPAM conference in 2014 and was submitted to the journal Heart.

6.3 Traditional Risk Factors and a Genetic Risk Score Are Associated with Age of First Acute Coronary Syndrome

Christopher Labos MD CM*, Rui Hao Leo Wang BSc†, Louise Pilote MD MPH PhD†, Peter Bogaty MD§, James M. Brophy MD PhD*†, James C. Engert PhD‡, George Thanassoulis MD MSc†

* Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, QC, Canada

† Departments of Medicine, McGill University, Montreal, QC, Canada

‡ Department of Human Genetics, McGill University, Montreal, QC, Canada

§ Institut Universitaire de Cardiologie et de Pneumologie, Laval University, Quebec City, QC, Canada

Word Count: 4157 words

Brief Title: Risk factors and age of first ACS

Financial support: Dr. Thanassoulis is supported by a FRQS Chercheur Boursier Clinicien Salary Award. This work was supported by CIHR grant CIHR MOP-119380 to Dr. Thanassoulis as well as, in part, by CIHR IGO- 86113 to Dr. Pilote. The RISCA cohort was supported in part by the Fonds de la Recherche en Santé du Québec, the Heart and Stroke Foundation of Canada, and unrestricted grants from Merck Frosst Canada and Pfizer Canada. None of the authors have any relationship with private industry relevant to this manuscript.

Address for correspondence:
George Thanassoulis MD MSc FRCPC
Director, Preventive and Genomic Cardiology
McGill University Health Center
687 Pine Ave W. H4.55
Montreal, QC
H3A 1A1
Telephone: 514-934-1934 x 35465
E-mail: george.thanassoulis@mcgill.ca

Abstract

Objectives: To examine the association between traditional risk factors (TRF) and a genetic risk score (GRS) with age of first acute coronary syndrome (ACS).

Background: Early onset ACS may occur due to a high burden of TRFs or to genetic factors that accelerate atherosclerosis. Whether recently discovered genetic variants for ACS are more prevalent at earlier age of first ACS remains unknown.

Methods: Participants were drawn from the Recurrence and Inflammation in the Acute Coronary Syndromes (RISCA) cohort. To construct a multilocus GRS, participants were genotyped for 30 single nucleotide polymorphisms identified from prior genome-wide association studies. Linear regression models were fit to estimate the association between TRFs and GRS with age of first ACS.

Results: We included 460 participants with a first ACS enrolled in the RISCA cohort. Several TRFs were significantly associated with earlier age of first ACS: male sex [6.9 years earlier (95% confidence interval [CI] 4.1-9.7)], current cigarette smoking [8.1 years (95% CI 6.1-10.0)], overweight (BMI >25) and obesity (BMI>30) [5.2 years (95% CI 2.6-7.9)]. In women, hormone replacement therapy was also associated with earlier age of first ACS [4.3 years earlier (95% CI 0.3,8.4)]. After multivariable adjustment for TRFs, a one standard deviation increment in the GRS was associated with a 1.0 (95%CI 0.1-2.0) year earlier age of first ACS.

Conclusion: Among individuals with a first ACS, a GRS composed of 30 SNPs is associated with younger age of presentation. Although genetic predisposition modestly contributes to earlier ACS, a greater number of traditional risk factors is associated with markedly earlier ACS.

Abbreviations:

ACS = acute coronary syndrome

ASA = acetylsalicylic acid

BMI = body mass index

CAD = coronary artery disease

GRS = genetic risk score

GWAS = genome wide association study

HRT = hormone replacement therapy

RISCA = Recurrence and Inflammation in the Acute Coronary Syndromes

STEMI = ST elevation myocardial infarction

TRF = traditional risk factors

Introduction

Premature myocardial infarction (MI), defined as an MI occurring prior to age 55, affects 2.2% of men and 1.0% of women annually [144]. Given the typical trajectory of traditional risk factors for myocardial infarction, an early onset of MI is frequently attributed to the possible presence of unique genetic factors that may accelerate the atherosclerotic process or predispose to a final step in the causal pathway of MI (i.e. thrombosis or plaque rupture). However, early MI may also occur due to a very high burden of traditional risk factors. Traditional risk factors, such as smoking, hypertension, dyslipidemia, and diabetes, have a well-established role in the development of coronary artery disease [3,145-153] and several such factors have been associated with earlier age of first MI [11,154-156].

Family history of MI is a marker of genetic risk [57] that can be easily ascertained and is associated with earlier MI[156], but remains a crude marker of genetic risk since family members share only 50% of their genetic material. High-throughput genotyping of large samples has permitted the discovery of many single nucleotide polymorphisms (SNPs) that are robustly associated with MI. [45-52,157] However, little data exists regarding the association between these SNPs and age of first MI. [158-160] In addition, the effect of a genetic risk score (GRS) on age of first MI has not been investigated. Accordingly, we examined the impact of traditional risk factors as well as a multi-locus GRS composed of SNPs strongly associated with MI from large-scale genome wide association studies on the age of a first acute coronary syndrome (ACS).

Methods

Study Sample

The study participants were from the previously published RISCA (Recurrence and Inflammation in the Acute Coronary Syndromes) cohort study.[100] Briefly, 1210 consecutive patients were recruited from four tertiary and four Canadian community hospitals (seven in Quebec and one in New Brunswick). To be eligible, patients had to have an urgent admission to hospital with a diagnosis of either acute MI or unstable angina. All basic demographic and clinical data including traditional risk factors, results of all diagnostic and therapeutic procedures performed in hospital (including biochemical tests) and all medications taken prior to admission and prescribed at discharge were recorded in a comprehensive case report form and stored in electronic format on a secure server. This information was independently verified for consistency and then systematically assessed by on-site visits.

For the purposes of our analysis, only patients presenting to hospital with a first acute coronary syndrome (ACS) comprising both ST elevation and non-ST elevation myocardial infarctions, as well as unstable angina were eligible. Therefore, of the 1210 individuals enrolled in RISCA, we excluded all patients with a prior history of myocardial infarction (n=341). To ensure we included only individuals with their first cardiovascular event, we also excluded any patients with a history of coronary artery bypass grafting (n=70), percutaneous coronary intervention (n=82), stable and unstable angina, (n=121), stroke (n=18), heart failure (n=8), or prior cardiovascular admission (n=49). Thus, this study cohort consists of patients in whom ACS was the inaugural cardiac event (n=521). An additional 61 individuals were excluded from the analysis for missing data (n = 59 for individuals that did not consent to genetic testing and n=2 for missing covariates).

Outcome and Covariate Definitions

The primary outcome was age at first ACS, as recorded at the time of enrolment in the RISCA study. For enrollment into RISCA, myocardial infarction was defined as a history of characteristic chest discomfort or pain with an elevation of creatinine kinase – myocardial band to greater than 1.5 times the upper limit of normal. The definition was subsequently amended to also include a history of characteristic chest pain and elevation in cardiac troponin levels above the upper limit of normal. A diagnosis of unstable angina required either one episode of characteristic discomfort or pain at rest or with minimal exertion lasting more than 10 minutes or two episodes lasting more than 5 minutes with negative cardiac biomarkers. To increase specificity, UA patients had to have electrocardiogram changes consisting of 0.5 mm ST-segment depression or transient ST-segment elevation or 2 mm T-wave inversion in 2 contiguous leads.

Patients were defined as having diabetes if they had a self-reported history of diabetes (not including glucose intolerance) or were prescribed oral hypoglycemic agents or insulin at admission. Similarly, patients were defined as having hypertension or hypercholesterolemia if they had a history of hypertension or hypercholesterolemia or were prescribed antihypertensive therapy or lipid lowering therapy at admission, respectively. Current smokers were defined as patients who continued to smoke (>1 cigarette per day) at the time of enrolment or who had quit within the past 30 days. Body mass index (BMI) was categorized into three groups: normal if BMI≤25, overweight if BMI>25 but ≤30, and obese if BMI>30. History of acetylsalicylic acid (ASA) use was defined as ASA therapy up to 325 mg daily and hormone replacement therapy

(HRT) as any hormonal therapy prescribed to a menopausal woman (excluding any treatment for a malignancy) at the time of ACS admission.

Genotyping and Calculation of the GRS

Blood samples for all patients were taken within 24 hours of admission. DNA extraction from buffy coats as well as quantification and plating of samples was performed using standard techniques. Extracted DNA was stored at -80°C. Genotyping for 28 SNPs was performed using iPLEX technology on a MassARRAY Compact Analyser (Sequenom Inc., San Diego, CA, USA). Two SNPs, rs4977574 and rs1412444, were genotyped using a Custom TaqMan SNP Genotyping Assay (Applied Biosystems, Foster City, CA, USA). For both the Sequenom and Taqman data, all SNP call rates were >97.5%. All samples had >80% successful genotypes and all SNPs were in Hardy-Weinberg equilibrium (p>0.002).

The genetic risk score (GRS) was constructed using genotypes from 30 SNPs robustly associated and replicated in published genomewide association studies (GWAS) of myocardial infarction or coronary artery disease [45-52,57,130]. As performed in prior work[90], a score for each individual was calculated as the unweighted sum of each risk allele across all 30 SNPs (i.e score of 2 for those homozygous for the risk allele, a score of 1 for heterozygotes, and a score of zero for the absence of the risk allele). Missing genotypes (<0.35% of all genotypes) were assumed to be missing at random (i.e. non-informative missingness) and were imputed as two times the risk allele frequency, using the risk allele frequencies from the entire data set. Thus, every individual could have a genetic risk score ranging from 0 to 60; the actual range observed was 18 to 40.

Statistical analysis

Continuous variables were reported as means with standard deviations. Categorical variables were reported as counts with proportions. The association between age of first ACS and all variables was assessed using multivariable linear regression. The GRS was considered as a continuous variable centered about the mean for the main analysis but we also expressed the results for the GRS divided into quintiles (1^{st} vs 5^{th} quintile). Linear regression models were fit with age at first MI as the dependent variable and GRS as the independent variable. The following covariates were included in the adjusted model: sex, hypertension, diabetes, hypercholesterolemia, current smoking, BMI (categorized as normal, overweight and obese), ASA use, and HRT use. The output of the regression model was reported as beta-coefficients representing the difference in the age of first ACS in years for each unit change of GRS or presence/absence of the covariate. The appropriateness of the linear model was assessed by plotting the model residuals against the covariates included in the model and by assessing the normality of the distribution of residuals.

To describe the risk factor burden and GRS scores across categories of age of first ACS, we divided the entire cohort into quintiles and calculated the mean number of risk factors and the GRS in each quintile. Means across quintiles were compared using one-way analysis of variance (ANOVA).

In secondary analyses, we examined each SNP comprising the GRS individually with age of ACS. We also conducted a sensitivity analysis of our main analysis including only patients with MI (i.e. excluding unstable angina). A p-value of 0.05 was considered significant for all analyses except in exploratory analyses where each SNP in the GRS was evaluated separately with age of ACS. For these per SNP associations, we considered a Bonferroni corrected p-value

of 0.002 (0.05/30 SNPs) to declare significance. All statistical testing was performed using STATA version 12 (StataCorp, College Station, Texas).

Results

A total of 460 individuals (mean age 59 +/- 12 years, 22.4% female) were included in the final analysis. Myocardial infarction was the admitting diagnosis for 360 (78.3%) individuals, of which there were 238 STEMIs. The remainder were diagnosed as having unstable angina. In the study sample, 32.6% had hypertension, 12.4% had diabetes, and 37.8% had hypercholesterolemia. Mean body mass index (BMI) was 27.0 +/- 4.4 and 40.7% of the population were current smokers. The mean GRS was 31.5 +/- 3.4 . Full details of the baseline characteristics are presented in Table 1.

Associations between traditional risk factors and age of first ACS

In fully adjusted models (all β coefficients expressed in years), male sex [β=-6.9 (95%CI -9.7,-4.1)], current smoking status [β=-8.1 95%CI(-10.0, -6.1)], HRT use [β=-4.3 (95%CI -8.4, -0.3)], and being either overweight (BMI>25) [β=-2.6 (95%CI -4.8, -0.3)] or obese (BMI>30) [β=-5.2 (95%CI -7.9, -2.6)] were all significantly associated ($p<0.05$) with younger age at first ACS (Table 2). ASA use [β=3.7 (95%CI 0.3, 7.0)] and hypertension [β=3.4 (95%CI 1.2-5.6)] were both associated with older age of first ACS. (Table 2) The number of traditional risk factors present (male sex, hypertension, diabetes, hypercholesterolemia, smoking, and obesity) was also strongly associated with the age of first ACS with a greater risk factor burden being associated with a younger age of first ACS, [β per additional risk factor =-2.2 (95%CI -3.2, -1.3) $p<0.001$].

A sensitivity analysis excluding unstable angina cases (i.e. limited only to first MI cases) demonstrated a similar point estimate.

Association between GRS and age of first ACS

In unadjusted analyses, the GRS was associated with a younger age at first ACS, [β−coefficient -0.4 yrs per unit change in GRS (95%CI -0.7 yrs,-0.1 yrs)]. In the fully adjusted model, the GRS remained significantly associated with a younger age at first ACS, [β= -0.3 yrs (95%CI -0.6 yrs,-0.01 yrs)], as did the comparison between the highest versus the lowest quintile of GRS, β= -3.0 yrs (95%CI -5.9 yrs,-0.03 yrs). A one standard deviation change (equivalent to 3.4 points) in the GRS was associated with an age at first ACS that was 1.0 years (95%CI 0.1, 2.0) earlier (**Figure 1**). Sensitivity analysis excluding unstable angina did not materially change these findings.

In exploratory analyses, we examined the association between individual SNPs that comprised the genetic risk score and age of ACS, two SNPs were nominally significant in multivariable analysis (at $p<0.05$): rs4977574 from chromosome 9p21 and rs9818870 from chromosome 3q22.3 but did not reach pre-specified thresholds for significance due to multiple testing (See Supplementary Table 4.3). The presence of each risk allele of 9p21 was associated with an age of first ACS that was 1.4 (95%CI 0.1, 2.7) years earlier and each risk allele of 3q22.3 was associated with an age of first ACS that was 1.9 (0.1, 3.7) years earlier.

Risk Factors and GRS by Age of ACS

To characterize the risk factor burden and GRS of earlier ACS cases as compared to later ACS, we examined the mean number of risk factors and the mean GRS across quintiles of age at

first presentation. We observed a linear trend toward both a higher number of risk factors ($p<0.001$) and a higher GRS with younger age of first ACS ($p=0.011$) (**Figure 2**). As compared to the oldest ACS cases (5^{th} age quintile; mean age = 78.3), the earliest ACS cases (1^{st} age quintile; mean age = 45.8 yrs) had a significantly higher mean number of risk factors (2.4 risk factors vs. 1.9 risk factors; $p=0.01$) and GRS (31.9 vs. 30.4 points; $p=0.038$).

Discussion

In our study of 460 individuals with a first ACS, a higher GRS was significantly associated with a younger age at first ACS. We found that for each standard deviation increment in GRS (3.4 units), age of first ACS was 1 year earlier. We also show that in the highest quintile of GRS, the mean age of first ACS was nearly 3 years earlier as compared to the age in the lowest quintile of GRS. We demonstrate that two individual SNPs in the GRS (at the 9p21 and 3q22 loci) were associated with significantly earlier ACS by 1.4 and 1.9 years per allele, respectively. Although these SNPs did not meet stringent criteria for significance due to multiple testing, our results corroborate the previously reported association between 9p21with a younger age of coronary artery disease onset and provide independent replication of this finding.[158-160] Lastly, we also confirm the association between several traditional risk factors and age of first ACS. Most notably, we found that current smoking and obesity were associated with markedly earlier first ACS (8 years and 5 years respectively). Based on our findings, age at first ACS in an obese, male smoker is markedly earlier than the age of first ACS in an individual without these risk factors and supports the notion that the presence of a heavy burden of risk factors is more common at earlier age of ACS. Our results demonstrate that both a GRS and traditional risk factors contribute to earlier age of ACS but that common genetic predisposition,

as measured by a GRS, is only modestly associated with earlier presentation, whereas a heavy burden of traditional risk factors appears to be a more strongly linked to earlier ACS. These results highlight the importance of a high burden of risk factors in early ACS and reinforce the need to promote risk factor reduction, especially cigarette smoking cessation and weight loss in young adults to prevent cardiovascular events at an earlier age.

Several studies have looked at the ability of a GRS to predict cardiovascular events. A GRS has been shown to be associated with prevalent coronary artery disease (CAD), [83] as well as incident [82,90] and recurrent cardiovascular events [93,95]. While some studies have demonstrated that a GRS offers modest incremental predictive ability [60,80,84-86], other studies have shown no benefit [79,81,87-89]. Most recently, it has been shown that a GRS improves risk reclassification by nearly 5% [161] and may improve treatment decisions with respect to statin use. In addition, a GRS also predicts severity of atherosclerosis on coronary angiography [162] as well as accelerated subclinical atherosclerosis[90], carotid intima medial thickness [163], and coronary artery calcium [164,165]. Our results extend these prior findings and demonstrate that a GRS also associates with an earlier age of first ACS. Although several studies have evaluated the impact of traditional risk factors with age of MI, there is very limited data on the role of genetic factors and age of ACS despite the strong expectation of a genetic effect in earlier ACS. Madala et al. have demonstrated that family history of cardiovascular disease was associated with a 3-year earlier age of first MI but did not examine genetic data.[156] Patel et al., have shown a trend between younger age and higher GRS among a cohort of participants undergoing angiography with known CAD but did not specifically examine the role of a GRS in first MI[166]. Thus, to our knowledge, our study is the first to specifically evaluate the impact of GRS as a correlate of age at first ACS.

Coronary heart disease in young individuals is usually assumed to have a genetic cause, [44] and while a family history of heart disease has been associated with premature MI [17,167] many traditional risk factors also play a role. Smoking has been shown to have the largest impact on the age of first MI. In several prior studies, smoking was a powerful predictor of age at first MI, with smokers having their first MI 9.2, 9.3 and 9.7 years earlier respectively.[154-156] This is largely in agreement with our results which showed smoking was associated with an age of first ACS that is 8.1 years earlier as compared to non-smokers. Male sex, as compared to female sex, was consistently associated with a 4-year earlier MI in all 3 prior studies, which is again consistent with our findings of a 6.9-year earlier ACS. BMI class has also been strongly associated with younger age at first MI. As compared to being normal weight, being overweight (BMI between 25 and 30) was associated with an ACS 2.6 years earlier in our study, which agrees with the data from Bahler et al. that also showed a 2.6-year earlier MI and the data from Madala et al. that showed a roughly 3-year earlier MI. Likewise obesity (BMI greater than 30), was associated with 5.2-year earlier ACS in our data and roughly 6.5 years earlier in the study by Madala et al. Furthermore, Madala et al. were able to show a clear linear relationship between age of first MI and increasing BMI levels into the extreme range above a BMI of 40.

Two medication classes were also significantly associated with age of first ACS. ASA use resulted in an age of first ACS that was 3.7 years later in our study, which is consistent with the findings from Bahler et al. that indicated 4.0 years later. We also demonstrated that use of HRT resulted in an earlier age of first ACS, a finding that has not been previously reported, to our knowledge. However, this finding is certainly compatible with the cardiac risks of HRT seen in the Women's Health Initiative trial. [168]

Although we demonstrate that a GRS is associated with earlier first ACS, the exact role of GWAS-identified variants has not been established. A high GRS may indicate an increased susceptibility to traditional atherosclerosis risk factors, a predisposing vascular defect and/or an increased propensity to thrombosis, [169] all which could increase the tendency for earlier ACS. Although the effect of the GRS on earlier ACS was modest (1 year earlier per SD of a GRS), this is within the range of other traditional risk factors. For example, a one standard deviation change in BMI was associated with an age at first ACS that was 2.2 years earlier. More importantly, at a population level, the impact of these modest changes in age of ACS may have important consequences. Earlier ACS leads to significant lifelong healthcare costs, potential loss of income in younger individuals and higher losses of quality-adjusted years of life.

Our results highlight both the importance of traditional risk factors as well as the contribution of common genetic markers to earlier ACS and may also have clinical implications in understanding the familial risk after premature ACS which is increased nearly 2-fold in first degree family members. [170] As we have shown, earlier ACS is strongly associated with a high burden of risk factors (i.e. male sex, obesity, smoking) with only a modest contribution from common genetic variation. The risk to family members (i.e. siblings) that do not share the same burden of risk factors may not be increased in these circumstances; however further studies directly evaluating the impact of risk factors in probands on familial risk are needed. Although a family history of premature MI is defined solely on the basis of age of first MI, our results suggest that a risk-based definition of "premature MI" for probands may be superior in determining which families are truly at higher risk. For example, a MI at 50 years of age in an obese male smoker may not be truly "premature" whereas a MI in a 68-year old woman with no risk factors may need to be considered premature with important implications for familial risk.

Whether future genetic studies using novel technologies such as exome and whole genome sequencing focusing specifically on unusually early MI cases based on lower traditional risk factor burden, rather than simply on age, will be more effective in identifying novel or rare predisposing variants remains to be seen.[171]

Our study has several strengths. The RISCA cohort contained very detailed data on subject phenotypes that allowed the incremental value of the GRS, independent of risk factors, to be calculated. Also the data was prospectively collected in multiple centers in both academic and community hospitals, ensuring good generalizability of the results. Also, we used 30 robust and validated genetic variants from GWAS studies of myocardial infarction or coronary artery disease to construct our GRS. In addition, the strong agreement between our findings and prior studies evaluating traditional risk factors and age of ACS, including the study of Madala et al., which included over 100,000 MI cases, demonstrates the validity of our study sample and the robustness of these associations.

Several limitations also deserve mention. First, risk factor data was obtained by self-report and chart review, which may have led to some misclassification. However, since any misclassification is likely to be non-differential it would tend to bias results toward the null and reduce the magnitude of observed associations. Second, certain traditional risk factors such as diabetes and hypercholesterolemia, which are important risk factors for ACS, were not associated with lower age of first ACS. While this may be due to misclassification or lack of power to detect such an effect, other studies also [155,156] found no association between younger age of first ACS and these risk factors. It is important to note that this does not indicate that these are not important risk factors for ACS; these results indicate only that the prevalence of these risk factors is not different in young as compared to old ACS cases. Whether refined

classification of these risk factors by severity (e.g. severe hypercholesterolemia) would have demonstrated a higher prevalence in earlier ACS requires further study. Our definition of hypercholesterolemia and hypertension also included statin or hypertension treatment at time of first ACS and may have contributed to the lack of association observed. Third, our study only included common genetic variants from MI GWAS published prior to 2012. We did not include more recently discovered common variants or rare variants that could have strengthened the association between the GRS and age. We also did not consider the known genetic contribution to TRFs as such contribution is captured by standard measures of the TRFs. Lastly, our study sample consisted of a cohort of ACS cases and therefore the incidence or relative risks for early ACS in the general population cannot be inferred. However, all risk factors used in our analysis, including the GRS, have been previously associated with increased risk of ACS.

In summary, a GRS composed of 30 cardiac SNPs was associated with a younger age of first ACS. Although common genetic predisposition modestly contributes to earlier ACS, several traditional risk factors are strongly associated with a markedly lower age of first ACS highlighting the importance of a heavy burden of risk factors as an important contributor to earlier age of first ACS.

Disclosures

No disclosures.

Acknowledgements

We acknowledge the collaboration of the principal investigators and research personnel of the participating centers in RISCA, in particular Luce Boyer, RN. We also acknowledge the genotyping expertise of Rosalie Frechette and the genotyping platform of the McGill University and Génome Québec Innovation Centre and Katia Desbiens & Michael Gritti for technical help.

Table 6.1: Baseline characteristics

Total n=460		
Age (years), mean (±SD)	58.6	±11.8
Male	357	77.6
Traditional Risk Factors		
Hypercholesterolemia	174	37.8
Hypertension	150	32.1
Diabetes	57	12.4
BMI kg/m^2 , mean(±SD)	27.0	±4.4
Smoking status		
Never	94	20.4
Former	181	39.4
Current	185	40.2
Diagnosis on Admission		
STEMI	238	51.7
NSTEMI	122	26.5
Unstable Angina	100	21.7
GRS, mean (±SD)	31.5	±3.4
Medications on Admission		
ASA	42	9.1
Thienopyridine	1	0.2
Beta-blocker	39	8.5
Statin	83	18.0
ACE-inhibitor	49	10.7
ARB	15	3.3
Diuretic	42	9.2
HRT	36	7.8%

All data expressed as numbers (n) and percentages except where otherwise specified.
SD – standard deviation; BMI – body mass index; STEMI – ST elevation myocardial infarction;
NSTEMI – Non-ST elevation myocardial infarction; GRS – genetic risk score;
ASA – acetylsalicylic acid; ACE-inhibitor – angiotensin converting enzyme inhibitor;
ARB – angiotensin receptor blocker; HRT – hormone replacement therapy

Table 6.2: Linear Associations between GRS and Risk Factors with Age of First ACS

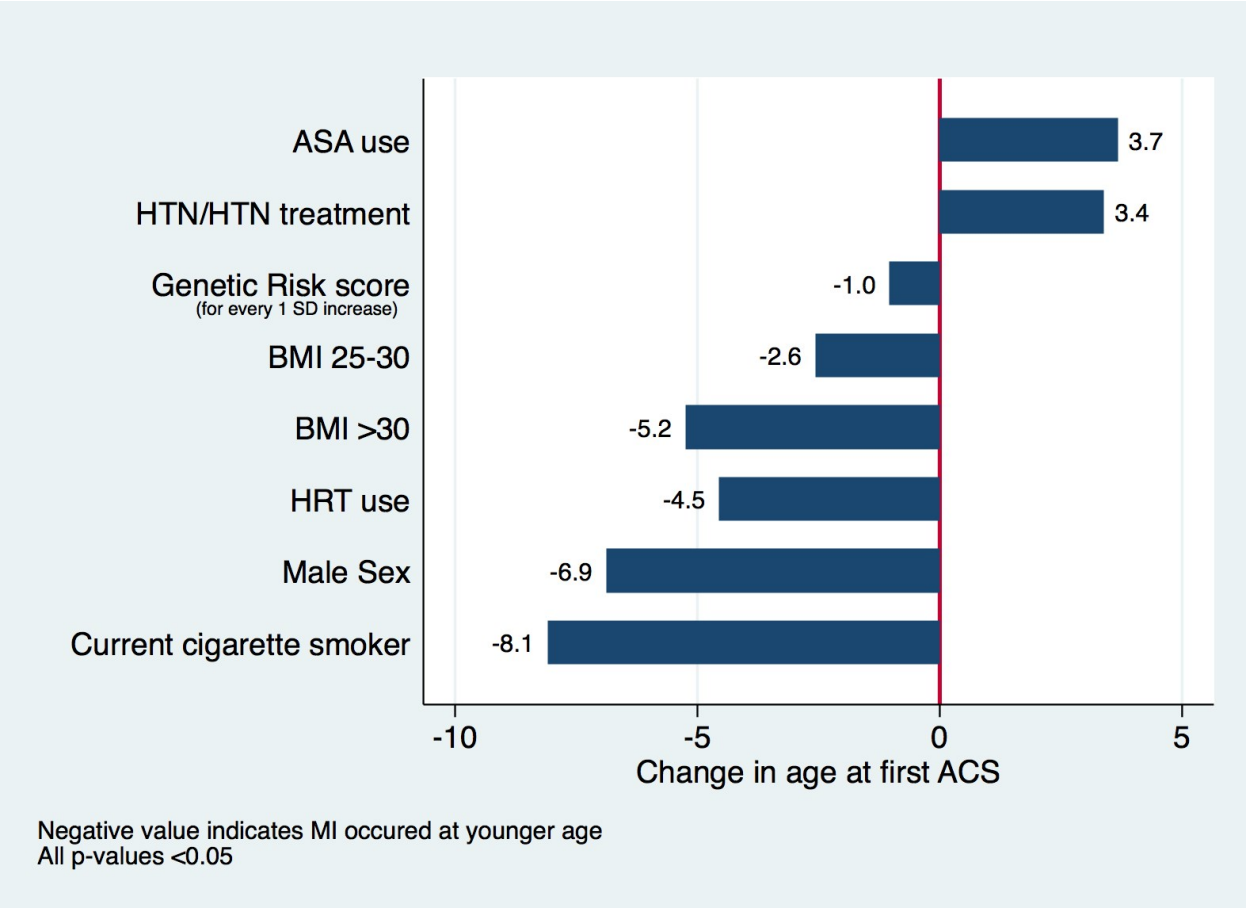
	Unadjusted			*Adjusted*[†]		
	β (years)	*95% Conf. Interval*	*P-value*	*β (years)*	*95% Conf. Interval*	*P-value*
GRS	-0.4	-0.7,-0.1	0.021	-0.3	-0.6,-0.01	0.040
Male sex	-8.2	-10.5, -5.8	<0.001	-6.9	-9.7, -4.1	<0.001
Hypertension	6.0	3.8, 8.2	<0.001	3.4	1.2, 5.6	0.003
Diabetes	3.1	-0.2, 6.4	0.062	0.3	-2.7, 3.3	0.846
Hypercholesterolemia	-0.4	-2.6, 1.9	0.745	-1.1	-3.1, 1.0	0.302
Current Smoker	-8.5	-10.6, -6.5	<0.001	-8.1	-10.0, -6.1	<0.001
Body Mass Index						
BMI 25-30	-2.1	-4.5, 0.3	0.090	-2.6	-4.8, -0.3	0.024
BMI >30	-4.5	-7.49, -1.59	0.003	-5.2	-7.9, -2.6	<0.001
ASA use	5.7	2.0, 9.4	0.003	3.7	0.3, 7.0	0.032
HRT use*	-4.3	-8.6, 0.02	0.051	-4.3	-8.4, -0.3	0.037

[†]Adjusted for age, sex, hypertension, diabetes, hypercholesterolemia, current smoking, body mass index, ASA use, HRT use and GRS. * Limited to women only

Table 6.3. Univariate and Multivariate Linear regression of individual SNPs

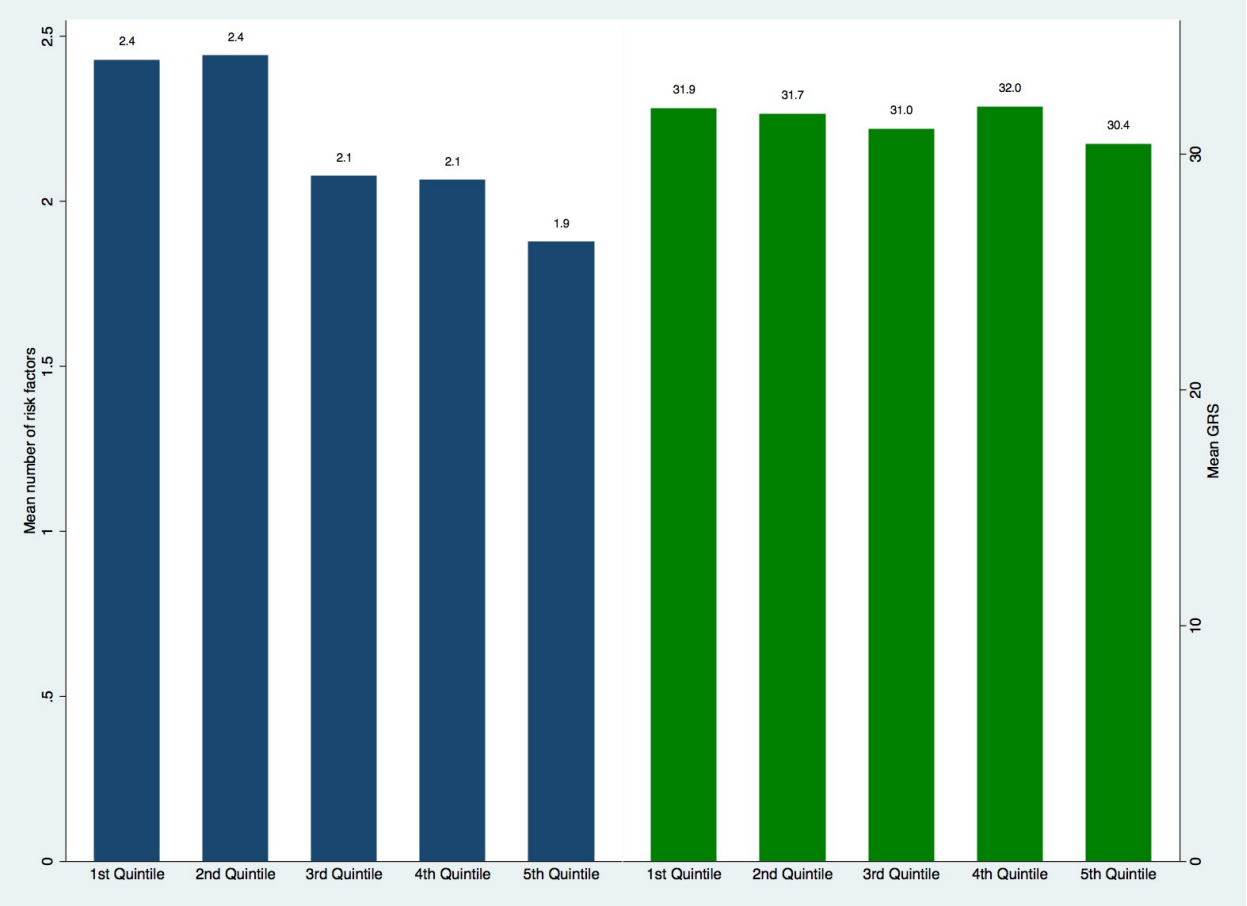
	Univariate				Multivariate			
	Coefficient	95% CI		p-value	Coefficient	95% CI		p-value
rs10953541	0.43	-1.32,	2.18	0.627	-0.17	-1.73	1.38	0.826
rs11206510	-0.39	-2.35,	1.58	0.700	0.17	-1.56	1.89	0.850
rs1122608	-1.03	-2.80,	0.75	0.258	-0.52	-2.08	1.05	0.516
rs11556924	-0.08	-1.24,	1.09	0.896	0.11	-0.92	1.15	0.829
rs12190287	-0.22	-1.86,	1.42	0.791	-0.31	-1.75	1.13	0.670
rs12413409	-0.71	-3.88,	2.47	0.661	-1.00	-3.81	1.80	0.482
rs12526453	0.58	-0.99,	2.14	0.469	0.15	-1.24	1.53	0.835
rs12936587	-0.82	-2.36,	0.71	0.294	-0.37	-1.74	1.00	0.592
rs1412444	-0.23	-1.82,	1.35	0.773	-0.42	-1.81	0.97	0.554
rs17114036	0.80	-1.87,	3.48	0.556	0.21	-2.15	2.57	0.860
rs1746048	1.04	-1.44,	3.52	0.411	0.67	-1.52	2.87	0.546
rs17465637	-1.59	-3.25,	0.06	0.059	-1.24	-2.70	0.22	0.096
rs17609940	1.02	-0.95,	3.00	0.309	-0.02	-1.78	1.74	0.983
rs216172	-0.29	-1.80,	1.22	0.705	0.01	-1.31	1.33	0.986
rs2259816	0.46	-1.10,	2.02	0.565	-0.77	-2.16	0.62	0.276
rs2505083	0.82	-0.71,	2.35	0.296	1.18	-0.16	2.53	0.085
rs2895811	-1.19	-2.74,	0.36	0.133	-0.95	-2.31	0.42	0.173
rs3184504	-1.03	-2.51,	0.45	0.173	-1.29	-2.59	0.01	0.052
rs3798220	0.32	-3.93,	4.57	0.882	-0.86	-4.63	2.90	0.653
rs3825807	0.94	-0.56,	2.45	0.217	1.19	-0.15	2.52	0.081
rs46522	-0.48	-2.02,	1.06	0.540	0.33	-1.03	1.70	0.631
rs4773144	-1.44	-2.99,	0.12	0.070	-1.26	-2.64	0.11	0.072
rs4977574	-1.62	-3.11,	-0.13	0.034	-1.37	-2.70	-0.05	0.042
rs579459	-0.34	-2.06,	1.38	0.696	0.17	-1.36	1.69	0.832
rs646776	-1.06	-2.90,	0.79	0.260	-0.40	-2.03	1.23	0.629
rs6725887	0.78	-1.52,	3.07	0.505	0.58	-1.45	2.60	0.576
rs964184	-1.83	-3.96,	0.31	0.093	-1.40	-3.30	0.50	0.148
rs974819	0.15	-1.41,	1.70	0.854	0.16	-1.22	1.54	0.823
rs9818870	-1.45	-3.51,	0.61	0.166	-1.89	-3.70	-0.07	0.041
rs9982601	-0.16	-2.26,	1.94	0.881	0.39	-1.48	2.27	0.680

Figure 6.1: Association of Risk Factors and Genetic Risk Score with Age at First ACS



* 1 SD deviation change in Genetic Risk score equals a 3.4 point change in GRS
ASA – acetylsalicylic acid; HTN – hypertension; BMI – body mass index; HRT – hormone replacement therapy; ACS – acute coronary syndrome

Figure 6.2. Mean Number of Risk Factors and GRS per quintile of Age of First ACS



$p<0.001$ and $p=0.011$ for trend across age quintile for mean number of risk factors and mean GRS, respectively

Chapter 7: Conclusion

In this thesis, we report on the clinical use of a genetic risk score composed of 30 previously validated SNPs associated with MI, as well as a 8-SNP GRS composed of LDL SNPs. There was no association between the 30-SNP GRS and recurrent events in either the RISCA or GENSIS-PRAXY cohorts. However in the younger GENSIS-PRAXY cohort, composed entirely of individuals less than 55 years of age, the results were almost statistically significant HR 1.07 (95%CI 0.97-1.09). This suggests the possibility that a GRS does have utility in younger individuals presenting with an ACS, if not in the population at large. This would certainly be plausible since it is often assumed that an ACS occurring at a younger age is more likely to be genetic in origin. Unfortunately, the same finding was not observed in RISCA when the analysis was confined to individuals less than 55 years old nor when the results were pooled across both cohorts. Therefore, we could find no conclusive evidence to support the notion that a 30-SNP GRS was associated with recurrent events or offered any predictive ability above and beyond well established risk models like the GRACE risk score.

The 8-SNP GRS composed of LDL SNPs was associated with recurrent events in GENESIS-PRAXY but not in RISCA and not when the results were pooled across both cohorts. Again, although the results in GENESIS-PRAXY are interesting they would need to be replicated before this limited form GRS could be used clinically.

One of the reasons, the GRS did not demonstrate statistical significance was that in the RISCA cohort the GRS consistently showed a mild protective effect with hazard ratios consistently less than 1, a result that was driven principally by the findings of two individual SNPs: PCSK9 and HNF1A. In this thesis we explore one possible explanation for this paradoxical finding, the occurrence of index event bias. By conditioning on the index event, in

our analysis the ACS leading to admission, one induces a collider stratification bias and induces a negative association between two risk factors for the disease under investigation and induces a negative association between the risk factor and subsequent recurrent events. Although this is a possible explanation, residual confounding and chance occurrence remain alternative explanations.

Nevertheless, the possible utility of a GRS in individuals presenting with ACS at a younger age, as was seen in the GENESIS-PRAXY, prompted our second manuscript which was an investigation between the GRS and other traditional risk factors and their association with the age of presentation for a first ACS. Although, the GRS was indeed associated with an earlier age of first ACS, other traditional risk factors such as smoking, obesity, and male sex had a much larger magnitude of association with age of first ACS. Thus, while it is often assumed that coronary disease that occurs at a young age is due to genetic factors, it is important to take into account that the traditional risk factor burden may also be playing a substantial role in determining the age of presentation for ACS.

In conclusion, we could find no conclusive proof to support the use of a genetic risk score for the prediction of recurrent events. Furthermore, the size of the effect measures is small and, even if such an association did exist, it is unlikely to be large and clinically significant. Although the GRS was associated with the age of first ACS, other traditional risk factors had a much larger and stronger association. Therefore, while genetic factors clearly play a role in cardiovascular disease, they do not seem to offer any advantage above traditional risk factors and risk models which seem to confer the greater degree of risk for the individual.

References

1. Ridker PM, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. *JAMA : the journal of the American Medical Association.* Feb 14 2007;297(6):611-619.
2. Ridker PM, Paynter NP, Rifai N, Gaziano JM, Cook NR. C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds Risk Score for men. *Circulation.* Nov 25 2008;118(22):2243-2251, 2244p following 2251.
3. D'Agostino RB, Sr., Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation.* Feb 12 2008;117(6):743-753.
4. Laslett LJ, Alagona P, Jr., Clark BA, 3rd, et al. The worldwide environment of cardiovascular disease: prevalence, diagnosis, therapy, and policy issues: a report from the American College of Cardiology. *Journal of the American College of Cardiology.* Dec 25 2012;60(25 Suppl):S1-49.
5. Lloyd-Jones DM, Larson MG, Beiser A, Levy D. Lifetime risk of developing coronary heart disease. *Lancet.* Jan 9 1999;353(9147):89-92.
6. Berry JD, Dyer A, Cai X, et al. Lifetime risks of cardiovascular disease. *The New England journal of medicine.* Jan 26 2012;366(4):321-329.
7. Dawber TR, Kannel WB, Revotskie N, Stokes J, 3rd, Kagan A, Gordon T. Some factors associated with the development of coronary heart disease: six years' follow-up experience in the Framingham study. *American journal of public health and the nation's health.* Oct 1959;49:1349-1356.
8. Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J, 3rd. Factors of risk in the development of coronary heart disease--six year follow-up experience. The Framingham Study. *Ann Intern Med.* Jul 1961;55:33-50.
9. Canto JG, Kiefe CI, Rogers WJ, et al. Number of coronary heart disease risk factors and mortality in patients with first myocardial infarction. *JAMA : the journal of the American Medical Association.* Nov 16 2011;306(19):2120-2127.
10. Greenland P, Knoll MD, Stamler J, et al. Major risk factors as antecedents of fatal and nonfatal coronary heart disease events. *JAMA : the journal of the American Medical Association.* Aug 20 2003;290(7):891-897.

11. Khot UN, Khot MB, Bajzer CT, et al. Prevalence of conventional risk factors in patients with coronary heart disease. *JAMA : the journal of the American Medical Association.* Aug 20 2003;290(7):898-904.
12. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation.* May 12 1998;97(18):1837-1847.
13. D'Agostino RB, Sr., Grundy S, Sullivan LM, Wilson P. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. *JAMA : the journal of the American Medical Association.* Jul 11 2001;286(2):180-187.
14. Yusuf S, Hawken S, Ounpuu S, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet.* Sep 11-17 2004;364(9438):937-952.
15. Andresdottir MB, Sigurdsson G, Sigvaldason H, Gudnason V. Fifteen percent of myocardial infarctions and coronary revascularizations explained by family history unrelated to conventional risk factors. The Reykjavik Cohort Study. *European heart journal.* Nov 2002;23(21):1655-1663.
16. Chow CK, Islam S, Bautista L, et al. Parental history and myocardial infarction risk across the world: the INTERHEART Study. *Journal of the American College of Cardiology.* Feb 1 2011;57(5):619-627.
17. Lloyd-Jones DM, Nam BH, D'Agostino RB, Sr., et al. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. *JAMA : the journal of the American Medical Association.* May 12 2004;291(18):2204-2211.
18. Murabito JM, Pencina MJ, Nam BH, et al. Sibling cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults. *JAMA : the journal of the American Medical Association.* Dec 28 2005;294(24):3117-3123.
19. Otaki Y, Gransar H, Berman DS, et al. Impact of family history of coronary artery disease in young individuals (from the CONFIRM registry). *The American journal of cardiology.* Apr 15 2013;111(8):1081-1086.
20. Roncaglioni MC, Santoro L, D'Avanzo B, et al. Role of family history in patients with myocardial infarction. An Italian case-control study. GISSI-EFRIM Investigators. *Circulation.* Jun 1992;85(6):2065-2072.
21. Sesso HD, Lee IM, Gaziano JM, Rexrode KM, Glynn RJ, Buring JE. Maternal and paternal history of myocardial infarction and risk of cardiovascular disease in men and women. *Circulation.* Jul 24 2001;104(4):393-398.

22. Sivapalaratnam S, Boekholdt SM, Trip MD, et al. Family history of premature coronary heart disease and risk prediction in the EPIC-Norfolk prospective population study. *Heart.* Dec 2010;96(24):1985-1989.
23. Conroy RM, Pyorala K, Fitzgerald AP, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *European heart journal.* Jun 2003;24(11):987-1003.
24. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ.* Jul 21 2007;335(7611):136.
25. Jneid H, Anderson JL, Wright RS, et al. 2012 ACCF/AHA focused update of the guideline for the management of patients with unstable angina/Non-ST-elevation myocardial infarction (updating the 2007 guideline and replacing the 2011 focused update): a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. *Circulation.* Aug 14 2012;126(7):875-910.
26. Antman EM, Cohen M, Bernink PJ, et al. The TIMI risk score for unstable angina/non-ST elevation MI: A method for prognostication and therapeutic decision making. *JAMA : the journal of the American Medical Association.* Aug 16 2000;284(7):835-842.
27. Mega JL, Morrow DA, Sabatine MS, et al. Correlation between the TIMI risk score and high-risk angiographic findings in non-ST-elevation acute coronary syndromes: observations from the Platelet Receptor Inhibition in Ischemic Syndrome Management in Patients Limited by Unstable Signs and Symptoms (PRISM-PLUS) trial. *American heart journal.* May 2005;149(5):846-850.
28. Damman P, Woudstra P, Kuijt WJ, et al. Short- and long-term prognostic value of the TIMI risk score after primary percutaneous coronary intervention for ST-segment elevation myocardial infarction. *Journal of interventional cardiology.* Feb 2013;26(1):8-13.
29. de Araujo Goncalves P, Ferreira J, Aguiar C, Seabra-Gomes R. TIMI, PURSUIT, and GRACE risk scores: sustained prognostic value and interaction with revascularization in NSTE-ACS. *European heart journal.* May 2005;26(9):865-872.
30. Morrow DA, Antman EM, Charlesworth A, et al. TIMI risk score for ST-elevation myocardial infarction: A convenient, bedside, clinical score for risk assessment at presentation: An intravenous nPA for treatment of infarcting myocardium early II trial substudy. *Circulation.* Oct 24 2000;102(17):2031-2037.

31. Granger CB, Goldberg RJ, Dabbous O, et al. Predictors of hospital mortality in the global registry of acute coronary events. *Archives of internal medicine.* Oct 27 2003;163(19):2345-2353.
32. Eagle KA, Lim MJ, Dabbous OH, et al. A validated prediction model for all forms of acute coronary syndrome: estimating the risk of 6-month postdischarge death in an international registry. *JAMA : the journal of the American Medical Association.* Jun 9 2004;291(22):2727-2733.
33. Tang EW, Wong CK, Herbison P. Global Registry of Acute Coronary Events (GRACE) hospital discharge risk score accurately predicts long-term mortality post acute coronary syndrome. *American heart journal.* Jan 2007;153(1):29-35.
34. Aragam KG, Tamhane UU, Kline-Rogers E, et al. Does simplicity compromise accuracy in ACS risk prediction? A retrospective analysis of the TIMI and GRACE risk scores. *PloS one.* 2009;4(11):e7947.
35. Center for Outcomes Research UoMMS. Methods and formulas used to calculate the GRACE Risk Scores for patients presenting to hospital with an acute coronary syndrome. 1998-2010; http://www.outcomes-umassmed.org/grace/files/GRACE_RiskModel_Coefficients.pdf. Accessed August 29, 2013.
36. Roe MT, Chen AY, Thomas L, et al. Predicting long-term mortality in older patients after non-ST-segment elevation myocardial infarction: the CRUSADE long-term mortality model and risk score. *American heart journal.* Nov 2011;162(5):875-883.e871.
37. Poci D, Hartford M, Karlsson T, Herlitz J, Edvardsson N, Caidahl K. Role of the CHADS2 score in acute coronary syndromes: risk of subsequent death or stroke in patients with and without atrial fibrillation. *Chest.* Jun 2012;141(6):1431-1440.
38. Fischer M, Broeckel U, Holmer S, et al. Distinct Heritable Patterns of Angiographic Coronary Artery Disease in Families With Myocardial Infarction. *Circulation.* February 22, 2005 2005;111(7):855-862.
39. Chan L, Boerwinkle E. Gene-Environment Interactions and Gene Therapy in Atherosclerosis. *Cardiology in Review.* 1994;2(3):130-137.
40. Nora JJ, Lortscher RH, Spangler RD, Nora AH, Kimberling WJ. Genetic--epidemiologic study of early-onset ischemic heart disease. *Circulation.* Mar 1980;61(3):503-508.
41. Berg AO, Baird MA, Botkin JR, et al. National Institutes of Health State-of-the-Science Conference Statement: Family History and Improving Health. *Ann Intern Med.* Dec 15 2009;151(12):872-877.

42. Genest J, McPherson R, Frohlich J, et al. 2009 Canadian Cardiovascular Society/Canadian guidelines for the diagnosis and treatment of dyslipidemia and prevention of cardiovascular disease in the adult - 2009 recommendations. *The Canadian journal of cardiology.* Oct 2009;25(10):567-579.
43. Murabito JM, Nam BH, D'Agostino RB, Sr., Lloyd-Jones DM, O'Donnell CJ, Wilson PW. Accuracy of offspring reports of parental cardiovascular disease history: the Framingham Offspring Study. *Ann Intern Med.* Mar 16 2004;140(6):434-440.
44. Topol EJ, McCarthy J, Gabriel S, et al. Single nucleotide polymorphisms in multiple novel thrombospondin genes may be associated with familial premature myocardial infarction. *Circulation.* Nov 27 2001;104(22):2641-2644.
45. Burton PR, Clayton DG, Cardon LR, et al. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007;447(7145):661-678.
46. Erdmann J, Großhennig A, Braund PS, et al. New susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nature Genetics.* 2009;41(3):280-282.
47. Gudbjartsson DF, Bjornsdottir US, Halapi E, et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nature Genetics.* 2009;41(3):342-347.
48. Helgadottir A, Thorleifsson G, Manolescu A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science.* 2007;316(5830):1491-1493.
49. Kathiresan S, Voight BF, Purcell S, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nature Genetics.* 2009;41(3):334-341.
50. McPherson R, Pertsemlidis A, Kavaslar N, et al. A common allele on chromosome 9 associated with coronary heart disease. *Science.* 2007;316(5830):1488-1491.
51. Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. *New England Journal of Medicine.* 2007;357(5):443-453.
52. Trégouët DA, König IR, Erdmann J, et al. Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. *Nature Genetics.* 2009;41(3):283-285.
53. Ahituv N, Kavaslar N, Schackwitz W, et al. A PYY Q62P variant linked to human obesity. *Human molecular genetics.* Feb 1 2006;15(3):387-391.

54. Ahituv N, Kavaslar N, Schackwitz W, et al. Medical sequencing at the extremes of human body mass. *American journal of human genetics.* Apr 2007;80(4):779-791.
55. Speliotes EK, Willer CJ, Berndt SI, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet.* Nov 2010;42(11):937-948.
56. International Consortium for Blood Pressure Genome-Wide Association S, Ehret GB, Munroe PB, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature.* Oct 6 2011;478(7367):103-109.
57. Schunkert H, Konig IR, Kathiresan S, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet.* Apr 2011;43(4):333-338.
58. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature.* Aug 5 2010;466(7307):707-713.
59. Willer CJ, Sanna S, Jackson AU, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet.* Feb 2008;40(2):161-169.
60. Kathiresan S, Melander O, Anevski D, et al. Polymorphisms associated with cholesterol and risk of cardiovascular events. *New England Journal of Medicine.* 2008;358(12):1240-1249.
61. Ridker PM, Pare G, Parker A, et al. Loci related to metabolic-syndrome pathways including LEPR,HNF1A, IL6R, and GCKR associate with plasma C-reactive protein: the Women's Genome Health Study. *American journal of human genetics.* May 2008;82(5):1185-1192.
62. Dehghan A, Dupuis J, Barbalic M, et al. Meta-analysis of genome-wide association studies in >80 000 subjects identifies multiple loci for C-reactive protein levels. *Circulation.* Feb 22 2011;123(7):731-738.
63. Collaboration CRPCHDG. Association between C reactive protein and coronary heart disease: mendelian randomisation analysis based on individual participant data. *BMJ.* 2011-02-15 23:34:41 2011;342.
64. Allan TM, Dawson AA. ABO blood groups and ischaemic heart disease in men. *British heart journal.* May 1968;30(3):377-382.
65. Reilly MP, Li M, He J, et al. Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. *Lancet.* Jan 29 2011;377(9763):383-392.
66. Visel A, Zhu Y, May D, et al. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature.* Mar 18 2010;464(7287):409-412.

67. Do R, Xie C, Zhang X, et al. The effect of chromosome 9p21 variants on cardiovascular disease may be modified by dietary intake: evidence from a case/control and a prospective study. *PLoS medicine.* Oct 2011;8(10):e1001106.
68. Gustavsson J, Mehlig K, Leander K, et al. Interaction of apolipoprotein E genotype with smoking and physical inactivity on coronary heart disease risk in men and women. *Atherosclerosis.* Feb 2012;220(2):486-492.
69. Humphries SE, Talmud PJ, Hawe E, Bolla M, Day IN, Miller GJ. Apolipoprotein E4 and coronary heart disease in middle-aged men who smoke: a prospective study. *Lancet.* Jul 14 2001;358(9276):115-119.
70. Webster AL, Yan MS, Marsden PA. Epigenetics and cardiovascular disease. *The Canadian journal of cardiology.* Jan 2013;29(1):46-57.
71. McPherson R. From Genome-Wide Association Studies to Functional Genomics: New Insights Into Cardiovascular Disease. *Canadian Journal of Cardiology.* 1// 2013;29(1):23-29.
72. Ioannidis JP. Prediction of cardiovascular disease outcomes and established cardiovascular risk factors by genome-wide association markers. *Circulation. Cardiovascular genetics.* Feb 2009;2(1):7-15.
73. Horne BD, Carlquist JF, Muhlestein JB, Bair TL, Anderson JL. Association of variation in the chromosome 9p21 locus with myocardial infarction versus chronic coronary artery disease. *Circulation. Cardiovascular genetics.* Dec 2008;1(2):85-92.
74. Ardissino D, Berzuini C, Merlini PA, et al. Influence of 9p21.3 genetic variants on clinical and angiographic outcomes in early-onset myocardial infarction. *Journal of the American College of Cardiology.* Jul 19 2011;58(4):426-434.
75. Patel RS, Su S, Neeland IJ, et al. The chromosome 9p21 risk locus is associated with angiographic severity and progression of coronary artery disease. *European heart journal.* Dec 2010;31(24):3017-3023.
76. Schunkert H, Gotz A, Braund P, et al. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation.* Apr 1 2008;117(13):1675-1684.
77. Morgan TM, Krumholz HM, Lifton RP, Spertus JA. Nonvalidation of reported genetic risk factors for acute coronary syndrome in a large-scale replication study. *JAMA : the journal of the American Medical Association.* Apr 11 2007;297(14):1551-1561.
78. Ntzani EE, Rizos EC, Ioannidis JP. Genetic effects versus bias for candidate polymorphisms in myocardial infarction: case study and overview of large-scale evidence. *Am J Epidemiol.* May 1 2007;165(9):973-984.

79. Paynter NP, Chasman DI, Buring JE, Shiffman D, Cook NR, Ridker PM. Cardiovascular disease risk prediction with and without knowledge of genetic variation at chromosome 9p21.3. *Annals of Internal Medicine.* 2009;150(2):65-72.
80. Talmud PJ, Cooper JA, Palmen J, et al. Chromosome 9p21.3 coronary heart disease locus genotype and prospective risk of CHD in healthy middle-aged men. *Clinical Chemistry.* 2008;54(3):467-474.
81. Paynter NP, Chasman DI, Pare G, et al. Association between a literature-based genetic risk score and cardiovascular events in women. *Jama.303.*631-637.
82. Ripatti S, Tikkanen E, Orho-Melander M, et al. A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. *Lancet.* Oct 23 2010;376(9750):1393-1400.
83. Davies RW, Dandona S, Stewart AF, et al. Improved prediction of cardiovascular disease based on a panel of single nucleotide polymorphisms identified through genome-wide association studies. *Circulation. Cardiovascular genetics.* Oct 2010;3(5):468-474.
84. Brautbar A, Ballantyne CM, Lawson K, et al. Impact of adding a single allele in the 9p21 locus to traditional risk factors on reclassification of coronary heart disease risk and implications for lipid-modifying therapy in the atherosclerosis risk in communities study. *Circulation: Cardiovascular Genetics.* 2009;2(3):279-285.
85. Morrison AC, Bare LA, Chambless LE, et al. Prediction of coronary heart disease risk using a genetic risk score: The atherosclerosis risk in communities study. *American Journal of Epidemiology.* 2007;166(1):28-35.
86. Humphries SE, Cooper JA, Talmud PJ, Miller GJ. Candidate gene genotypes, along with conventional risk factor assessment, improve estimation of coronary heart disease risk in healthy UK men. *Clinical Chemistry.* 2007;53(1):8-16.
87. Junyent M, Tucker KL, Shen J, et al. A composite scoring of genotypes discriminates coronary heart disease risk beyond conventional risk factors in the Boston Puerto Rican Health Study. *Nutrition, Metabolism and Cardiovascular Diseases.* 2010;20(3):157-164.
88. Trichopoulou A, Yiannakouris N, Bamia C, Benetou V, Trichopoulos D, Ordovas JM. Genetic predisposition, nongenetic risk factors, and coronary infarct. *Archives of internal medicine.* 2008;168(8):891-896.
89. Yamada Y, Izawa H, Ichihara S, et al. Prediction of the risk of myocardial infarction from polymorphisms in candidate genes. *New England Journal of Medicine.* 2002;347(24):1916-1923.

90. Thanassoulis G, Peloso GM, Pencina MJ, et al. A genetic risk score is associated with incident cardiovascular disease and coronary artery calcium: the Framingham Heart Study. *Circulation. Cardiovascular genetics.* Feb 1 2012;5(1):113-121.
91. Consortium CADCDG. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet.* Apr 2011;43(4):339-344.
92. Pencina MJ, D'Agostino RB, Sr., Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Statistics in medicine.* Jan 15 2011;30(1):11-21.
93. Buysschaert I, Carruthers KF, Dunbar DR, et al. A variant at chromosome 9p21 is associated with recurrent myocardial infarction and cardiac death after acute coronary syndrome: the GRACE Genetics Study. *European heart journal.* May 2010;31(9):1132-1141.
94. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Statistics in medicine.* Jan 30 2008;27(2):157-172; discussion 207-112.
95. Wauters E, Carruthers KF, Buysschaert I, et al. Influence of 23 coronary artery disease variants on recurrent myocardial infarction or cardiac death: the GRACE Genetics Study. *European heart journal.* Apr 2013;34(13):993-1001.
96. Cook NR, Ridker PM. Advances in measuring the effect of individual predictors of cardiovascular risk: the role of reclassification measures. *Ann Intern Med.* Jun 2 2009;150(11):795-802.
97. Ioannidis JP, Patsopoulos NA, Evangelou E. Heterogeneity in meta-analyses of genome-wide association investigations. *PloS one.* 2007;2(9):e841.
98. van der Net JB, Janssens AC, Sijbrands EJ, Steyerberg EW. Value of genetic profiling for the prediction of coronary heart disease. *American heart journal.* Jul 2009;158(1):105-110.
99. Janssens AC, Aulchenko YS, Elefante S, Borsboom GJ, Steyerberg EW, van Duijn CM. Predictive testing for complex diseases using multiple genes: fact or fiction? *Genetics in medicine : official journal of the American College of Medical Genetics.* Jul 2006;8(7):395-400.
100. Bogaty P, Boyer L, Simard S, et al. Clinical utility of C-reactive protein measured at admission, hospital discharge, and 1 month later to predict outcome in patients with acute coronary disease. The RISCA (recurrence and inflammation in the acute coronary syndromes) study. *Journal of the American College of Cardiology.* Jun 17 2008;51(24):2339-2346.

101. Pilote L, Karp I. GENESIS-PRAXY (GENdEr and Sex determInantS of cardiovascular disease: From bench to beyond-Premature Acute Coronary SYndrome). *American heart journal.* May 2012;163(5):741-746 e742.
102. RUBIN DB. Inference and missing data. *Biometrika.* December 1, 1976 1976;63(3):581-592.
103. Johansson AM, Karlsson MO. Comparison of methods for handling missing covariate data. *The AAPS journal.* Oct 2013;15(4):1232-1241.
104. Heitjan DF, Basu S. Distinguishing "Missing at Random" and "Missing Completely at Random". *The American Statistician.* 1996;50(3):207-213.
105. Cox DR. Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological).* 1972;34(2):187-220.
106. *EPICONF: Stata module to assess confounding effects in epidemiological studies* [computer program]. Boston College Department of Economics; 1998.
107. Greene WH. *Econometric Analysis. 7th edition.* Uppeer SAddle River, New Jersey: Prentice Hall; 2012.
108. Newson RB. Comparing the predictive powers of survival models using Harrell's C or Somers' D. *Stata Journal.* 2010;10(3):339-358.
109. Sundström J, Byberg L, Gedeborg R, Michaëlsson K, Berglund L. Useful tests of usefulness of new risk factors: Tools for assessing reclassification and discrimination. *Scandinavian Journal of Public Health.* June 1, 2011 2011;39(4):439-441.
110. Perneger TV. What's wrong with Bonferroni adjustments. *BMJ.* 1998-04-18 00:00:00 1998;316(7139):1236-1238.
111. Writing Group M, Lloyd-Jones D, Adams RJ, et al. Heart disease and stroke statistics--2010 update: a report from the American Heart Association. *Circulation.* Feb 23 2010;121(7):e46-e215.
112. Mulders TA, Meyer Z, van der Donk C, et al. Patients with premature cardiovascular disease and a positive family history for cardiovascular disease are prone to recurrent events. *International journal of cardiology.* Nov 17 2011;153(1):64-67.
113. Consortium W. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007;447: 661–678.
114. Erdmann J, Groszhennig A, Braund PS, et al. New susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat Genet.* 2009;41(3):280-282.
115. Helgadottir A. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science.* 2007;316:1491-1493.

116. Myocardial Infarction Genetics C, Kathiresan S, Voight BF, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet.* Mar 2009;41(3):334-341.
117. Samani NJ, Erdmann J, Hall AS, et al. Genomewide Association Analysis of Coronary Artery Disease. *N Engl J Med.* August 2, 2007 2007;357(5):443-453.
118. Tregouet DA, Konig IR, Erdmann J, et al. Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. *Nat Genet.* Mar 2009;41(3):283-285.
119. Consortium CD. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet.* 2011;43(4):339-344.
120. Schunkert H, Konig IR, Kathiresan S, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet.* 2011;43(4):333-338.
121. Ai D, Chen C, Han S, et al. Regulation of hepatic LDL receptors by mTORC1 and PCSK9 in mice. *The Journal of clinical investigation.* Apr 2 2012;122(4):1262-1270.
122. Li H, Dong B, Park SW, Lee HS, Chen W, Liu J. Hepatocyte nuclear factor 1alpha plays a critical role in PCSK9 gene transcription and regulation by the natural hypocholesterolemic compound berberine. *The Journal of biological chemistry.* Oct 16 2009;284(42):28885-28895.
123. Wu M, Dong B, Cao A, Li H, Liu J. Delineation of molecular pathways that regulate hepatic PCSK9 and LDL receptor expression during fasting in normolipidemic hamsters. *Atherosclerosis.* Oct 2012;224(2):401-410.
124. Dong B, Wu M, Li H, et al. Strong induction of PCSK9 gene expression through HNF1alpha and SREBP2: mechanism for the resistance to LDL-cholesterol lowering effect of statins in dyslipidemic hamsters. *Journal of lipid research.* Jun 2010;51(6):1486-1495.
125. Xu C, Wang F, Wang B, et al. Minor allele C of chromosome 1p32 single nucleotide polymorphism rs11206510 confers risk of ischemic stroke in the Chinese Han population. *Stroke; a journal of cerebral circulation.* Aug 2010;41(8):1587-1592.
126. Dahabreh IJ, Kent DM. Index event bias as an explanation for the paradoxes of recurrence risk research. *JAMA : the journal of the American Medical Association.* Feb 23 2011;305(8):822-823.
127. Hernandez-Diaz S, Wilcox AJ, Schisterman EF, Hernan MA. From causal diagrams to birth weight-specific curves of infant mortality. *European journal of epidemiology.* 2008;23(3):163-166.

128. Hernandez-Diaz S, Schisterman EF, Hernan MA. The birth weight "paradox" uncovered? *Am J Epidemiol.* Dec 1 2006;164(11):1115-1120.
129. Baglin T. Unraveling the thrombophilia paradox: from hypercoagulability to the prothrombotic state. *Journal of thrombosis and haemostasis : JTH.* Feb 2010;8(2):228-233.
130. Rich JD, Cannon CP, Murphy SA, Qin J, Giugliano RP, Braunwald E. Prior aspirin use and outcomes in acute coronary syndromes. *Journal of the American College of Cardiology.* Oct 19 2010;56(17):1376-1385.
131. Banack HR, Kaufman JS. The obesity paradox: Understanding the effect of obesity on mortality among individuals with cardiovascular disease. *Preventive medicine.* Feb 10 2014;62C:96-102.
132. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology (Cambridge, Mass.).* Jul 2009;20(4):488-495.
133. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology (Cambridge, Mass.).* Sep 2004;15(5):615-625.
134. Rothman KJ GS, Lash T. *Modern Epidemiology.* 3rd Edition ed. Philadelphia, PA: Lipincott Williams & Wilkins; 2008
135. Pearl J. Causal diagrams for empirical research. *Biometrika.* 1995;82:669-710.
136. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology (Cambridge, Mass.).* Jan 1999;10(1):37-48.
137. Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology (Cambridge, Mass.).* May 2003;14(3):300-306.
138. Smits LJM, van Kuijk SMJ, Leffers P, Peeters LL, Prins MH, Sep SJS. Index event bias—a numerical example. *Journal of Clinical Epidemiology.* 2// 2013;66(2):192-196.
139. Whitcomb BW, Schisterman EF, Perkins NJ, Platt RW. Quantification of collider-stratification bias and the birthweight paradox. *Paediatric and Perinatal Epidemiology.* 2009;23(5):394-402.
140. Horton JD, Cohen JC, Hobbs HH. Molecular biology of PCSK9: its role in LDL metabolism. *Trends in biochemical sciences.* Feb 2007;32(2):71-77.
141. Sullivan D, Olsson AG, Scott R, et al. Effect of a monoclonal antibody to pcsk9 on low-density lipoprotein cholesterol levels in statin-intolerant patients: The gauss randomized trial. *JAMA : the journal of the American Medical Association.* 2012;308(23):2497-2506.
142. Cohen JC, Boerwinkle E, Mosley TH, Jr., Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *The New England journal of medicine.* Mar 23 2006;354(12):1264-1272.

143. Fewell Z, Davey Smith G, Sterne JA. The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study. *Am J Epidemiol.* Sep 15 2007;166(6):646-655.
144. Towfighi A, Zheng L, Ovbiagele B. Sex-specific trends in midlife coronary heart disease risk and prevalence. *Archives of internal medicine.* Oct 26 2009;169(19):1762-1766.
145. Cole JH, Miller JI, 3rd, Sperling LS, Weintraub WS. Long-term follow-up of coronary artery disease presenting in young adults. *Journal of the American College of Cardiology.* Feb 19 2003;41(4):521-528.
146. Berenson GS, Srinivasan SR, Bao W, Newman WP, 3rd, Tracy RE, Wattigney WA. Association between multiple cardiovascular risk factors and atherosclerosis in children and young adults. The Bogalusa Heart Study. *The New England journal of medicine.* Jun 4 1998;338(23):1650-1656.
147. Iribarren C, Jacobs DR, Jr., Slattery ML, et al. Epidemiology of low total plasma cholesterol concentration among young adults: the CARDIA study. Coronary Artery Risk Development in Young Adults. *Preventive medicine.* Jul-Aug 1997;26(4):495-507.
148. Zimmerman FH, Cameron A, Fisher LD, Ng G. Myocardial infarction in young adults: angiographic characterization, risk factors and prognosis (Coronary Artery Surgery Study Registry). *Journal of the American College of Cardiology.* Sep 1995;26(3):654-661.
149. Chen L, Chester M, Kaski JC. Clinical factors and angiographic features associated with premature coronary artery disease. *Chest.* Aug 1995;108(2):364-369.
150. Barbash GI, White HD, Modan M, et al. Acute myocardial infarction in the young--the role of smoking. The Investigators of the International Tissue Plasminogen Activator/Streptokinase Mortality Trial. *European heart journal.* Mar 1995;16(3):313-316.
151. Malmberg K, Bavenholm P, Hamsten A. Clinical and biochemical factors associated with prognosis after myocardial infarction at a young age. *Journal of the American College of Cardiology.* Sep 1994;24(3):592-599.
152. Hoit BD, Gilpin EA, Henning H, et al. Myocardial infarction in young patients: an analysis by age subsets. *Circulation.* Oct 1986;74(4):712-721.
153. Rosenberg L, Kaufman DW, Helmrich SP, Miller DR, Stolley PD, Shapiro S. Myocardial infarction and cigarette smoking in women younger than 50 years of age. *JAMA : the journal of the American Medical Association.* May 24-31 1985;253(20):2965-2969.
154. Bähler C, Gutzwiller F, Erne P, Radovanovic D. Lower age at first myocardial infarction in female compared to male smokers. *European Journal of Preventive Cardiology.* October 1, 2012 2012;19(5):1184-1193.

155. Grundtvig M, Hagen TP, German M, Reikvam Å. Sex-based differences in premature first myocardial infarction caused by smoking: twice as many years lost by women as by men. *European Journal of Cardiovascular Prevention & Rehabilitation.* April 1, 2009 2009;16(2):174-179.
156. Madala MC, Franklin BA, Chen AY, et al. Obesity and age of first non-ST-segment elevation myocardial infarction. *Journal of the American College of Cardiology.* Sep 16 2008;52(12):979-985.
157. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet.* Apr 2011;43(4):339-344.
158. Ellis KL, Pilbrow AP, Frampton CM, et al. A common variant at chromosome 9P21.3 is associated with age of onset of coronary disease but not subsequent mortality. *Circulation. Cardiovascular genetics.* Jun 2010;3(3):286-293.
159. Meng W, Hughes AE, Patterson CC, Belton C, Kee F, McKeown PP. Chromosome 9p21.3 is associated with early-onset coronary heart disease in the Irish population. *Disease markers.* 2008;25(2):81-85.
160. Palomaki GE, Melillo S, Bradley LA. Association between 9p21 genomic markers and heart disease: a meta-analysis. *JAMA : the journal of the American Medical Association.* Feb 17 2010;303(7):648-656.
161. Tikkanen E, Havulinna AS, Palotie A, Salomaa V, Ripatti S. Genetic Risk Prediction and a 2-Stage Risk Screening Strategy for Coronary Heart Disease. *Arteriosclerosis, thrombosis, and vascular biology.* April 18, 2013 2013;33:2261-2266.
162. Dandona S, Stewart AF, Chen L, et al. Gene dosage of the common variant 9p21 predicts severity of coronary artery disease. *Journal of the American College of Cardiology.* Aug 3 2010;56(6):479-486.
163. Liao Y-C, Lin H-F, Rundek T, et al. Segment-Specific Genetic Effects on Carotid Intima-Media Thickness: The Northern Manhattan Study. *Stroke; a journal of cerebral circulation.* December 1, 2008 2008;39(12):3159-3165.
164. Raffield LM, Cox AJ, Hsu FC, et al. Impact of HDL genetic risk scores on coronary artery calcified plaque and mortality in individuals with type 2 diabetes from the Diabetes Heart Study. *Cardiovascular diabetology.* 2013;12:95.
165. Tsao CW, Preis SR, Peloso GM, et al. Relations of long-term and contemporary lipid levels and lipid genetic risk scores with coronary artery calcium in the framingham heart study. *Journal of the American College of Cardiology.* Dec 11 2012;60(23):2364-2371.

166. Patel RS, Sun YV, Hartiala J, et al. Association of a genetic risk score with prevalent and incident myocardial infarction in subjects undergoing coronary angiography. *Circulation. Cardiovascular genetics.* Aug 1 2012;5(4):441-449.
167. Bao W, Srinivasan SR, Wattigney WA, Berenson GS. The relation of parental cardiovascular disease to risk factors in children and young adults. The Bogalusa Heart Study. *Circulation.* Jan 15 1995;91(2):365-371.
168. Writing Group for the Women's Health Initiative I. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: Principal results from the women's health initiative randomized controlled trial. *JAMA : the journal of the American Medical Association.* 2002;288(3):321-333.
169. Musunuru K, Post WS, Herzog W, et al. Association of single nucleotide polymorphisms on chromosome 9p21.3 with platelet reactivity: a potential mechanism for increased vascular disease. *Circulation. Cardiovascular genetics.* Oct 2010;3(5):445-453.
170. Vaidya D, Yanek LR, Moy TF, Pearson TA, Becker LC, Becker DM. Incidence of coronary artery disease in siblings of patients with premature coronary artery disease: 10 years of follow-up. *The American journal of cardiology.* Nov 1 2007;100(9):1410-1415.
171. Wieser S, Ruthemann I, De Boni S, et al. Cost of acute coronary syndrome in Switzerland in 2008. *Swiss medical weekly.* 2012;142:w13655.

Apppendix A: SNPs used in the GRS

30 SNPs included in MI GRS

SNP	Gene	Chromosome	Susceptibility Allele	Included 8 SNP LDL GRS
rs10953541	7q22/BCAP29	7	C	
rs11206510	PCSK9/BSND	1	T	YES
rs1122608	LDLR/SMARCA4	19	G	YES
rs11556924	ZC3HC1	7	G	
rs12190287	TCF21	6	C	
rs12413409	CYP17A1/CNNM2/NT5C2	10	C	
rs12526453	PHACTR1	6	C	
rs12936587	PEMT	17	G	
rs1412444	LIPA	10	T	YES
rs17114036	PPAP2B	1	T	
rs1746048	CXCL12/HNRNPA3P1	10	C	
rs17465637	MIA3	1	G	
rs17609940	ANKS1A	6	G	
rs216172	SMG6-SRR	17	C	
rs2259816	HNF1A	12	A	YES
rs2505083	KIAA1462	10	G	
rs2895811	HHIPL1	14	C	
rs3184504	SH2B3	12	T	
rs3798220	LPA	6	C	YES
rs3825807	ADAMTS7	15	T	
rs46522	UBE2Z/GIP	17	T	
rs4773144	COL4A1-A2	13	C	
rs4977574	9p21	9	G	
rs579459	ABO	9	C	YES
rs646776	CELSR2/PSRC1-SORT1	1	T	YES
rs6725887	WDR12	2	C	
rs964184	APOA5/APOC3/ZNF259	11	C	YES
rs974819	PDGFD	11	A	
rs9818870	MRAS	3	T	
rs9982601	SLC5A3/MRPS6/KCNE2/C21orf82	21	A	

Apppendix B: Test of the Proportional Hazards Assumption

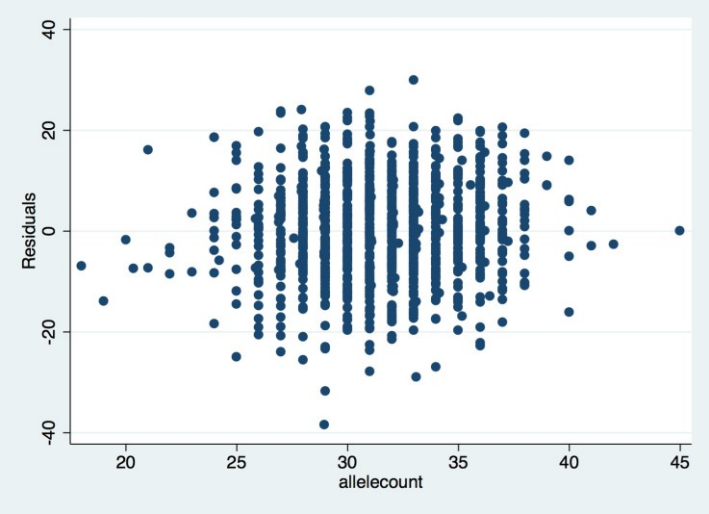
```
. estat phtest, detail

      Test of proportional-hazards assumption

      Time:  Time
      ----------------------------------------------------------------
                  |       rho            chi2       df       Prob>chi2
      ------------+---------------------------------------------------
      allelecount |     -0.17541          1.65        1          0.1990
      ecg_stelev  |      0.04659          0.16        1          0.6892
      age         |     -0.00658          0.00        1          0.9530
      0.gender    |     -0.08885          0.51        1          0.4748
      1b.gender   |            .             .        1               .
      prior_cvd   |      0.05289          0.20        1          0.6563
      prior_revasc|     -0.10436          0.85        1          0.3568
      hx_hta      |     -0.00083          0.00        1          0.9945
      hx_diabetes |      0.13517          1.27        1          0.2602
      hx_lipid    |      0.07292          0.37        1          0.5452
      bmi         |     -0.32977          7.24        1          0.0071
      currentsmo~r|     -0.00095          0.00        1          0.9934
      med_asa     |      0.03378          0.09        1          0.7668
      med_othplat |      0.01709          0.02        1          0.8831
      med_bblock  |     -0.16192          2.07        1          0.1504
      med_lipid   |     -0.11814          0.91        1          0.3405
      med_aceorarb|      0.06491          0.29        1          0.5894
      ------------+---------------------------------------------------
      global test |                      15.15       16          0.5134
      ----------------------------------------------------------------
```

The p-value is large p=0.5134 for the global test, so there is no evidence for interaction between the covariates and time, which indicates that the proportional hazards assumption is satisfied.

Even the individual level covariates do not show any evidence for violations of the proportional hazards assumption.

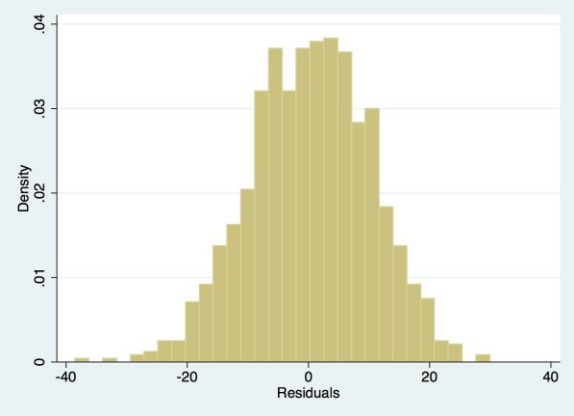
Apppendix C: Test of the Appropriateness of the Linear Model

1. Linear models assume that the variance of the error is constant and does not depend on any covariate. This can be assessed by plotting the residuals against the covariate in question. In this situation, plotting the residuals of the regression model of age of first ACS yields the following plot:



There is clearly no pattern here, and therefore we can conclude that the assumption was met.

2. Linear models also assume that residuals are normally distributed. This can be assessed visually with a histogram. The residuals appear to be roughly normally distributed and we can assume that this assumption was met as well.



There does not appear to be any major violation of the linear model assumptions present and we can assume that the linear model is valid in this circumstance.

Apppendix D: Discrimination and Calibration Indices for the GRS Added to the GRACE Risk Model

Multiple tests were considered to assess to the predictive capacity of the GRS when added to a standard risk score such as the GRACE risk score. Recently, Pencina et al.[94] suggested that the newer metrics such as IDI and NRI rather than area under the curve should be used to assess the value of a new marker over and above the predictive ability of standard risk factors or models.

In our data, the GRS offered no predictive advantage over the GRACE risk score regardless of the measure used.

	RISCA		PRAXY	
	Test Statistic	P-value	Test Statistic	P-value
Calibration				
Likelihood ratio test	LR χ^2 1.46 $_{df=1}$	0.2267	LR χ^2 0.16 $_{df=1}$	0.6852
Discrimination				
Harrell's C-statistic GRS+GRACE vs. GRACE alone	0.68 vs 0.66	0.876	0.54 vs 0.52	0.514
IDI	0.0017	0.4464	0.0021	0.2471
NRI: Category free	0.1834	0.1552	0.1718	0.1209

GRS: genetic risk score, IDI: integrated discrimination improvement, NRI: net reclassification improvement