Arteriosclerosis, Thrombosis, and Vascular Biology

JOURNAL OF THE AMERICAN HEART ASSOCIATION

Multilocus Genetic Risk Scores for Coronary Heart Disease Prediction

Andrea Ganna, Patrik K.E. Magnusson, Nancy L. Pedersen, Ulf de Faire, Marie Reilly, Johan Ärnlöv, Johan Sundström, Anders Hamsten and Erik Ingelsson

Arterioscler Thromb Vasc Biol. 2013;33:2267-2272; originally published online May 16, 2013;
doi: 10.1161/ATVBAHA.113.301218
Arteriosclerosis, Thrombosis, and Vascular Biology is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 1079-5642. Online ISSN: 1524-4636

Multilocus Genetic Risk Scores for Coronary Heart Disease Prediction

Andrea Ganna, Patrik K.E. Magnusson, Nancy L. Pedersen, Ulf de Faire, Marie Reilly, Johan Ärnlöv, Johan Sundström, Anders Hamsten, Erik Ingelsson

Objective—Current guidelines do not support the use of genetic profiles in risk assessment of coronary heart disease (CHD). However, new single nucleotide polymorphisms associated with CHD and intermediate cardiovascular traits have recently been discovered. We aimed to compare several multilocus genetic risk score (MGRS) in terms of association with CHD and to evaluate clinical use.

Approach and Results—We investigated 6 Swedish prospective cohort studies with 10612 participants free of CHD at baseline. We developed 1 overall MGRS based on 395 single nucleotide polymorphisms reported as being associated with cardiovascular traits, 1 CHD-specific MGRS, including 46 single nucleotide polymorphisms, and 6 trait-specific MGRS for each established CHD risk factors. Both the overall and the CHD-specific MGRS were significantly associated with CHD risk (781 incident events; hazard ratios for fourth versus first quartile, 1.54 and 1.52; $P<0.001$) and improved risk classification beyond established risk factors (net reclassification improvement, 4.2% and 4.9%; $P=0.006$ and 0.017). Discrimination improvement was modest (C-index improvement, 0.004). A polygene MGRS performed worse than the CHD-specific MGRS. We estimate that 1 additional CHD event for every 318 people screened at intermediate risk could be saved by measuring the CHD-specific genetic score in addition to the established risk factors.

Conclusions—Our results indicate that genetic information could be of some clinical value for prediction of CHD, although further studies are needed to address aspects, such as feasibility, ethics, and cost efficiency of genetic profiling in the primary prevention setting. **(*Arterioscler Thromb Vasc Biol*. 2013;33:2267-2272.)**

Key Words: association studies ■ genetics ■ risk assessment ■ risk prediction ■ risk score

Since the dissemination of the HapMap project in 2005,[1] genetic researchers have performed a large number of genome-wide association studies to identify genetic determinants of complex diseases. In the cardiovascular field, recent examples include the publication from the CARDIoGRAMplusC4D consortium, which increased the number of loci robustly associated with coronary heart disease (CHD) in individuals of Northern European descent to 46,[2] and studies of related traits, such as lipid fractions[3] or fasting glucose and insulin.[4] The improved biological understanding of these traits has yet to be followed by clinical applications of the discoveries. Although direct-to-consumer tests of recently discovered genetic markers are already available on the market, especially in the United States, both US[5] and European[6] guidelines for prevention of cardiovascular diseases in clinical practice advise against prognostic use of DNA-based tests in the primary prevention setting.

See accompanying article on page 2049

Several studies have investigated whether the introduction of a multilocus genetic risk score (MGRS) in addition to an established CHD risk algorithm, such as the Framingham Heart Study (FHS) risk score,[7] would improve disease prediction. Using 13 CHD-associated single nucleotide polymorphisms (SNPs), Ripatti et al[8] showed that an MGRS was highly associated with CHD, but was unable to improve the risk classification beyond what was achieved with traditional risk factors. Paynter et al[9] considered a larger genetic score (101 SNPs), including SNPs associated with cardiovascular disease–related phenotypes, and failed to demonstrate a significant association with CHD. However, to the best of our knowledge, no previous studies have examined the use of an MGRS, including the most recently reported CHD-related SNPs. Furthermore, an alternative scoring approach has been proposed, which includes all SNPs associated with the outcome in an external population up to a certain probability value, instead of using only published genome-wide significant SNPs. This so-called polygene approach, which has shown

Received on: January 21, 2013; final version accepted on: May 5, 2013.

From the Department of Medical Epidemiology and Biostatistics (A.G., P.K.E.M., N.L.P., M.R.), Division of Cardiovascular Epidemiology, Institute of Environmental Medicine (U.d.F.), Cardiovascular Genetics and Genomics Group, Atherosclerosis Research Unit, Department of Medicine Solna (A.H.), Karolinska Institutet, Stockholm, Sweden; Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory (A.G., J.Ä., E.I.) and Department of Medical Sciences, Cardiovascular Epidemiology (J.S.), Uppsala University, Uppsala, Sweden; School of Health and Social Studies, Dalarna University, Falun, Sweden (J.Ä.); and Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom (E.I.).

The online-only Data Supplement is available with this article at http://atvb.ahajournals.org/lookup/suppl/doi:10.1161/ATVBAHA.113.301218/-/DC1.

Correspondence to Erik Ingelsson, MD, PhD, FAHA, Department of Medical Sciences, Molecular Epidemiology, Uppsala University Hospital, SE-751 85 Uppsala, Sweden. E-mail erik.ingelsson@medsci.uu.se

Arterioscler Thromb Vasc Biol **is available at http://atvb.ahajournals.org** **DOI: 10.1161/ATVBAHA.113.301218**

promising results for other traits in cross-sectional studies,[10] has so far not been tested for CHD prediction.

In the present study, including 6 longitudinal, population-based Swedish cohorts of >10 000 participants, our primary aim was to compare an overall MGRS and a CHD-specific MGRS in terms of disease association and clinical use. Our secondary aims were to investigate the clinical use of a polygene MGRS and the associations of trait-specific MGRS with CHD and established cardiovascular risk factors.

Materials and Methods

Materials and Methods are available in the online-only Supplement.

Results

Genetic Score Generation Procedures

In Figure 1A, we have schematically outlined the process for selection of SNPs to be included in the MGRS from the National Human Genome Research Institute GWA study catalog.[11] Briefly, 119 traits that are related to CHD in a direct or more broad sense (Table I in the online-only Data Supplement) were selected from the National Human Genome Research Institute catalog by a medical doctor (E.I.) with expertise in cardiovascular epidemiology that was blinded to the results of the association analyses in our data. All genome-wide association studies investigating these traits underwent an accurate quality control to assess the validity of the SNPs reported in the catalog. After quality control, 615 single SNPs from 92 studies investigating 49 traits were used to calculate the scores. We eliminate correlated SNPs in high linkage disequilibrium using the pruning technique as implemented in PLINK,[12] and we created 3 weighted MGRS using the log(odds ratio) for CHD association in Wellcome Trust Case Control Consortium[13] (score 1 and 3) or in a published article[2] (score 2) to calculate weights and 7 unweighted trait-specific MGRS.

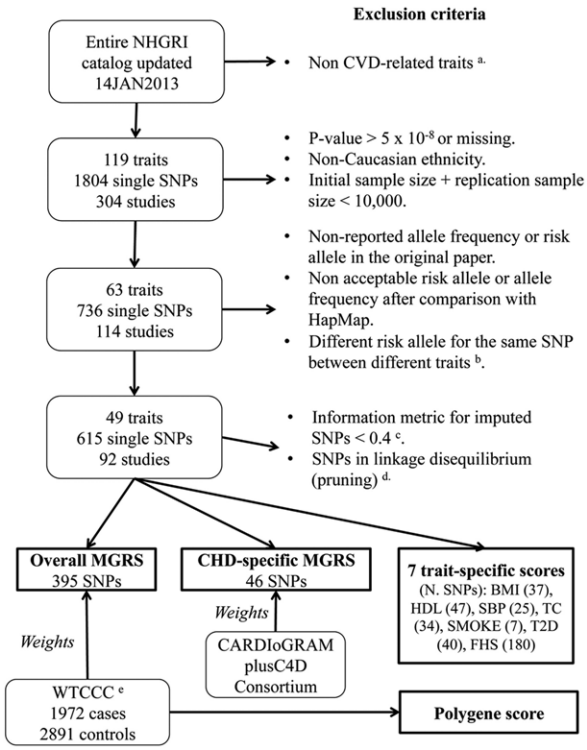


Figure. Outline of the procedure for selecting multilocus genetic risk score (MGRS) single nucleotide polymorphisms (SNPs) from the National Human Genome Research Institute (NHGRI) Catalog. [a]Cardiovascular disease-related traits were defined a priori from traits reported in the catalog (see Table I in the online-only Supplement for a full list). [b]In case of contrasting risk alleles for different traits, the risk allele for the trait that was most strongly associated with outcome was kept. [c]For SNPs with low imputation quality in CardioMetabochip, we set the number of alleles equal to the average number of risk alleles in TwinGene (to keep the SNPs in the score). [d]SNPs within 250 kb with R^2 higher than 0.2 were LD-pruned based on HapMap CEU rel22. Within each cluster, the SNPs with the lowest *P* value for association with CHD in the Wellcome Trust Case Control Consortium (WTCCC) data were selected. [e]WTCCC underwent quality control and was imputed using HapMap CEU release 22 as reference panel.

1. An overall MGRS obtained from 395 SNPs associated with CHD or CHD-related traits (Tables I and II in the online-only Data Supplement);
2. A CHD-specific MGRS obtained from 46 SNPs that have been found associated with CHD by the CARDIoGRAMplusC4D consortium, which is the largest genome-wide association studies meta-analysis on CHD to date[2];
3. A polygene MGRS obtained as the weighted sum of risk alleles for SNPs with $P<0.2$ in the Wellcome Trust Case Control Consortium study. This threshold corresponds to the highest increment in the C-statistic over a basic model, as shown graphically in Figure I in the online-only Data Supplement;
4. An unweighted trait-specific MGRS for each of 6 established cardiovascular risk factors: body mass index (BMI; 37 SNPs), high-density lipoprotein (HDL)-cholesterol (47 SNPs), systolic blood pressure (25 SNPs), total cholesterol (TC; 34 SNPs), smoking (7 SNPs), type 2 diabetes mellitus (40 SNPs), and 1 FHS MGRS, including 180 SNPs associated with any of the aforementioned traits.

Descriptive Statistics

A total of 10 612 participants, free of CHD at baseline and with complete information on FHS risk score components and BMI, were included in our study. A total of 781 CHD events (539 in men and 242 in women) were observed during a median follow-up of 4.3 years (interquartile range, 3.6–5.8). In Table 1, we report the descriptive statistics of our sample, together with the associations of the FHS risk score components and BMI with incident CHD.

Association With CHD and Risk Factors

In a single SNP analysis (Table I in the online-only Data Supplement), *MECOM* rs419076 and *IGFBP3* rs7784776 showed the strongest associations with CHD after adjustment for FHS risk factors. However, the associations were not significant after multiple-testing correction in the univariate analysis. Both the overall MGRS and the CHD-specific MGRS were highly significantly associated with CHD. Among the risk factors, the CHD-specific MGRS was associated only with TC; interestingly, no association was detected with the FHS. Each trait-specific MGRS was significantly associated with the corresponding trait, except for the smoking MGRS (Table 2). The FHS MGRS was associated with all the risk factors for CHD, except for smoking, but not directly with CHD. The HDL-cholesterol MGRS was strongly associated with BMI and TC. No association with CHD

Table 1. Baseline Descriptive Statistics and Associations With CHD for FHS Risk Factors and BMI

Characteristics	Descriptive Statistics		HR for Association with CHD†
	Men	Women	
Female sex, n (%)	5251 (50)	5361 (50)	0.46*
Age in years, mean (SD)	68 (8)	67 (10)	
Prevalent type 2 diabetes mellitus, n (%)	484 (9)	340 (6)	1.90*
Current smokers, n (%)	894 (17)	894 (17)	1.46*
Use of antihypertensive drugs, n (%)	1164 (22)	1178 (22)	1.47*
Systolic blood pressure, mm Hg, mean (SD)	144 (21)	143 (23)	1.01*
Total cholesterol, mmol/L, mean (SD)	5.7 (1.1)	6.2 (1.1)	1.28*
High-density lipoprotein-cholesterol, mmol/L, mean (SD)	1.3 (0.3)	1.6 (0.4)	0.43*
BMI, kg/m^2, mean (SD)	26.1 (3.5)	25.6 (4.3)	0.99

BMI indicates body mass index; CHD indicates coronary heart disease; FHS, Framingham Heart Study; and HR, hazard ratio.

*P<0.001.

†All estimates reported are from Cox proportional hazard analyses with age as timescale and adjusted for sex, systolic blood pressure, antihypertensive treatment, diabetes mellitus, current smoking, total cholesterol, body mass index, and study of origin. HRs are given for unit increase for continuous traits.

was found for the polygene MGRS. Once the SNPs included in the CHD-specific score were removed from the overall MGRS, this score was still significantly associated with CHD (hazard ratio, 1.091; P=0.008).

Reclassification, Discrimination, and Calibration

Participants who were in the upper quartile of the distribution of the overall MGRS had 1.54× increased risk for CHD compared with individuals in the lowest quartile (P<0.001; Table 3). The overall MGRS significantly improved risk classification beyond established FHS risk factors (net reclassification improvement [NRI], 4.2%; P=0.006), but the discrimination improvement was modest (C-index improvement, 0.002). Higher reclassification and discrimination were observed for the CHD-specific MGRS (NRI, 4.9%; C-index improvement, 0.004). Participants' distribution within risk categories and reclassification after addition of the CHD-specific MGRS are reported in Figure II in the online-only Data Supplement. Intermediate- and high-risk subjects evidenced

Table 2. *P* Values for Association Between MGRS and CHD or Main CHD Risk Factors

	No. of SNPs*	Outcomes							
		CHD§	BMI‖	HDL‖	SBP‖	SMOKE¶	T2D¶	TC‖	FHS‖,#
Overall†	395	2×10^{-5}**	0.13	5×10^{-16}**	1×10^{-3}††	0.85	0.10	7×10^{-10}**	3×10^{-4}**
CHD-specific‡	46	4×10^{-6}**	0.48	0.21	0.11	0.76	0.57	1×10^{-2}††	0.11
Polygene (P<0.2)†	Metabochip≈14 162 HumanOmniExpress ≈117 482	0.17	0.79	0.08	0.69	0.27	0.11	0.09	0.28
Trait-specific									
BMI	37	0.35	4×10^{-18}**	2×10^{-4}**	0.45	3×10^{-2}††	0.26	3×10^{-2}††	0.62
HDL	47	0.88	0.81	4×10^{-90}**	1.00	0.13	0.22	3×10^{-4}**	1×10^{-2}††
SBP	25	0.54	0.60	0.51	3×10^{-9}**	0.42	4×10^{-2}††	0.23	0.16
Smoking	7	0.05	0.47	0.21	0.64	0.10	0.72	0.09	0.61
T2D	40	0.75	0.23	2×10^{-2}††	0.29	0.23	7×10^{-11}**	0.37	5×10^{-4}††
TC	34	0.24	0.71	1×10^{-4}**	0.49	0.18	0.54	1×10^{-51}**	2×10^{-3}††
FHS	180	0.43	1×10^{-2}††	1×10^{-22}**	6×10^{-4}††	0.88	2×10^{-5}**	1×10^{-3}††	5×10^{-6}**

CHD indicates coronary heart disease; FHS, Framingham Heart Study; MGRS, multilocus genetic risk score; and T2D, type 2 diabetes mellitus.

*Number of single nucleotide polymorphisms (SNPs) selected after quality controls and pruning.

†Weighted with log(odds ratio) from Wellcome Trust Case Control Consortium (WTCCC) study.

‡Weighted with log(odds ratio) from CARDIoGRAMplusC4D consortium.

§All P values reported are from Cox proportional hazard analyses with age as timescale and adjusted for sex, systolic blood pressure (SBP), antihypertensive treatment, diabetes mellitus, current smoking, total cholesterol (TC), body mass index (BMI), and study of origin.

‖All P values reported are from a linear regression model adjusted for age, sex, and study of origin. In high-density lipoprotein (HDL) and TC analyses, we adjusted also for lipid-lowering treatment.

¶All P values reported are from a logistic regression model adjusted for age, sex, and study of origin.

#Individual risk at 10 y obtained from Cox proportional hazard analyses adjusted for age, sex, SBP, antihypertensive treatment, diabetes mellitus, current smoking, TC, and study of origin. TwinGene study was not included in these analyses because of short follow-up.

**P<5×10^{-4}, ††P<5×10^{-2}.

Table 3. Association and Prediction Measures for Main MGRS and CHD

		Outcome: CHD							
		Association (HR, 95 CI)§	Reclassification‖			Discrimination‖			Calibration¶
	No. of SNPs*	Fourth vs First Quartile	Among Events, %	Among Nonevents, %	Total (NRI % [95 CI])	Base Model C-Index	C-Index Increment	Integrated Discrimination Improvement, % [95 CI]	Gronnesby–Borgan *P* Value
Overall†	395	1.54 [1.25; 1.92]	2.9	1.3	4.2 [1.2; 7.1]		0.002	0.2 [–0.1; 0.5]	0.96
CHD-specific‡	46	1.52 [1.24; 1.87]	2.8	2.1	4.9 [1.1; 8.7]	0.702	0.004	0.4 [0.1; 0.7]	0.99
Polygene (P<0.2)†	Metabochip≈14 162 HumanOmniExpress ≈117 482	1.24 [0.82; 1.87]	1.3	0.0	1.3 [–0.3; 2.9]		0.001	0.1 [–0.1; 0.1]	0.73

CHD indicates coronary heart disease; CI, confidence interval; HR, hazard ratio; MGRS, multilocus genetic risk score; and NRI, net reclassification improvement.

*Number of single nucleotide polymorphisms (SNPs) selected after quality controls and pruning.

†Weighted with log(odds ratio) from Wellcome Trust Case Control Consortium study.

‡Weighted with log(odds ratio) from CARDIoGRAMplusC4D consortium.

§Cox proportional hazard analyses with age as timescale and adjusted for sex, systolic blood pressure, antihypertensive treatment, diabetes mellitus, current smoking, total cholesterol, body mass index, and study of origin.

‖Prediction measures calculated from individual risk of CHD at 10 y obtained from Cox proportional hazard analyses adjusted for age, sex, systolic blood pressure, antihypertensive treatment, diabetes mellitus, current smoking, total cholesterol, and study of origin. TwinGene study was not included in these analyses because of short follow-up.

¶Nonsignificant *P* values indicate good model calibration.

the largest changes in individual risk, when the CHD-specific MGRS was added to the FHS risk factors (Figure III in the online-only Data Supplement). Good calibration was observed for all scores (Table 3). The TC MGRS had the highest R^2, explaining 6.2% of the total variance for this phenotype in our sample (Figure IV in the online-only Data Supplement). When calculating the discriminatory abilities of the CHD-specific MGRS without adjusting for established risk factors, we obtained a C-index of 0.54.

Number of Events and Event-Free Life Years Prevented

Among the 1272 (Figure II in the online-only Data Supplement) participants with a 10-year risk between 10% and 20%, 83 (6.5%) participants would have been reclassified as high risk (>20%), when adding the CHD-specific score and, therefore, would have been eligible for statins treatment according to the adult treatment panel-III guidelines.[14] Twenty of these participants experienced an event during the 10 years follow-up. That is, assuming a risk reduction of 20% for individuals treated with statins, the targeted assessment of the genetic risk score among intermediate risk subjects could help to prevent ≈4 (ie, 0.20×20) additional CHD events during a 10-year period, which corresponds to 1 avoided event for every 318 people screened (ie, 1272/4; see online-only Data Supplement for details).

Furthermore, if the CHD-specific genetic score was measured on the entire study population in addition to the established risk factors, 3.15 [confidence interval, −1.30; 7.60] event-free life years per 1000 people screened would have been saved (Table III in the online-only Supplement).

Discussion

In the present longitudinal study of 10 612 participants, we investigated 10 MGRS for association with CHD and established cardiovascular risk factors. Three scores were further investigated in terms of clinical use (ie, the ability to discriminate between individuals with and without the disease and to correctly classify them in clinical categories of risk). We found that both a large comprehensive genetic score and a CHD-specific score were able to significantly improve the correct classification of study participants. In addition, we showed that a polygene MGRS, developed in an external data set, including all SNPs up to a nonsignificant *P* value threshold, did not outperform a literature-based score. Finally, we found that the genetic risk score for HDL-cholesterol was associated both with BMI and TC, potentially indicating pleiotropic effects of HDL-associated loci. The association between the CHD-specific MGRS and the 10-year predicted risk calculated from established risk factors (FHS) was not significant. Indeed, only 14 of 44 CHD-specific loci have been observed in a previous large study[2] to be associated with lipid-related traits, type 2 diabetes mellitus, or blood pressure. Similarly, in an older study,[15] only 1 of 14 CHD-associated SNPs was associated with a wide range of established risk factors and novel biomarkers. These previous results, together with our findings, support the hypothesis that most of the CHD loci are not involved in pathways perturbing currently known risk factors, making a genetic score of CHD-specific loci a good candidate for future improvements of current prediction algorithms, and CHD loci an interesting starting point for further studies into the pathophysiology of atherosclerosis.

Our results indicate that use of genetic profiling, including SNPs associated with intermediate traits, might improve allocation of patients to correct risk strata. Although we calculated prediction measures in a relatively small- and high-risk subsample of the study, in which we had long enough follow-up time (n=3014), the NRI (4.9%; *P*=0.01) was significant. This proportion of correct reclassification is clinically relevant, and is larger than that observed for HDL-cholesterol in our study (NRI, 2.7%) and for C-reactive protein and fibrinogen

in a recent meta-analysis (NRI, 1.52% and 0.83%, respectively).[16] However, caution should be exercised when comparing estimates from a large meta-analysis with that observed in a single study. In a hypothetical population similar to that investigated in this study, the routine use of genetic profiling in primary prevention of CHD among individuals at intermediate risk could prevent 1 event for every 300 to 400 people screened. If the risk assessment with genetic profiling is performed on the entire population, and not only on individual at intermediate risk, 3.15 [−1.30; 7.60] years free of CHD events could be saved per 1000 people undergoing the risk assessment. Clinicians and health policy makers should evaluate whether these benefits outbalance the costs of implementing such testing in clinical practice.

Genetic profiling has some advantages compared with other biomarkers that are routinely used in clinical practice. For example, genetic information remains stable throughout life and is, therefore, not sensitive to regression dilution bias. Hence, although the strength of the genotype–phenotype associations has been suggested to be modified by age,[17] genetic markers are likely to be predictive throughout life. This might allow risk prediction to be performed much earlier in life, to allow for earlier primary prevention in high-risk individuals. However, to investigate this aspect, study populations with younger participants and longer follow-up are needed than what was available for the present study. Another potential advantage is that no invasive blood sampling is needed, because DNA is routinely extracted from saliva. However, several important obstacles remain to be considered before genetic profiling can be introduced in routine healthcare. These include demonstration of usefulness and cost-effectiveness for risk prediction in several independent samples and the fact that genetic markers, in contrast to modifiable risk markers, are not useful for assessing efficacy of treatment or other risk-reduction strategies.

The alternative scoring approach (polygene approach), which includes all SNPs associated with the outcome in an external population up to a certain probability value, has shown promising results for other traits and has been proposed to increase the variance explained for common complex diseases.[18] For the first time, we applied this approach to CHD. In our study, this score was not associated with CHD, probably because of the relatively small data set used for generating the score, resulting in lower accuracy of the weights. Theoretical frameworks have predicted polygenic scores to substantially increase the discriminatory ability of the model, with sample size larger than those currently available from meta-analysis consortia.[19,20] Therefore, future studies using larger development data sets are needed to understand whether a large number of SNPs with small effect size not exceeding the genome-wide threshold could be able to improve the clinical use beyond established genetic variants. The polygene score and the overall MGRS embrace a similar approach; they include a large number of SNPs, both true signals and noise, trying to capture the polygene architecture of CHD. The SNP selection strategy for these 2 scores was, however, different. The polygene score uses a probability value threshold for SNP selection and, therefore, the predictive power is much dependent on the sample size of the study where the selection was made. In contrast, the overall MGRS uses a priori information from the literature and, therefore, the prediction ability is determined by the quality of the literature selection. The overall MGRS was significantly associated with CHD, even after the exclusion of CHD-specific SNPs, indicating that a selection strategy based on literature-based information can be useful to improve the prediction ability of genetic risk scores. Moreover, we performed 2 sensitivity analyses to investigate whether the inclusion of a large number of SNPs in the overall MGRS might have caused a spurious association with CHD. First, we plotted the association results across the 6 investigated studies to address consistency (Figure V in the online-only Supplement). Second, we created 100 random overall MGRS and studied their association with CHD in TwinGene (Figure VI in the online-only Supplement). The association was consistent across studies and almost identical in the 2 larger studies. The average association of the 100 random scores was significantly lower (P=0.01) than the associations observed in the original score. These analyses suggest that the association between the overall MGRS and CHD is unlikely to be explained by chance.

Previous studies did not find a significant association between an overall MGRS and a CHD.[21,22] A smaller number of traits included in the score, different outcomes definitions, and heterogeneity between study populations might explain such differences. In addition, strengths of our study include the large number of incident CHD events investigated, as well as the up-to-date literature-based MGRS, including genetic variants from the most recent studies. Moreover, we used information from an external study population to assign weights in our scoring. This is in contrast to previous studies that have used association coefficients reported in the literature[8]; an approach that biases the score attributable to the winner's curse. We also acknowledge several limitations to our study. We do not have any information on family history. However, this information is not included in the FHS risk score, and a previous report indicates that inclusion of family history does not influence the magnitude of the association between a CHD-specific MGRS and CHD.[8,21] Moreover, although family history remains an important contributing factor to risk prediction, it has recently been shown that genetic scores can substantially increase the discriminatory abilities of the model above family history alone.[20] However, larger sample sizes for deriving these scores are needed.

Furthermore, our study was undertaken in Swedish middle-aged to elderly individuals; hence, the generalizability to other ethnicities or age groups is unknown. We expect early onset CHD events to be more influenced by genetic factors compared with late-onset CHD. Therefore, the predictive ability of a genetic score in younger individuals is likely to be higher.[23] Finally, we only included myocardial infarction and unstable angina in our outcome, but not stable angina, because the validity of this diagnosis in Swedish National Patient Register is unknown. Differences in the definition of the cardiovascular outcomes across studies might limit the comparability and generalizability of the results.

In conclusion, using data from 6 Swedish prospective cohort studies with 10612 healthy participants from the community, we have investigated the clinical use of genetic

scores in primary prevention of cardiovascular diseases. Current efforts to discover additional genetic loci associated with CHD and related traits, as well as the use of sequencing approaches along with integration of other omics technologies, are likely to further improve the performance of existing predictive equations. Our results indicate that genetic information could be of some clinical value for prediction of CHD, although further studies are needed to address aspects such as feasibility, ethics, and cost efficiency of genetic profiling in the primary prevention setting.

Acknowledgments

We thank Dr Rapsomaniki for the methodological contribution.

Sources of Funding

This study was supported by grants from the Swedish Foundation for Strategic Research (ICA08-0047), the Swedish Research Council (project grants no. 2012-1397 and 2009–2298), the Swedish Heart-Lung Foundation (2012-0197, 2010-0401 and 2008-0326), the National Institutes of Health (AG04563, AG10175, AG08861, AG08724, AG028555, and DK 066134), the Swedish Society of Medicine, the Knut and Alice Wallenberg Foundation, the Torsten and Ragnar Söderberg Foundation, and the Strategic Cardiovascular Program of Karolinska Institutet and the Stockholm County Council.

Disclosures

None.

References

1. The International HapMap Consortium. A haplotype map of the human genome. *Nature*. 2005;437:1299–1320.
2. CARDIoGRAMplusC4D Consortium, Deloukas P, Kanoni S, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet*. 2013;45:25–33.
3. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010;466:707–713.
4. Scott RA, Lagou V, Welch RP, et al.; DIAbetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet*. 2012;44:991–1005.
5. Greenland P, Alpert JS, Beller GA, et al.; American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: executive summary: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2010;122:2748–2764.
6. Graham I, Atar D, Borch-Johnsen K, et al. European guidelines on cardiovascular disease prevention in clinical practice: Executive summary: Fourth joint task force of the European society of cardiology and other societies on cardiovascular disease prevention in clinical practice (constituted by representatives of nine societies and by invited experts). *Eur Heart J*. 2007;28:2375–2414
7. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97:1837–1847.
8. 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*. 2010;376:1393–1400.
9. Paynter NP, Chasman DI, Paré G, Buring JE, Cook NR, Miletich JP, Ridker PM. Association between a literature-based genetic risk score and cardiovascular events in women. *JAMA*. 2010;303:631–637.
10. Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, Sklar P. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. 2009;460:748–752.
11. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA*. 2009;106:9362–9367.
12. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–575.
13. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447:661–678.
14. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive summary of the third report of the national cholesterol education program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel iii). *JAMA*. 2001;285:2486–2497.
15. Angelakopoulou A, Shah T, Sofat R, et al. Comparative analysis of genome-wide association studies signals for lipids, diabetes, and coronary heart disease: Cardiovascular Biomarker Genetics Collaboration. *Eur Heart J*. 2012;33:393–407.
16. Emerging Risk Factors Collaboration, Kaptoge S, Di Angelantonio E, et al. C-reactive protein, fibrinogen, and cardiovascular disease prediction. *The N Engl J Med.*. 2012;367:1310–1320.
17. Genin E, Hannequin D, Wallon D, et al. APOE and Alzheimer disease: a major gene with semi-dominant inheritance. *Mol Psychiatry*. 2011;16:903–907.
18. Evans DM, Visscher PM, Wray NR. Harnessing the information contained within genome-wide association studies to improve individual prediction of complex disease risk. *Hum Mol Genet*. 2009;18:3525–3531.
19. Dudbridge F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet*. 2013;9:e1003348.
20. Chatterjee N, Wheeler B, Sampson J, Hartge P, Chanock SJ, Park JH. Projecting the performance of risk prediction based on polygenic analyses of genome-wide association studies. *Nat Genet*. 2013;45:400–5, 405e1.
21. Thanassoulis G, Peloso GM, Pencina MJ, Hoffmann U, Fox CS, Cupples LA, Levy D, D'Agostino RB, Hwang SJ, O'Donnell CJ. A genetic risk score is associated with incident cardiovascular disease and coronary artery calcium: the Framingham Heart Study. *Circ Cardiovasc Genet*. 2012;5:113–121.
22. Vaarhorst AA, Lu Y, Heijmans BT, et al. Literature-based genetic risk scores for coronary heart disease: the Cardiovascular Registry Maastricht (CAREMA) prospective cohort study. *Circ Cardiovasc Genet*. 2012;5:202–209.
23. Nora JJ, Lortscher RH, Spangler RD, Nora AH, Kimberling WJ. Genetic–epidemiologic study of early-onset ischemic heart disease. *Circulation*. 1980;61:503–508.

Significance

Both US and European guidelines for prevention of cardiovascular diseases in clinical practice advise against prognostic use of DNA-based tests in the primary prevention setting. In this study, we show that genetic information could be of some clinical value for prediction of coronary artery disease. By measuring the coronary artery disease–specific genetic score in addition to the established risk factors, 1 additional coronary artery disease event for every 318 people screened at intermediate risk can be saved. Both a score including 46 well-established coronary artery disease–associated SNPs and a score including 395 SNPs associated with cardiovascular risk factors have similar prediction performances.

SUPPLEMENTAL MATERIAL

Multi-locus genetic risk scores for coronary heart disease prediction

Supplementary Table I. Cardiovascular Related Traits Selected from the NHGRI Catalog and Original Number of SNPs for Each Trait.

Trait selected	N. SNPs reported in the catalog
abdominal aortic aneurysm	4
adiponectin levels	48
adiposity	8
thoracic aortic aneurysms and dissections	1
aortic root size	13
arterial stiffness	1
atrial fibrillation	21
atrial fibrillation/atrial flutter	2
atrioventricular conduction	5
blood pressure	74
body mass (lean)	1
body mass index	112
body mass index and fat mass	1
c-reactive protein	54
cardiac structure and function	4
cardiovascular disease risk factors	31
cholesterol	2
cholesterol, total	67
coronary artery calcification	6
coronary heart disease	185
coronary spasm	1
diabetes (incident)	1
diabetes related insulin traits	4
diastolic blood pressure	32
echocardiographic traits	8
electrocardiographic conduction measures	7
electrocardiographic traits	23
endothelial function traits	1
fasting glucose-related traits	30
fasting insulin-related traits	4
fasting plasma glucose	16
fibrinogen	15
glycated hemoglobin levels	17
hdl cholesterol	119
heart failure	17
heart rate variability traits	4
hemostatic factors and hematological phenotypes	27
hypertension	35

hypertension (young onset)	1
hypertriglyceridemia	5
insulin resistance/response	15
interleukin-18 levels	2
ldl cholesterol	104
left ventricular mass	5
proinsulin levels	10
lipoprotein-associated phospholipase A2 activity and mass	12
major cvd	1
metabolic syndrome	45
myocardial infarction	3
myocardial infarction (early onset)	9
obesity	29
obesity (early onset extreme)	1
obesity (extreme)	19
obesity-related traits	4
pr interval	10
peripheral artery disease	2
qt interval	45
rr interval (heart rate)	20
resting heart rate	10
smoking behavior	48
stroke	9
subarachnoid aneurysmal hemorrhage	1
subclinical atherosclerosis traits	8
subclinical brain infarct	1
sudden cardiac arrest	52
systolic blood pressure	30
triglycerides	98
two-hour glucose challenge	5
type 1 diabetes	97
type 2 diabetes	199
type 2 diabetes and 6 quantitative traits	1
type 2 diabetes and other traits	6
ventricular conduction	25
ventricular fibrillation	1
waist circumference	17
waist circumference and related phenotypes	4
waist-hip ratio	17
weight	33
cardiac hypertrophy	19

dilated cardiomyopathy	2
hdl cholesterol - triglycerides (hdlg-tg)	8
metabolic syndrome (bivariate traits)	16
peripartum cardiomyopathy	1
sick sinus syndrome	1
triglycerides-blood pressure (tg-bp)	6
waist circumference - triglycerides (wc-tg)	7
natriuretic peptide levels	3
coronary restenosis	1
carotid intima media thickness	11
metabolic traits	48
ankle-brachial index	2
aortic stiffness	1
body mass index and cholesterol (psychopharmacological treatment)	2
cardiac repolarization	4
creatinine levels	6
diabetes (gestational)	3
hypertension risk in short sleep duration	3
inflammatory biomarkers	26
insulin-related traits	7
lipid levels in hepatitis c treatment	1
lipid metabolism phenotypes	18
lipid traits	4
lp (a) levels	5
matrix metalloproteinase levels	4
obesity and blood pressure	5
polyunsaturated fatty acid levels	1
postoperative ventricular dysfunction	1
preeclampsia	1
type 2 diabetes and gout	1
urate levels	23
uric acid levels	32
vwf and fviii levels	11
stroke (ischemic)	12
pericardial fat	1
plasminogen activator inhibitor type 1 levels (pai-1)	4
subcutaneous adipose tissue	39
visceral adipose tissue adjusted for bmi	45
visceral adipose tissue/subcutaneous adipose tissue ratio	42
visceral fat	40

Supplementary Table II. Single-SNP Association with CHD for All the SNPs Included in at Least One MGRS Ordered by P-value.

SNPs	Risk allele	Risk allele frequency	Hazard Ratio [a]	P-value [a]	Trait-specific MGRS	FHS MGRS y/n [b]	Overall MGRS y/n [c]	Trait in the catalog	Gene reported in the original paper	Reference (PubMed ID)
rs419076	T	0.45	0.83	0.0005	SBP	1	1	blood pressure	MECOM	21909110
rs7784776	G	0.36	1.23	0.0013	-	0	1	ventricular conduction	IGFBP3	21076409
rs1051730	A	0.34	0.85	0.0042	Smoke	1	0	smoking behavior	CHRNA3	20418890
rs11755527	G	0.38	1.19	0.0043	-	0	1	type 1 diabetes	BACH2	18978792
rs11222084	T	0.36	0.86	0.0071	-	0	1	blood pressure	ADAMTS-8	21909110
rs13292136	C	0.91	0.79	0.0078	T2D	1	1	type 2 diabetes	CHCHD9	20581827
rs2126259	C	0.89	1.27	0.0080	-	0	1	ldl cholesterol	PPP1R3B	20864672
rs4148008	G	0.30	1.15	0.0103	HDL	1	1	lipids	-	20686565
rs941576	A	0.52	0.89	0.0188	-	0	1	type 1 diabetes	DLK1, MEG3, RTL1, DIO3	19966805
rs13376333	T	0.33	0.86	0.0197	-	0	1	atrial fibrillation	KCNN3	20173747
rs340029	T	0.63	1.15	0.0206	-	0	1	c-reactive protein	RORA	21300955
rs2191349	T	0.50	1.12	0.0211	-	0	1	fasting glucose-related traits	DGKB, TMEM195	20081858
rs2847281	A	0.59	0.88	0.0215	-	0	1	c-reactive protein	PTPN2	21300955
rs9987289	A	0.10	0.82	0.0238	HDL	1	0	lipids	-	20686565
rs3903239	G	0.48	1.14	0.0255	-	0	1	atrial fibrillation	PRRX1	22544366
rs10889353	A	0.67	1.14	0.0283	TC	1	0	cholesterol, total	DOCK7	19060911
rs1539019	C	0.60	0.81	0.0304	-	0	1	fibrinogen	NLRP3	20031576
rs9976767	G	0.44	1.15	0.0319	-	0	1	type 1 diabetes	UBASH3A	18840781
rs7202877	T	0.89	1.22	0.0322	T2D	1	1	type 2 diabetes	BCAR1	22885922
rs1167998	A	0.67	1.13	0.0341	-	0	1	triglycerides	DOCK7	19060911
rs9515203	T	0.70	1.13	0.0363	CHD	0	1	coronary heart disease	COL4A1/COL4A2	23202125
rs7072268	T	0.46	0.90	0.0368	-	0	1	glycated hemoglobin levels	HK1	19096518
rs1387153	T	0.28	0.89	0.0418	T2D	1	0	fasting plasma glucose	MTNR1B	19060909

rs737337	C	0.11	1.17	0.0447	HDL	1	1	lipids	-	20686565
rs2047009	G	0.51	1.11	0.0486	CHD	0	1	coronary heart disease	CXCL12	23202125
rs10923931	T	0.10	1.18	0.0512	T2D	1	1	type 2 diabetes	NOTCH2, ADAM30	18372903
rs251253	T	0.61	0.90	0.0519	-	0	1	pr interval	NKX2-5, C5orf41	20062060
rs445925	G	0.88	1.18	0.0557	CHD	0	1	coronary heart disease	ApoE/ApoC1	23202125
rs12190287	C	0.62	1.11	0.0562	CHD	0	1	coronary heart disease	TCF21	23202125
rs2228671	C	0.90	1.18	0.0608	TC	1	0	cholesterol, total	LDLR	19060911
rs11605924	A	0.49	1.10	0.0633	-	0	1	fasting glucose-related traits	CRY2	20081858
rs4759375	C	0.94	1.25	0.0652	HDL	1	1	lipids	-	20686565
rs7903146	T	0.26	1.11	0.0656	T2D	1	1	proinsulin levels	TCF7L2	21873549
rs6987702	C	0.30	1.11	0.0715	TC	1	1	cholesterol, total	TRIB1	19060911
rs2078267	C	0.48	0.91	0.0722	-	0	1	urate levels	SLC22A11	20884846
rs17609940	G	0.77	0.90	0.0749	CHD	0	1	coronary heart disease	ANKS1A	21378990
rs11191548	T	0.90	1.17	0.0755	SBP	1	0	blood pressure	CYP17A1, NT5C2	21909110
rs7173743	T	0.59	1.10	0.0767	CHD	0	0	coronary heart disease	ADAMTS7	23202125
rs7692387	G	0.80	1.13	0.0778	CHD	0	1	coronary heart disease	GUCY1A3	23202125
rs12130333	C	0.79	1.12	0.0834	-	0	1	triglycerides	ANGPTL3, DOCK7, ATG4C	18193044
rs1004467	A	0.89	1.15	0.0877	SBP	1	1	systolic blood pressure	CYP17A1	19430479
rs6495122	A	0.42	1.09	0.0890	-	0	1	diastolic blood pressure	CSK, ULK3	19430479
rs11591147	G	0.98	1.50	0.0912	-	0	1	ldl cholesterol	PCSK9	18193044
rs3810291	A	0.70	1.11	0.0946	BMI	1	1	body mass index	TMEM160, ZC3H4	20935630
rs4687718	G	0.88	1.19	0.0951	-	0	1	ventricular conduction	TKT, PRKCD, CACNA1D	21076409
rs9319428	A	0.33	1.09	0.0969	CHD	0	1	coronary heart disease	FLT1	23202125
rs2967605	T	0.17	1.12	0.0971	HDL	1	1	hdl cholesterol	ANGPTL4	19060906
rs6601530	G	0.46	0.92	0.0980	-	0	1	carotid intima media thickness	PINX1	21909108

rs7422339	C	0.68	1.10	0.1002	-	0	1	fibrinogen	CPS1	20031577
rs1805017	T	0.24	0.88	0.1093	-	0	1	lipoprotein-associated phospholipase a2 activity and mass	PLA2G7	22003152
rs501120	T	0.87	1.13	0.1121	CHD	0	0	coronary heart disease	CXCL12	23202125
rs3001032	T	0.69	0.92	0.1136	-	0	1	adiponectin levels	LYPLAL1	22479202
rs13238203	C	0.97	1.38	0.1140	-	0	1	lipids	-	20686565
rs1173771	G	0.59	1.09	0.1147	SBP	1	1	blood pressure	NPR3, C5orf23	21909110
rs731839	G	0.33	1.09	0.1162	-	0	1	adiponectin levels	PEPD	22479202
rs9982601	T	0.14	1.12	0.1183	CHD	0	1	coronary heart disease	gene_desert/KCNE2	23202125
rs560887	C	0.70	1.09	0.1195	-	0	1	fasting glucose-related traits	G6PC2	20081858
rs3817334	T	0.41	0.92	0.1226	BMI	1	1	body mass index	MTCH2, NDUFS3, CUGBP1	20935630
rs6474412	T	0.84	0.86	0.1249	Smoke	1	1	smoking behavior	CHRNB3,CHRNA6	20418888
rs2048327	C	0.39	1.08	0.1267	CHD	0	1	coronary heart disease	SLC22A3/LPAL2/LPA	23202125
rs1042034	T	0.79	1.11	0.1277	-	0	1	lipids	-	20686565
rs11191593	T	0.90	1.14	0.1280	-	0	1	blood pressure	CYP17A1, NT5C2	21909110
rs478222	A	0.55	0.92	0.1318	-	0	1	type 1 diabetes	EFR3B, 3NCOA1, C2orf79, CENPO, ADCY3, DNAJC27, POMC, DN	21980299
rs2814944	A	0.15	0.89	0.1343	HDL	0	0	lipids	-	20686565
rs2384550	G	0.64	0.93	0.1416	-	0	1	diastolic blood pressure	TBX3, TBX5	19430479
rs8068318	C	0.25	0.91	0.1439	-	0	1	creatinine levels	TBX2	20383145
rs1127065	C	0.55	1.08	0.1441	-	0	1	metabolic syndrome	CAMK2B	22399527
rs6015450	G	0.14	1.12	0.1453	SBP	1	1	blood pressure	GNAS, EDN3	21909110
rs1362212	A	0.12	1.15	0.1466	-	0	1	ventricular conduction	TBX20	21076409
rs11206510	T	0.81	1.10	0.1492	CHD	0	1	ldl cholesterol	PCSK9	19060906

rs12413409	G	0.90	1.13	0.1493	CHD	0	0	coronary heart disease	CYP17A1, CNNM2, NT5C2	21378990
rs673548	G	0.80	1.10	0.1494	HDL	0	0	metabolic syndrome	APOB	22399527
rs1333049	C	0.44	1.08	0.1496	CHD	0	0	coronary artery calcification	CDKN2B	22144573
rs2521501	T	0.32	1.09	0.1497	SBP	1	0	blood pressure	FES	21909110
rs10913469	C	0.22	0.91	0.1501	-	0	1	weight	SEC16B, RASAL2	19079260
rs633185	C	0.70	0.92	0.1502	SBP	1	1	blood pressure	FLJ32810, TMEM133	21909110
rs3733829	G	0.33	0.91	0.1504	Smoke	1	1	smoking behavior	CYP2A6,EGLN2	20418890
rs1799884	T	0.15	1.10	0.1516	-	0	1	glycated hemoglobin levels	GCK	20858683
rs12670798	C	0.26	1.09	0.1553	TC	1	1	lipids	-	20686565
rs12936587	G	0.52	1.08	0.1572	CHD	0	1	coronary heart disease	RAI1/PEMT/RASD1	23202125
rs442177	T	0.56	1.08	0.1588	-	0	1	triglycerides	AFF1	20864672
rs11556924	C	0.63	1.08	0.1618	CHD	0	1	coronary heart disease	ZC3HC1	23202125
rs2200733	T	0.09	0.88	0.1620	-	0	1	atrial fibrillation/atrial flutter	PITX2,ENPEP	17603472
rs11047543	G	0.90	0.85	0.1624	-	0	1	pr interval	SOX5, C12orf67	20062060
rs7025486	A	0.26	1.09	0.1649	-	0	1	abdominal aortic aneurysm	DAB2IP	20622881
rs1501908	C	0.65	0.93	0.1659	-	0	1	ldl cholesterol	TIMD4, HAVCR1	19060906
rs268	G	NA	1.44	0.1666	HDL	1	1	metabolic syndrome	LPL	22399527
rs6474359	T	0.97	0.83	0.1703	-	0	1	glycated hemoglobin levels	ANK1	20858683
rs9488822	A	0.70	0.93	0.1728	TC	1	1	lipids	-	20686565
rs16942887	G	0.86	0.91	0.1738	HDL	1	1	lipids	-	20686565
rs7562790	G	0.41	1.07	0.1789	-	0	1	ventricular conduction	CRIM1	21076409
rs6882076	C	0.64	0.93	0.1805	TC	1	0	lipids	-	20686565
rs1321311	A	0.23	0.92	0.1816	-	0	1	electrocardiographic traits	CDKN1A	20062063
rs174548	G	0.32	0.93	0.1828	HDL	1	0	hdl cholesterol	FADS1	20864672
rs319690	T	0.66	0.92	0.1842	-	0	1	blood pressure	MAP4	21909110
rs10903129	G	0.55	1.07	0.1854	TC	1	1	cholesterol, total	TMEM57	19060911

rs7342028	T	0.15	1.16	0.1865	-	0	1	ventricular conduction	VTI1A	21076409
rs7692808	G	0.71	0.93	0.1951	-	0	1	pr interval	ARHGAP24	20062060
rs2642442	T	0.70	0.93	0.1965	TC	1	1	lipids	-	20686565
rs4846914	G	0.39	1.07	0.1988	HDL	1	1	hdl cholesterol	GALNT2	19060906
rs539514	T	0.53	1.08	0.1990	-	0	1	type 1 diabetes	LMO7	21980299
rs4689388	A	0.56	1.07	0.2002	T2D	1	1	type 2 diabetes and other traits	WFS1, PPP2R2C	19734900
rs1402837	T	0.23	1.08	0.2011	-	0	1	glycated hemoglobin levels	G6PC2	19096518
rs9436640	T	0.51	1.12	0.2033	-	0	1	ventricular conduction	NFIA	21076409
rs12129861	G	0.50	0.93	0.2063	-	0	1	uric acid levels	PDZK1	19503597
rs12425791	A	0.16	1.09	0.2083	-	0	1	stroke	NINJ2	19369658
rs742132	A	0.71	1.07	0.2117	-	0	1	uric acid levels	LRRC16A, SCGN	19503597
rs7961581	C	0.26	1.07	0.2126	T2D	1	1	type 2 diabetes	TSPAN8,LGR5	18372903
rs10968576	G	0.31	1.07	0.2175	BMI	1	1	body mass index	LRRN6C	20935630
rs10519210	G	0.03	1.20	0.2223	-	0	1	heart failure	USP3	20445134
rs3025343	A	0.09	1.11	0.2233	Smoke	1	1	smoking behavior	DBH	20418890
rs2241423	G	0.79	0.93	0.2238	BMI	1	1	body mass index	MAP2K5, LBXCOR1	20935630
rs6725887	C	0.12	0.91	0.2252	CHD	0	1	coronary heart disease	WDR12	23202125
rs7138803	A	0.42	1.07	0.2260	BMI	1	1	body mass index	FAIM2	20935630
rs2929282	T	0.04	0.83	0.2265	-	0	1	lipids	-	20686565
rs2247056	C	0.72	1.07	0.2295	-	0	1	lipids	-	20686565
rs4790333	T	0.45	1.07	0.2320	-	0	1	proinsulin levels	SGSM2	21873549
rs11084753	G	0.68	1.07	0.2359	BMI	1	1	body mass index	KCTD15	19079261
rs381815	T	0.29	1.07	0.2379	SBP	1	1	blood pressure	PLEKHA7	21909110
rs231362	G	0.51	0.94	0.2421	T2D	1	1	type 2 diabetes	KCNQ1	20581827
rs386000	G	0.75	1.07	0.2448	HDL	1	1	lipids	-	20686565
rs7819412	A	0.51	1.06	0.2457	-	0	1	triglycerides	XKR6, AMAC1L2	19060906
rs1549318	T	0.58	1.08	0.2461	-	0	1	proinsulin levels	LARP6	21873549
rs3184504	T	0.46	1.06	0.2469	CHD	0	0	coronary heart disease	SH2B3	23202125
rs3184504	T	0.46	1.06	0.2469	SBP	0	0	coronary heart disease	SH2B3	23202125

rs4929949	C	0.50	0.94	0.2475	BMI	1	1	body mass index	RPL27A, TUB	20935630
rs1178979	T	0.80	0.93	0.2477	-	0	1	triglycerides	BAZ1B, BCL7B, TBL2, MLXIPL	20864672
rs492602	G	0.45	0.94	0.2489	TC	1	1	lipids	-	20686565
rs2814982	C	0.88	1.10	0.2538	TC	1	1	lipids	-	20686565
rs11065987	A	0.59	0.94	0.2546	TC	1	1	lipids	-	20686565
rs1555543	C	0.57	1.06	0.2546	BMI	1	1	body mass index	PTBP2	20935630
rs174546	T	0.34	0.94	0.2609	-	0	1	lipids	-	20686565
rs1990760	T	0.63	1.07	0.2613	-	0	1	type 1 diabetes	IFIH1	17554260
rs9851724	T	0.67	1.07	0.2637	-	0	1	ventricular conduction	SCN10A,SCN5A	21076409
rs4660293	G	0.24	0.93	0.2675	HDL	1	1	lipids	-	20686565
rs2304256	C	0.65	0.90	0.2763	-	0	1	type 1 diabetes	TYK2	19966805
rs9299	T	0.64	1.06	0.2893	BMI	1	1	obesity	HOXB5	22484627
rs4944092	A	0.63	0.94	0.2930	-	0	1	pr interval	WNT11	20062060
rs7941030	C	0.40	1.06	0.2961	TC	1	1	lipids	-	20686565
rs6235	G	0.29	1.06	0.2971	-	0	1	proinsulin levels	PCSK1	21873549
rs7756935	C	0.15	0.88	0.2989	-	0	1	lipoprotein-associated phospholipase a2 activity and mass	PLA2G7	22003152
rs7957197	T	0.79	0.94	0.3002	T2D	1	1	type 2 diabetes	HNF1A	20581827
rs4026608	T	0.60	0.91	0.3020	-	0	1	aortic root size	HMGA2	19584346
rs2954029	A	0.51	1.06	0.3031	CHD	0	1	coronary heart disease	TRIB1	23202125
rs4149268	T	0.36	0.95	0.3033	HDL	1	1	hdl cholesterol	ABCA1	18193043
rs7206971	A	0.47	1.05	0.3111	-	0	1	lipids	-	20686565
rs2542151	G	0.17	1.07	0.3123	-	0	1	type 1 diabetes	PTPN2	17554260
rs10150332	C	0.21	1.06	0.3174	BMI	1	1	body mass index	NRXN3	20935630
rs3217992	T	0.32	1.06	0.3186	CHD	0	0	coronary heart disease	CDKN2BAS	23202125
rs2112347	T	0.63	1.06	0.3186	BMI	1	1	body mass index	FLJ35779, HMGCR	20935630
rs805303	G	0.62	0.95	0.3208	SBP	1	1	diastolic blood pressure	BAT2, BAT5	21909115
rs2072183	C	0.24	1.06	0.3226	TC	1	1	lipids	-	20686565
rs505802	C	0.30	1.06	0.3228	-	0	1	uric acid levels	SLC22A12	19503597
rs2068888	G	0.56	1.05	0.3237	-	0	1	lipids	-	20686565

rs35767	G	0.84	1.07	0.3249	-	0	1	fasting glucose-related traits	IGF1	20081858
rs4252120	T	0.71	1.06	0.3257	CHD	0	1	coronary heart disease	PLG	23202125
rs675209	T	0.25	0.94	0.3294	-	0	1	urate levels	RREB1	20884846
rs1446468	C	0.59	1.05	0.3311	-	0	1	blood pressure	FIGN	21909110
rs2254287	G	0.37	0.95	0.3329	-	0	1	ldl cholesterol	B3GALT4	18193043
rs1800588	C	0.78	1.07	0.3336	HDL	1	1	hdl cholesterol	LIPC	18193044
rs4771122	G	0.25	1.06	0.3336	BMI	1	1	body mass index	MTIF3, GTF3A	20935630
rs2107595	A	0.17	1.08	0.3379	-	0	1	stroke (ischemic)	HDAC9	23041239
rs556621	T	0.38	1.10	0.3380	-	0	1	stroke (ischemic)	CDC5L, SUPT3H	22941190
rs17696736	G	0.42	1.05	0.3405	-	0	1	type 1 diabetes	C12orf30	17554260
rs2890652	C	0.18	1.07	0.3437	BMI	1	1	body mass index	LRP1B	20935630
rs511154	A	0.21	0.94	0.3443	-	0	1	fibrinogen	PCCB	20031576
rs281868	A	0.53	0.95	0.3452	-	0	1	resting heart rate	SLC35F1	20639392
rs9398652	C	0.91	0.91	0.3485	-	0	1	resting heart rate	GJA1	20639392
rs11848785	A	0.77	1.06	0.3548	-	0	1	ventricular conduction	SIPA1L1	21076409
rs17782313	C	0.25	1.06	0.3549	BMI	1	1	body mass index	MC4R	19079261
rs2242285	A	0.41	1.05	0.3591	-	0	1	ventricular conduction	LRIG1,SLC25A26	21076409
rs925946	T	0.35	0.95	0.3595	BMI	1	1	body mass index	BDNF	19079260
rs763361	T	0.47	1.05	0.3602	-	0	1	type 1 diabetes	CD226	17554260
rs11220462	A	0.15	1.07	0.3628	-	0	1	lipids	-	20686565
rs1878406	T	0.11	0.93	0.3650	CHD	0	1	coronary heart disease	EDNRA	23202125
rs16948048	G	0.40	1.05	0.3663	-	0	1	diastolic blood pressure	ZNF652, PHB	19430483
rs2923084	G	0.18	0.94	0.3707	HDL	1	1	lipids	-	20686565
rs4457053	G	0.25	0.95	0.3771	T2D	1	1	type 2 diabetes	ZBED3	20581827
rs12967135	A	0.25	1.06	0.3786	HDL	0	0	lipids	-	20686565
rs883079	C	0.28	1.05	0.3790	-	0	1	ventricular conduction	TBX5	21076409
rs6734238	G	0.40	0.95	0.3796	-	0	1	c-reactive protein	IL1F10	21300955
rs2479409	G	0.33	0.95	0.3803	-	0	1	lipids	-	20686565
rs4105144	C	0.63	1.10	0.3874	Smoke	1	1	smoking behavior	CYP2A6,RAB4D	20418888
rs1084651	G	0.81	0.95	0.3898	HDL	1	1	lipids	-	20686565

rs543874	G	0.22	0.95	0.3987	BMI	1	0	body mass index	SEC16B	20935630
rs1733724	A	0.16	0.91	0.4009	-	0	1	ventricular conduction	DKK1	21076409
rs11897119	C	0.41	0.96	0.4048	-	0	1	pr interval	MEIS1	20062060
rs7034200	A	0.48	0.96	0.4071	-	0	1	fasting glucose-related traits	GLIS3	20081858
rs1886512	T	0.68	0.95	0.4099	-	0	1	ventricular conduction	KLF12	21076409
rs4845625	T	0.43	1.04	0.4102	CHD	0	1	coronary heart disease	IL6R	23202125
rs12328675	T	0.88	0.94	0.4129	HDL	1	1	lipids	-	20686565
rs6450176	A	0.25	0.95	0.4129	HDL	1	1	lipids	-	20686565
rs4765127	G	0.66	0.96	0.4178	HDL	1	1	lipids	-	20686565
rs10953541	C	0.75	1.05	0.4181	CHD	0	1	coronary heart disease	Intergenic	21378988
rs11869286	G	0.32	0.96	0.4197	HDL	1	1	lipids	-	20686565
rs1122608	G	0.76	1.05	0.4200	CHD	0	1	coronary heart disease	LDLR	23202125
rs4607103	C	0.77	1.05	0.4206	T2D	1	1	type 2 diabetes	ADAMTS9	18372903
rs10842994	C	0.80	0.95	0.4211	T2D	1	1	type 2 diabetes	KLHDC5	22885922
rs4373814	C	0.44	0.96	0.4221	SBP	1	1	diastolic blood pressure	CACNB2	21909115
rs508487	T	0.05	0.90	0.4225	-	0	1	cardiovascular disease risk factors	PCSK7	21943158
rs987237	G	0.18	1.05	0.4263	BMI	1	1	body mass index	TFAP2B	20935630
rs6800541	C	0.38	1.04	0.4287	-	0	1	pr interval	SCN10A	20062060
rs223116	G	0.82	1.09	0.4305	-	0	1	resting heart rate	MYH7,NDNG	20639392
rs4074536	T	0.73	1.05	0.4317	-	0	1	ventricular conduction	CASQ2	21076409
rs782590	T	0.53	0.96	0.4324	SBP	1	1	metabolic syndrome	SMEK2	22399527
rs10838687	T	0.78	0.95	0.4339	-	0	1	proinsulin levels	MADD	21873549
rs10938397	G	0.41	0.96	0.4370	BMI	1	1	body mass index	GNPDA2	20935630
rs4731702	C	0.50	0.96	0.4394	HDL	1	1	lipids	-	20686565
rs17367504	A	0.85	1.06	0.4434	SBP	1	1	blood pressure	MTHFR, NPPB	21909110
rs2199936	A	0.10	1.07	0.4482	-	0	1	urate levels	ABCG2	20884846
rs11781551	G	0.53	1.04	0.4485	-	0	1	carotid intima media thickness	ZHX2	21909108
rs9818870	T	0.15	0.94	0.4499	CHD	0	0	coronary heart disease	MRAS	23202125

rs229541	A	0.41	1.05	0.4513	-	0	1	type 1 diabetes	C1QTNF6	18978792
rs10947789	T	0.74	0.96	0.4523	CHD	0	1	coronary heart disease	KCNK5	23202125
rs924043	C	0.87	0.94	0.4537	-	0	1	type 1 diabetes	WDR27, C6orf120, PHF10, TCTE3, C6orf208, LOC154449, DLL	21980299
rs10206899	T	0.78	0.95	0.4540	-	0	1	creatinine levels	NAT8,NAT8B,ALMS1,DUSP11,TPRKB	20383145
rs713586	C	0.47	0.96	0.4566	BMI	1	0	body mass index	RBJ, ADCY3, POMC	20935630
rs2518491	T	0.27	0.96	0.4584	-	0	1	glycated hemoglobin levels	SPTA1	20858683
rs12051272	T	0.02	1.24	0.4587	-	0	1	adiponectin levels	CDH13	22479202
rs972283	G	0.50	0.96	0.4604	T2D	0	0	type 2 diabetes	KLF14	20581827
rs10770612	G	0.16	1.08	0.4644	-	0	1	aortic root size	PDE3A	19584346
rs7164883	G	0.17	0.95	0.4656	-	0	1	atrial fibrillation	HCN4	22544366
rs1514175	A	0.43	0.96	0.4675	BMI	1	1	body mass index	TNNI3K	20935630
rs17114036	A	0.92	1.08	0.4689	CHD	0	1	coronary heart disease	PPAP2B	23202125
rs4698036	T	0.76	1.06	0.4692	-	0	1	cardiovascular disease risk factors	Intergenic	21943158
rs11558471	A	0.68	1.04	0.4744	-	0	1	fasting glucose-related traits	SLC30A8	20081858
rs8090011	G	0.38	0.96	0.4753	T2D	1	1	type 2 diabetes	LAMA1	22693455
rs9568856	A	0.14	1.06	0.4753	BMI	1	1	obesity	OLFM4	22484627
rs12779790	G	0.19	0.95	0.4766	T2D	1	1	type 2 diabetes	CDC123,CAMK1D	18372903
rs11118316	A	0.46	0.96	0.4805	BMI	1	0	visceral adipose tissue/subcutaneous adipose tissue ratio	LYPLAL1	22589738
rs1800562	G	0.94	0.93	0.4807	-	0	1	glycated hemoglobin levels	HFE	20858683
rs4773144	G	0.40	0.96	0.4810	CHD	0	1	coronary heart disease	COL4A1/COL4A2	23202125
rs2796441	G	0.58	0.97	0.4814	T2D	1	1	type 2 diabetes	TLE1	22885922
rs1152591	A	0.50	1.06	0.4827	-	0	1	atrial fibrillation	SYNE2	22544366
rs896854	T	0.54	0.96	0.4877	T2D	1	1	type 2 diabetes	TP53INP1	20581827

rs2782980	C	0.70	0.96	0.4885	-	0	1	blood pressure	ADRB1	21909110
rs7679	C	0.19	0.96	0.4916	HDL	1	1	hdl cholesterol	PLTP	19060906
rs1561198	T	0.46	1.04	0.4917	CHD	0	1	coronary heart disease	GGCX/VAMP8	23202125
rs10824026	A	0.84	1.05	0.4941	-	0	1	atrial fibrillation	SYNPO2L	22544366
rs11136341	G	0.35	1.04	0.4949	-	0	1	lipids	-	20686565
rs2081687	T	0.34	1.04	0.4984	TC	1	1	lipids	-	20686565
rs964184	G	0.13	1.05	0.4999	HDL	1	1	coronary heart disease	ZNF259, APOA5, APOA4, APOC3, APOA1	21378990
rs964184	G	0.13	1.05	0.4999	CHD	1	1	coronary heart disease	ZNF259, APOA5, APOA4, APOC3, APOA1	21378990
rs2106261	T	0.18	0.96	0.5006	-	0	1	atrial fibrillation	ZFHX3	22544366
rs181362	T	0.24	1.04	0.5016	HDL	1	1	lipids	-	20686565
rs13002573	A	0.77	1.04	0.5031	-	0	1	blood pressure	FIGN	21909110
rs7570971	A	0.25	0.96	0.5035	TC	1	1	lipids	-	20686565
rs7578326	A	0.64	1.04	0.5046	T2D	1	1	type 2 diabetes	IRS1	20581827
rs10830962	G	0.36	0.96	0.5113	-	0	1	metabolic syndrome	MTNR1B	22399527
rs174570	C	0.85	1.05	0.5154	TC	1	0	cholesterol, total	FADS2, FADS3	19060911
rs247617	C	0.19	0.96	0.5175	HDL	1	1	metabolic syndrome	CETP	22399527
rs1466535	G	0.64	1.04	0.5197	-	0	1	abdominal aortic aneurysm	LRP1	22055160
rs1800961	T	0.05	0.93	0.5294	HDL	1	1	hdl cholesterol	HNF4A	19060906
rs8042680	A	0.29	1.04	0.5343	T2D	1	1	type 2 diabetes	PRC1	20581827
rs273909	G	0.15	1.05	0.5351	CHD	0	1	coronary heart disease	SLC22A4/SLC22A5	23202125
rs1378942	C	0.32	1.03	0.5405	SBP	1	0	blood pressure	CSK	21909110
rs9912468	G	0.40	1.03	0.5416	-	0	1	ventricular conduction	PRKCA	21076409
rs7359397	T	0.42	0.97	0.5433	BMI	0	0	body mass index	SH2B1, APOB48R, SULT1A2, AC138894.2, ATXN2L, TUFM	20935630
rs7395662	G	0.63	0.97	0.5460	HDL	1	1	hdl cholesterol	MADD, FOLH1	19060911
rs7177055	A	0.72	1.04	0.5462	T2D	1	1	type 2 diabetes	HMG20A	22885922
rs11649653	C	0.59	0.97	0.5501	-	0	1	lipids	-	20686565
rs2902940	A	0.72	1.03	0.5515	TC	1	1	lipids	-	20686565

rs932764	G	0.45	1.03	0.5569	SBP	1	1	hypertension	PLCE1	21909115
rs514230	T	0.51	0.97	0.5600	TC	1	1	lipids	-	20686565
rs4737009	A	0.23	1.04	0.5610	-	0	1	glycated hemoglobin levels	ANK1	20858683
rs4420065	C	0.62	1.03	0.5619	-	0	1	c-reactive protein	LEPR	21300955
rs2071518	T	0.24	0.96	0.5723	-	0	1	blood pressure	NOV	21909110
rs2116830	G	0.87	1.07	0.5743	BMI	1	1	obesity	KCNMA1	21708048
rs2925979	T	0.30	1.03	0.5761	HDL	1	1	adiponectin levels	CMIP	22479202
rs1803274	C	0.86	0.94	0.5817	-	0	1	cardiovascular disease risk factors	BCHE	21943158
rs3846662	G	0.44	1.03	0.5892	TC	1	1	cholesterol, total	HMGCR	19060911
rs2304130	A	0.89	1.04	0.5899	TC	1	1	cholesterol, total	NCAN	19060911
rs6864049	G	0.53	1.03	0.5906	BMI	1	1	body mass index	ZNF608	20935630
rs10033464	T	0.11	0.96	0.5914	-	0	1	atrial fibrillation/atrial flutter	PITX2,ENPEP	17603472
rs515135	C	0.84	1.04	0.5925	CHD	0	0	coronary heart disease	APOB	23202125
rs3825932	T	0.33	1.03	0.5974	-	0	1	type 1 diabetes	CTSH	18978792
rs9369640	A	0.65	1.03	0.5991	CHD	0	0	coronary heart disease	PHACTR1	23202125
rs1424233	T	0.46	1.03	0.5995	BMI	1	1	obesity	MAF	19151714
rs3177928	A	0.14	0.96	0.6006	TC	1	1	lipids	-	20686565
rs11634397	G	0.65	1.03	0.6019	T2D	1	1	type 2 diabetes	ZFAND6	20581827
rs2844479	A	0.68	1.03	0.6027	-	0	1	weight	AIF1, NCR3	19079260
rs13082711	C	0.22	0.97	0.6047	-	0	1	blood pressure	SLC4A7	21909110
rs10468017	C	0.72	1.03	0.6049	HDL	1	1	hdl cholesterol	LIPC	19060906
rs1564348	T	0.85	0.96	0.6054	-	0	1	lipids	-	20686565
rs2522056	G	0.81	0.97	0.6094	-	0	1	fibrinogen	IRF1	20031576
rs459193	G	0.75	1.03	0.6123	T2D	1	1	type 2 diabetes	ANKRD55	22885922
rs17391905	T	0.97	0.93	0.6142	-	0	1	ventricular conduction	C1orf185, RNF11, CDKN2C,FAF1	21076409
rs2505083	C	0.43	0.97	0.6143	CHD	0	0	coronary heart disease	KIAA1462	23202125
rs2000999	A	0.21	1.03	0.6170	TC	1	1	lipids	-	20686565
rs562338	G	0.84	1.04	0.6206	-	0	1	ldl cholesterol	APOB	18262040
rs13078807	G	0.18	1.03	0.6223	BMI	1	1	body mass index	CADM2	20935630

rs1169288	C	0.32	0.97	0.6227	TC	1	0	lipids	-	20686565
rs2737229	A	0.69	1.03	0.6232	TC	1	1	lipids	-	20686565
rs2972146	T	0.62	1.03	0.6237	HDL	0	0	lipids	-	20686565
rs2932538	G	0.74	0.97	0.6253	SBP	1	1	diastolic blood pressure	MOV10	21909115
rs605066	C	0.43	0.97	0.6268	HDL	1	1	lipids	-	20686565
rs2142672	G	0.71	0.97	0.6287	-	0	1	ldl cholesterol	MYLIP,GMPR	20864672
rs4805834	C	0.88	1.05	0.6289	-	0	1	creatinine levels	SLC7A9	20383145
rs887912	T	0.28	0.97	0.6312	BMI	1	1	body mass index	FANCL	20935630
rs2568958	A	0.58	1.03	0.6328	BMI	1	1	body mass index	NEGR1	19079260
rs9686661	T	0.15	0.96	0.6329	-	0	1	lipids	-	20686565
rs6102059	C	0.73	0.97	0.6346	-	0	1	ldl cholesterol	MAFB	19060906
rs2290159	G	0.78	1.03	0.6374	TC	1	1	lipids	-	20686565
rs581080	G	0.20	0.97	0.6417	HDL	1	1	lipids	-	20686565
rs2794520	C	0.67	1.03	0.6453	-	0	1	c-reactive protein	CRP	21300955
rs1165205	T	0.48	0.98	0.6457	-	0	1	urate levels	SLC17A3	18834626
rs7498665	G	0.42	0.98	0.6485	-	1	1	body mass index	SH2B1, ATP2A1	19079260
rs17514846	A	0.45	1.02	0.6491	CHD	0	1	coronary heart disease	FURIN/FES	23202125
rs17465637	C	0.73	1.03	0.6502	CHD	0	1	coronary heart disease	MIA3	21378990
rs6901250	A	0.31	1.03	0.6520	-	0	1	c-reactive protein	GPRC6A	21300955
rs1495741	G	0.23	1.03	0.6525	-	0	1	lipids	-	20686565
rs206936	G	0.20	1.03	0.6532	BMI	1	1	body mass index	NUDT3, HMGA1	20935630
rs10745954	A	0.43	0.98	0.6553	-	0	1	c-reactive protein	ASCL1	21300955
rs6029526	A	0.49	1.02	0.6568	-	0	1	lipids	-	20686565
rs17020136	C	0.21	1.03	0.6585	-	0	1	ventricular conduction	HEATR5B, STRN	21076409
rs12970134	A	0.28	1.03	0.6586	T2D	0	0	body mass index	MC4R	19079260
rs7578597	T	0.92	0.96	0.6694	T2D	1	1	type 2 diabetes	THADA	18372903
rs11603334	A	0.18	1.03	0.6717	-	0	1	proinsulin levels	ARAP1	21873549
rs2652834	A	0.21	1.03	0.6721	HDL	1	1	lipids	-	20686565
rs8050136	A	0.41	1.02	0.6790	T2D	0	0	adiposity	FTO	21706003
rs693	A	0.53	1.02	0.6808	TC	1	0	cholesterol, total	APOB	19060911
rs243021	A	0.46	1.02	0.6821	T2D	1	1	type 2 diabetes	BCL11A	20581827
rs10821415	A	0.40	0.98	0.6868	-	0	1	atrial fibrillation	C9orf3	22544366

rs3099844	A	0.12	0.97	0.6881	HDL	1	0	metabolic syndrome	HCG26, MICB	22399527
rs10811661	T	0.85	1.03	0.6882	T2D	1	1	type 2 diabetes	CDKN2A,CDKN 2B	17463246
rs1106766	C	0.74	0.97	0.6901	-	0	1	urate levels	R3HDM2,INHBC	20884846
rs2075292	G	0.11	1.03	0.6938	-	0	1	triglycerides	APOA1,KIAA09 99,LOC645044	18193046
rs5756931	T	0.63	1.02	0.6943	-	0	1	lipids	-	20686565
rs1532624	C	0.56	1.02	0.6956	HDL	1	1	hdl cholesterol	CETP	19060911
rs10401969	T	0.89	1.03	0.6959	T2D	1	1	lipids	-	20686565
rs10401969	T	0.89	1.03	0.6959	TC	1	1	lipids	-	20686565
rs2877716	C	0.76	1.02	0.7106	-	0	1	two-hour glucose challenge	ADCY5	20081857
rs2412710	A	0.02	0.93	0.7241	-	0	1	lipids	-	20686565
rs7241918	G	0.19	0.98	0.7255	HDL	1	1	lipids	-	20686565
rs12666989	G	0.79	1.02	0.7255	-	0	1	resting heart rate	UFSP1	20639392
rs439401	C	0.66	0.98	0.7342	-	0	1	triglycerides	TOMM40, APOE	19060911
rs12571751	A	0.53	1.02	0.7358	T2D	1	1	type 2 diabetes	ZMIZ1	22885922
rs12258967	C	0.72	1.02	0.7379	-	0	1	blood pressure	CACNB2	21909110
rs991014	T	0.43	1.02	0.7393	-	0	1	ventricular conduction	SETBP1	21076409
rs340874	C	0.54	1.02	0.7397	-	0	1	fasting glucose-related traits	PROX1	20081858
rs602633	G	0.76	1.02	0.7458	CHD	0	1	coronary heart disease	SORT1	23202125
rs10512597	C	0.86	0.97	0.7489	-	0	1	fibrinogen	CD300LF, SLC9A3R1, NAT9	20031577
rs10838681	G	0.72	0.98	0.7508	HDL	1	0	metabolic syndrome	NR1H3	22399527
rs13165478	G	0.63	1.02	0.7568	-	0	1	ventricular conduction	HAND1,SAP30L	21076409
rs1552224	A	0.82	0.98	0.7576	T2D	1	0	type 2 diabetes	CENTD2	20581827
rs5215	C	0.40	1.02	0.7628	T2D	1	1	type 2 diabetes	KCNJ11	17463249
rs516946	C	0.79	0.98	0.7688	T2D	1	1	type 2 diabetes	ANK1	22885922
rs1046896	T	0.30	1.02	0.7689	-	0	1	glycated hemoglobin levels	FN3K	20858683
rs855791	A	0.43	0.99	0.7838	-	0	1	glycated hemoglobin levels	TMPRSS6	20858683
rs17470137	A	0.28	0.98	0.7850	-	0	1	aortic root size	CCDC100, PPIC	19584346
rs10761731	A	0.60	0.99	0.7856	-	0	1	lipids	-	20686565

rs4402960	T	0.30	1.02	0.7887	T2D	1	1	type 2 diabetes	IG2BP2	22693455
rs1327235	G	0.45	1.01	0.7891	SBP	1	1	blood pressure	JAG1	21909110
rs7944584	A	0.75	0.99	0.7981	-	0	1	fasting glucose-related traits	MADD	20081858
rs2867125	C	0.83	0.98	0.8080	BMI	1	1	body mass index	TMEM18	20935630
rs452036	G	0.63	1.03	0.8081	-	0	1	resting heart rate	MYH6	20639392
rs12444979	C	0.88	0.98	0.8081	BMI	1	1	body mass index	GPRC5B, IQCK	20935630
rs998584	A	0.47	0.99	0.8088	-	0	1	adiponectin levels	VEGFA	22479202
rs12708716	A	0.67	1.01	0.8156	-	0	1	type 1 diabetes	KIAA0350	17554260
rs2650000	A	0.38	0.99	0.8164	-	0	1	ldl cholesterol	HNF1A	19060906
rs7998202	G	0.14	0.98	0.8174	-	0	1	glycated hemoglobin levels	ATP11A,TUBGCP3	20858683
rs4502156	T	0.53	0.99	0.8219	-	0	1	proinsulin levels	VPS13C, C2CD4A, C2CD4B	21873549
rs12946454	T	0.25	0.99	0.8269	SBP	1	1	systolic blood pressure	PLCD3, ACBD4, HEXIM1, HEXIM2	19430483
rs3136441	T	0.86	0.98	0.8302	HDL	1	1	lipids	-	20686565
rs7593730	T	0.21	1.01	0.8338	T2D	1	1	type 2 diabetes	RBMS1, ITGB6	20418489
rs7134375	C	0.57	1.01	0.8378	HDL	1	1	lipids	-	20686565
rs5015480	C	0.55	0.99	0.8402	T2D	1	1	type 2 diabetes	HHEX	22693455
rs2075650	A	0.85	0.99	0.8465	TC	1	1	cardiovascular disease risk factors	TOMM40	21943158
rs2075650	A	0.85	0.99	0.8465	CHD	1	1	cardiovascular disease risk factors	TOMM40	21943158
rs6265	C	0.82	0.99	0.8500	BMI	1	1	body mass index	BDNF	19079260
rs6265	C	0.82	0.99	0.8500	Smoke	1	1	body mass index	BDNF	19079260
rs17477177	C	0.22	1.01	0.8504	-	0	1	blood pressure	PIK3CG	21909110
rs3757354	C	0.74	1.01	0.8510	-	0	1	lipids	-	20686565
rs4129767	G	0.48	1.01	0.8515	HDL	1	1	lipids	-	20686565
rs2681472	A	0.87	1.02	0.8519	-	0	1	diastolic blood pressure	ATP2B1	19430479
rs646776	T	0.76	0.99	0.8571	TC	1	0	cholesterol, total	CELSR2	19060911
rs2083637	A	0.75	1.01	0.8609	HDL	1	1	hdl cholesterol	LPL	19060911
rs2277862	C	0.89	1.01	0.8609	TC	1	1	lipids	-	20686565
rs6810075	C	0.36	1.01	0.8633	-	0	1	adiponectin levels	ADIPOQ	22479202
rs838880	T	0.65	0.99	0.8641	HDL	1	1	lipids	-	20686565

rs864745	T	0.52	1.01	0.8641	T2D	1	1	type 2 diabetes	JAZF1	18372903
rs2023938	C	0.10	0.99	0.8686	CHD	0	0	coronary heart disease	HDAC9	23202125
rs974819	T	0.25	0.99	0.8696	CHD	0	1	coronary heart disease	PDGFD	23202125
rs11847697	T	0.04	0.98	0.8729	BMI	1	1	body mass index	PRKD1	20935630
rs13226650	A	0.85	0.98	0.8745	-	0	1	metabolic syndrome	MLXIPL	22399527
rs1458038	T	0.37	1.01	0.8798	SBP	1	1	blood pressure	FGF5	21909110
rs947474	G	0.18	1.01	0.8799	-	0	1	type 1 diabetes	PRKCQ	18978792
rs10946398	C	0.31	0.99	0.8806	T2D	1	1	type 2 diabetes	CDKAL1	17463249
rs2287019	C	0.78	1.01	0.8813	BMI	1	1	body mass index	QPCTL, GIPR	20935630
rs11920090	T	0.87	1.01	0.8814	-	0	1	fasting glucose-related traits	SLC2A2	20081858
rs13333226	A	0.82	1.01	0.8816	-	0	1	hypertension	UMOD	21082022
rs2895811	C	0.46	1.01	0.8819	CHD	0	1	coronary heart disease	HHIPL1	23202125
rs13266634	C	0.56	1.01	0.8823	T2D	1	0	glycated hemoglobin levels	SLC30A8	19096518
rs4665058	A	0.02	1.03	0.8828	-	0	1	sudden cardiac arrest	BAZ2B	21738491
rs264	G	0.87	0.99	0.8879	CHD	0	1	coronary heart disease	LPL	23202125
rs579459	C	0.23	0.99	0.8917	CHD	0	1	coronary heart disease	ABO	23202125
rs8017377	A	0.47	0.99	0.8944	-	0	1	lipids	-	20686565
rs7134594	C	0.47	0.99	0.8947	HDL	1	1	lipids	-	20686565
rs13389219	C	0.58	0.99	0.8967	T2D	1	1	type 2 diabetes	GRB14	22885922
rs16890979	C	0.76	1.01	0.8989	-	0	1	urate levels	SLC2A9	18834626
rs2292239	T	0.33	0.99	0.8993	-	0	1	type 1 diabetes	ERBB3	17554260
rs7515577	A	0.77	0.99	0.9016	TC	1	1	lipids	-	20686565
rs1329650	G	0.63	0.99	0.9069	Smoke	1	1	smoking behavior	LOC100188947	20418890
rs7129220	A	0.10	0.99	0.9085	SBP	1	1	systolic blood pressure	ADM	21909115
rs2293889	T	0.43	0.99	0.9094	HDL	1	1	lipids	-	20686565
rs17608766	C	0.14	1.01	0.9149	SBP	1	1	blood pressure	GOSR2	21909110
rs10885122	G	0.88	1.01	0.9158	-	0	1	fasting glucose-related traits	ADRA2A	20081858
rs13407662	T	0.02	0.98	0.9263	-	0	1	stroke (ischemic)	Intergenic	23041239
rs6544713	T	0.28	0.99	0.9268	CHD	0	1	coronary heart disease	ABCG5/ABCG8	23202125

rs11953630	C	0.71	1.01	0.9287	SBP	1	1	diastolic blood pressure	EBF1	21909115
rs2590838	G	0.49	1.00	0.9316	-	0	1	adiponectin levels	GNL3	22479202
rs6499640	A	0.61	1.00	0.9317	BMI	1	1	body mass index	FTO	19079260
rs11708996	C	0.14	1.01	0.9349	-	0	1	pr interval	SCN5A	20062060
rs6756629	G	0.95	0.99	0.9363	TC	1	1	cholesterol, total	ABCG5	19060911
rs1260326	T	0.34	1.00	0.9429	-	0	1	c-reactive protein	GCKR	21300955
rs6679677	A	0.12	1.01	0.9433	-	0	1	type 1 diabetes	PHTF1, PTPN22	17554260
rs1367117	A	0.34	1.00	0.9477	-	0	1	lipids	-	20686565
rs2281727	G	0.37	1.00	0.9528	CHD	0	1	coronary heart disease	SMG6	23202125
rs3807989	G	0.58	1.00	0.9583	-	0	1	atrial fibrillation	CAV1	22544366
rs9349379	G	0.44	1.00	0.9594	-	0	1	coronary artery calcification	PHACTR1	22144573
rs1799945	G	0.12	1.00	0.9640	SBP	1	1	diastolic blood pressure	HFE	21909115
rs1530440	C	0.81	1.00	0.9654	-	0	1	diastolic blood pressure	c10orf107, TMEM26, RTKN2, RHOBTB1, ARID5B	19430483
rs13107325	C	0.96	1.01	0.9720	SBP	1	1	blood pressure	SLC39A8	21909110
rs2681492	T	0.87	1.00	0.9738	SBP	1	0	systolic blood pressure	ATP2B1	19430479
rs2255141	A	0.27	1.00	0.9780	TC	1	1	lipids	-	20686565
rs9816226	T	0.84	1.00	0.9797	BMI	1	1	body mass index	ETV5	20935630
rs1800789	A	0.15	1.00	0.9812	-	0	1	fibrinogen	FGB	20031576
rs10128711	C	0.72	1.00	0.9815	TC	1	1	lipids	-	20686565
rs1689800	G	0.33	1.00	0.9840	HDL	1	1	lipids	-	20686565
rs871606	T	0.92	1.00	0.9863	-	0	1	blood pressure	CHIC2	21909110
rs12356193	A	0.85	1.00	0.9902	-	0	1	uric acid levels	SLC16A9	19503597
rs1531343	C	0.08	1.00	0.9974	T2D	1	1	type 2 diabetes	HMGA2	20581827
rs9815354	A	0.15718055	0.999963663	0.99958085	-	0	1	diastolic blood pressure	ULK4	19430479

[a] Cox proportional hazard analyses with age as timescale and adjusted for sex, systolic blood pressure, antihypertensive treatment, diabetes, current smoking, total cholesterol, body max index and study of origin.

[b] Value equal to 1 indicates that the SNP has been included in the FHS trait-specific risk score.

[c] Value equal to 1 indicates that the SNP has been included in the overall MGRS.

Supplementary Table III. Net Benefit calculations.

Model	N. treated[a]	N. events treated[a]	10-year risk among treated	Benefit[b]	Costs[b]	Net benefit[b]	k[c]
Base model	1242	237	28%	62.96	46.45	16.51	2.13%
Base model + CHD-specific MGRS	1253	243	29%	66.22	46.56	19.66	2.13%

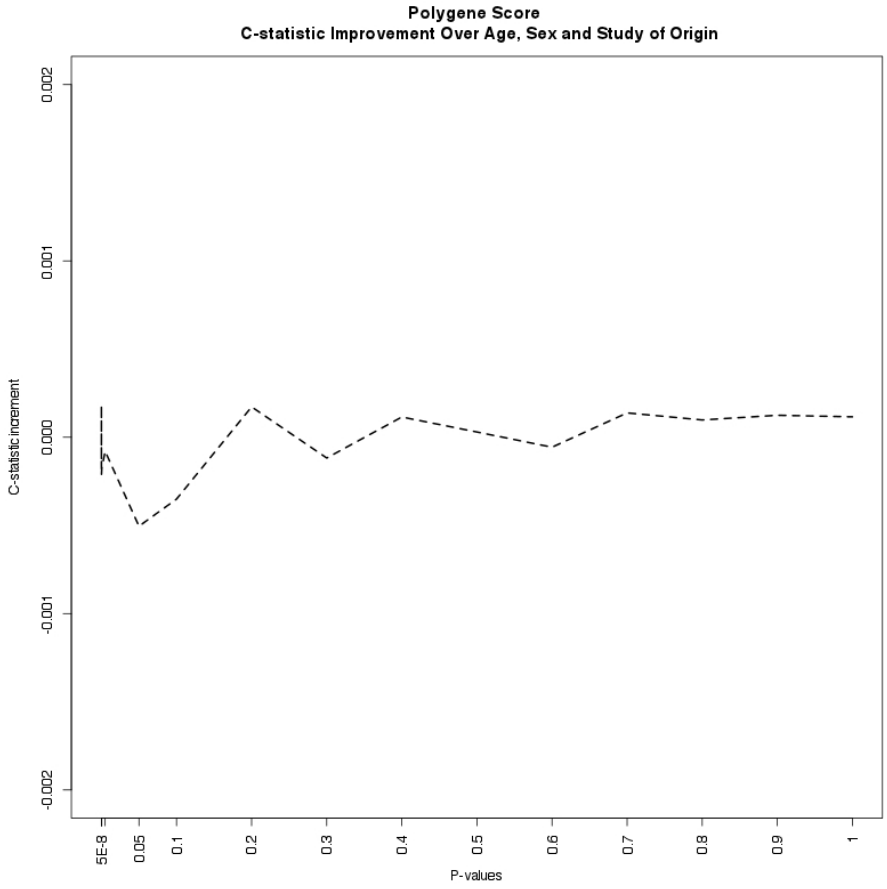
[a] Individuals are treated if 10-year risk is higher than 20%.

[b] In EFLYs per 1,000 screened.

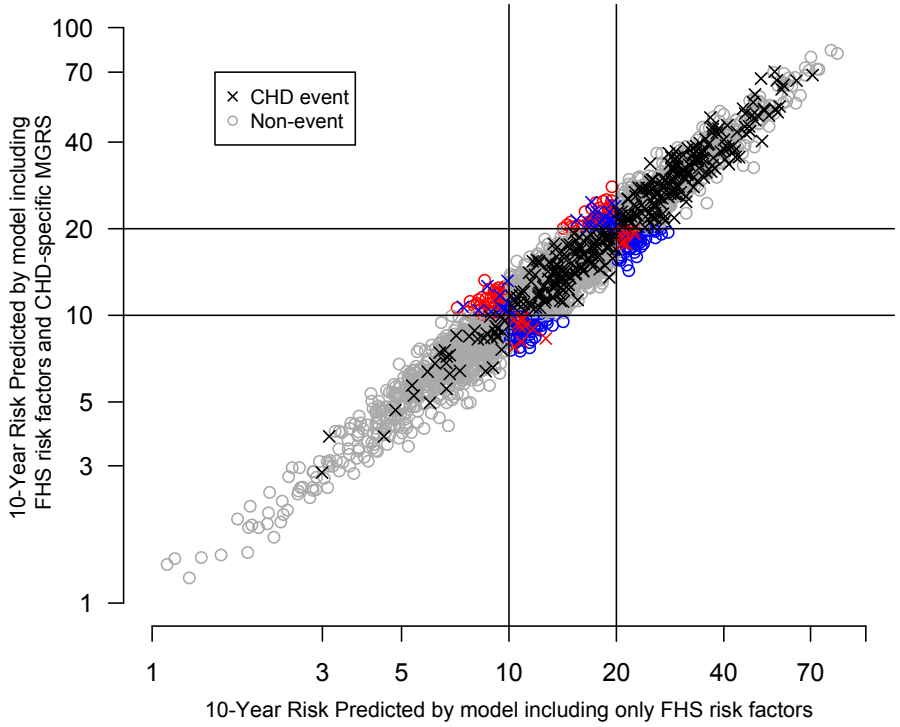
[c] Treatment cost relative to EFLY cost.

SUPPLEMENTAL FIGURES

Supplementary Figure I. Polygene Approach. We first studied the association between SNPs and CHD in the WTCCC study. We used these results to rank the SNPs up to different P-values thresholds (*x-axis*) obtaining 17 different lists of SNPs, risk alleles and association coefficients (OR). In our study population, we summed the number of risk alleles for the corresponding SNPs in each list weighting for the log(OR) and obtaining 17 polygene scores. We added each polygene score to a logistic model for association of age, sex and study of origin with CHD and we calculated the difference in C-statistics (*y-axis*).



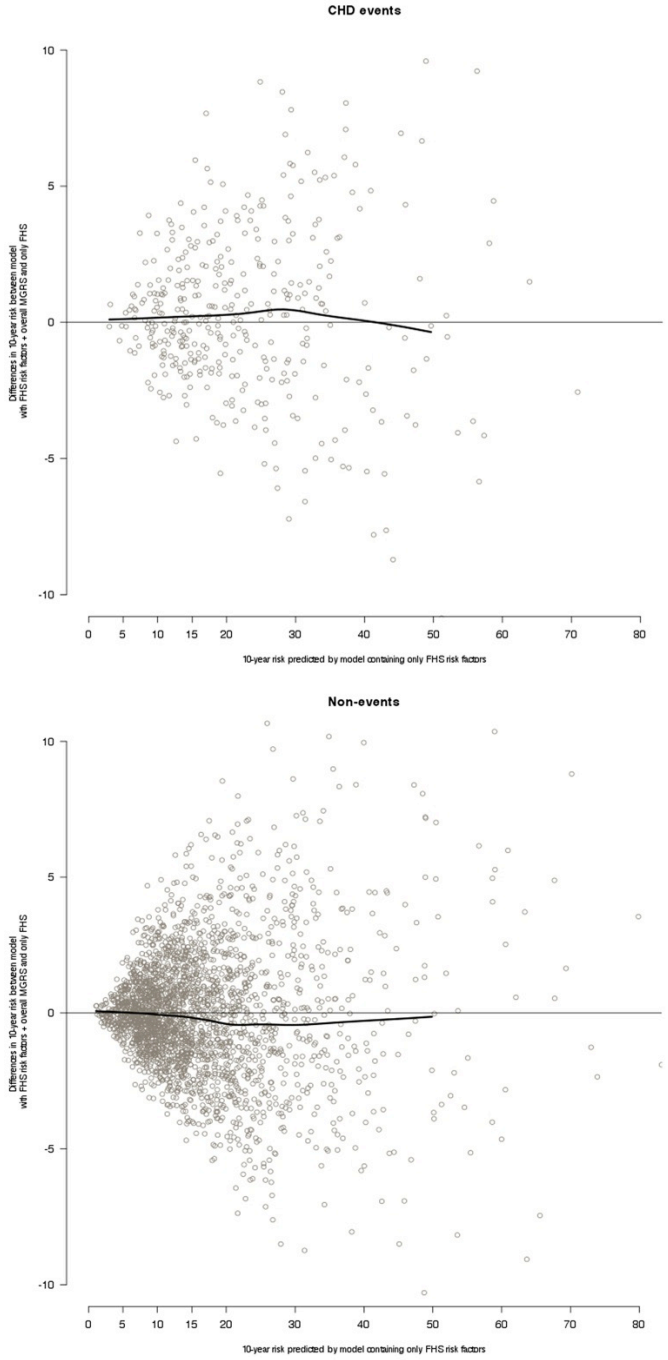
Supplemental Figure II. Reclassification Graph and Reclassification Table Representing the Individual Risk at 10 Years for a Model with FHS Risk Factors *vs.* Individual Risk at 10 Years for a Model with FHS Risk Factors + CHD-specific MGRS. The Axes are on Log-scale. Blue marks are correctly reclassified, red marks are incorrectly reclassified between categories of clinical interest (<10%, 10% to 20%, >20%). The TwinGene study was not included in these analyses due to short follow-up.



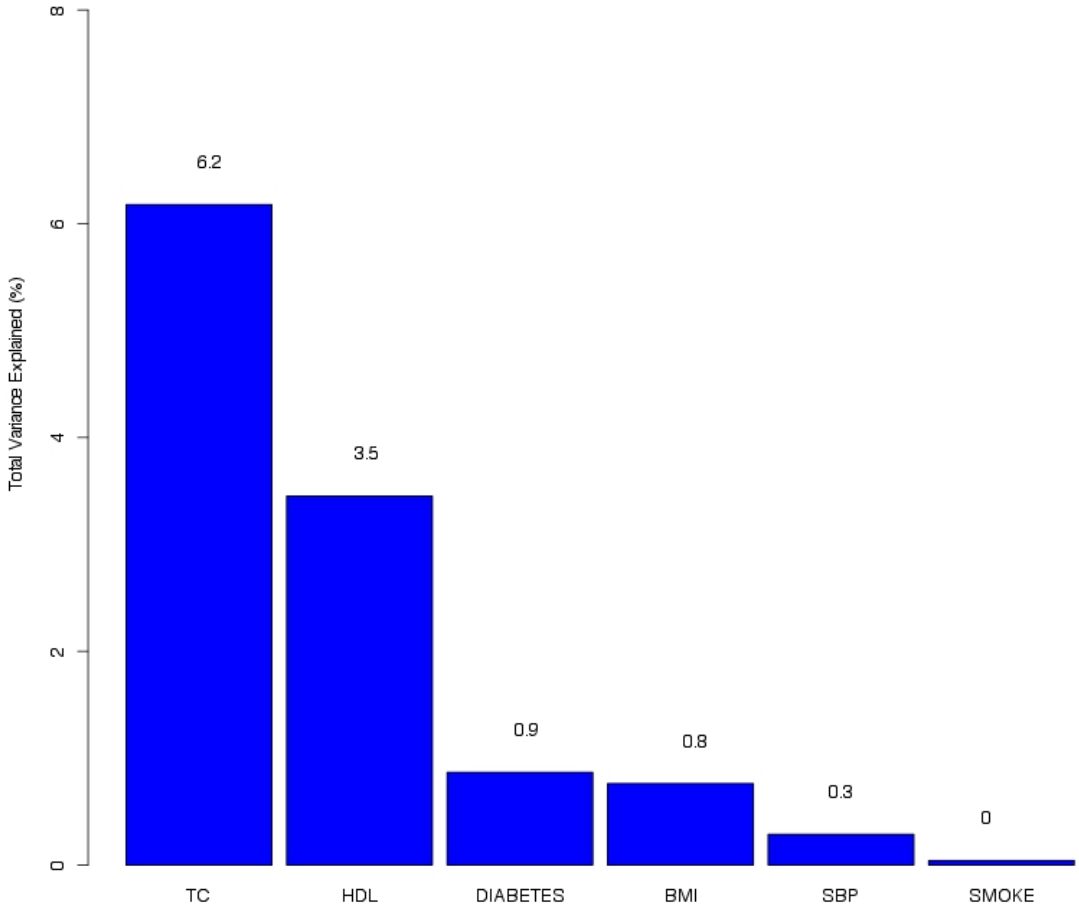
Original risk categories	**Reclassification**		
	<10%	10% to 20%	>20%
	number of participants		
Participants without CHD events			
<10%	752	84	0
10% to 20%	109	946	63
>20%	0	92	580
Participants with CHD events			
<10%	36	10	0
10% to 20%	11	123	20
>20%	0	8	180

Net reclassification for participants without CHD: 54 of 2626 (2.1%).
Net reclassification for participants with CHD: 11 of 388 (2.8%).

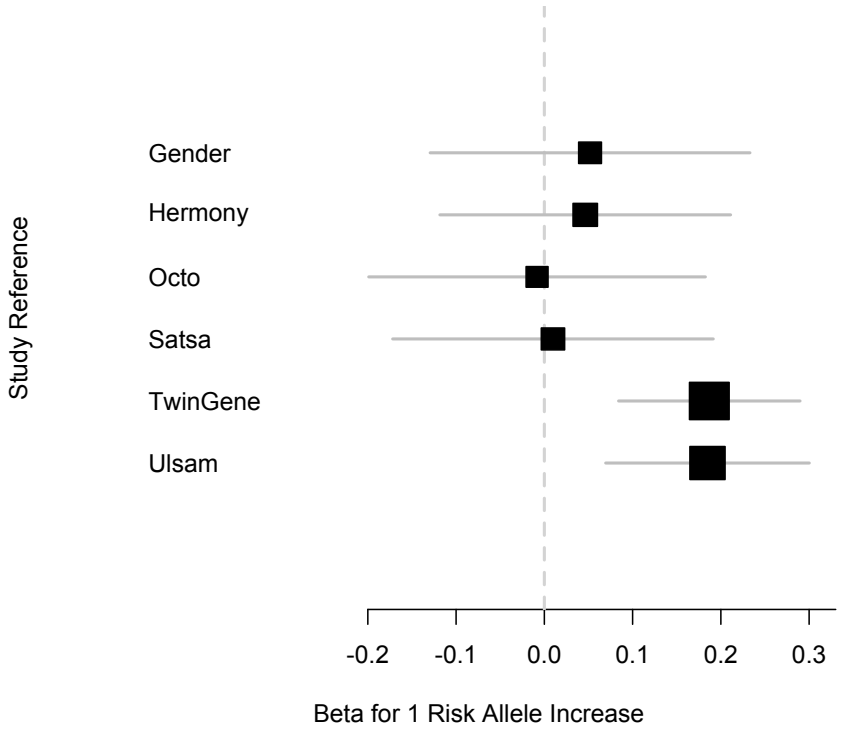
Supplementary Figure III. Scatter Plot of the Individual Risk Without Overall MGRS (*x-axis*) vs. Difference between Individual Risk With and Without the Overall MGRS (*y-axis*), and Estimated Lowess Curve. The lowess curve provides an intuitive representation of the data trend. In the top panel we can observe an enrichment of positive values indicating an increment of the individual risk after the inclusion of the overall MGRS in the model. In the bottom panel we can observe a graphically detectable decrement of the individual risk only in participants with individual risk greater than 15%.



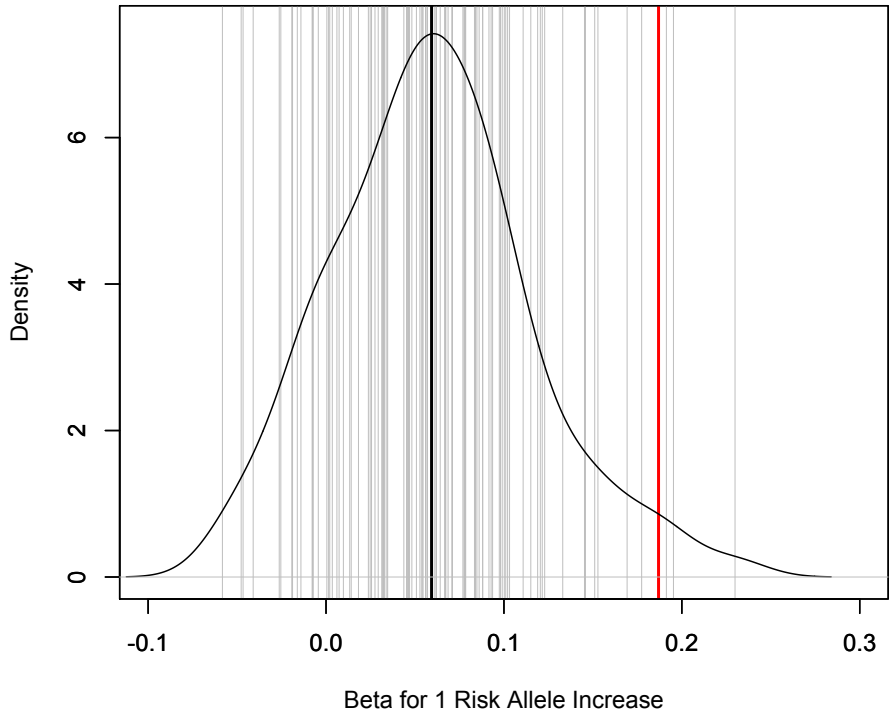
Supplementary Figure IV. Proportion of Total Variance Explained for FHS Risk Factors and BMI. We calculated several trait-specific genetic scores (**Supplementary Table 2**) and for each score (*x-axis*) we evaluated the proportion of total variance explained (*y-axis*). This was calculated as the improvement in the R^2 after including each trait-specific score over a logistic regression model (for dichotomous phenotypes) or linear regression model (for continuous phonotypes) including age, sex and study of origin. In HDL and TC analyses, we adjusted also for lipid lowering treatment.



Supplementary Figure V. Association between overall MGRS and CHD for each study population.



Supplementary Figure VI. Distribution of betas for association with CHD for 100 random overall MGRSs (grey lines). Black thick line: average of the betas obtained from the 100 random overall MGRSs. Red Line: betas from the original overall MGRS. Each random overall MGRS was a mixture of two components: (1) SNPs directly associated with CHD; and (2) randomly selected SNPs with similar (+/- 0.01) allele frequency as in the original overall MGRS. Scores were calculated as the weighted sum of risk alleles and weights are obtained from WTCCC. TwinGene was used to calculate these scores and to study the association with CHD.



MATERIALS AND METHODS

Study populations

The Swedish Twin Registry is a population-based national register including over 170,000 Swedish twins born from 1886 to 2000 [1]. For this study, we included five separate sub-studies with availability of DNA and data on cardiovascular risk factors that have been conducted within the registry (SATSA, OCTO-Twin, GENDER, HARMONY and TwinGene). We also included the Uppsala Longitudinal Study of Adult Men (ULSAM) [2], which is an ongoing, longitudinal, epidemiologic study based on men born between 1920 and 1924 in Uppsala County, Sweden. The total number of individuals with DNA and data on cardiovascular risk factors when combining these studies was 12,566. After exclusion of individuals without complete data on FHS variables or BMI (N=1,163) or with CHD events prior to the baseline examination (N=859), 10,612 individuals remained eligible for the present study. Participants excluded because of missing data had a similar risk factors distribution compared to those included in the analysis (data not shown). All participants gave informed written consent and the Ethics Committees of Karolinska Institutet or Uppsala University approved the study.

We collected information on FHS risk factors [3] (age, sex, anti-hypertensive treatment, total cholesterol [TC], HDL-cholesterol [HDL], systolic blood pressure [SBP], diabetes prevalence [T2D] and current smoking), body mass index (BMI) and lipid-lowering treatment from all participants who were free of prior CHD at the baseline examination.

In order to assess the weights for the MGRS and to develop the polygene MGRS, we used 1,972 CHD cases and 2,891 controls from the Wellcome Trust Case-Control Consortium (WTCCC) [4].

Detailed study descriptions

SATSA

The SATSA sample comprises all pairs of twins who indicated that they had been separated before the age of 11 and reared apart, and a sample of twins reared together matched on gender and date and county of birth [5]. SATSA twins aged 50 and older were invited to participate in in-person testing (IPT) sessions in which questionnaires including items concerning health-related behaviors (e.g. alcohol, tobacco, and dietary habits), cognitive tests and physical health measures were administrated. SATSA twins have been followed longitudinally with up to nine IPT sessions across 24 years. In this study, we used blood samples for DNA extraction, lipid measurements and other cardiovascular risk factors from the third IPT (1992-1994). In total, 569 participants had information on Framingham Heart Study (FHS) risk factors available at the third IPT, and 518 of them were successfully genotyped. For the purpose of the present study, we excluded 83 subjects with any missing FHS risk factor or body mass index (BMI; N=64), or with previous CHD event (N=20). Hence, 435 individuals were included in the analyses.

OCTO-Twin

The OCTO-Twin sample included all twin pairs in Sweden aged 80 years or older in 1991-1994 (i.e. birth years 1913 or earlier) [6]. Up to five waves of IPT sessions at 2-year intervals were conducted on all living twins who agreed to participate, irrespective of co-twin's vital status. Blood samples for subsequent extraction of DNA were collected between the first and second IPT. Questionnaires were similar to the ones used in SATSA. Cardiovascular risk factors were measured at the first IPT (1991-1994). A total of 702 participants participated at the first IPT, and 506 of them were successfully genotyped. We excluded 96 subjects with missing at least one FHS risk factors or BMI (N=76), or with previous CHD event (N=25). Thus, we included 410 subjects in the analyses.

GENDER

All living pairs of unlike-sex twins born between 1906 and 1925 were identified through the Swedish Twin Registry in 1995 [7]. Twin pairs were sent surveys assessing health and other factors. A subset of this population-based sample aged 70-79 years completed IPT similar to those in SATSA and OCTO-Twin. Three waves of IPT sessions were carried out at four years intervals during 1995-2004 Blood samples for DNA extraction and cardiovascular risk factors were collected during the first IPT session (1995-1997). A total of 496 participants participated at the first IPT, and 467 of them were successfully genotyped. We excluded 46 subjects with missing at least one FHS risk factors or BMI (N=16), or with previous CHD event (N=31), resulting in 421 subjects eligible for the present analyses.

HARMONY

All twins from the Swedish Twin Registry aged 65 and older were screened by telephone for cognitive dysfunction. This included any surviving twins from the SATSA, OCTO- Twin, and GENDER studies described above. Among those screened, 11.5% were positive for suspicion of dementia and were referred for complete clinical evaluation and blood sampling by a physician or a nurse (1999-2001). Once the preliminary IPT suggested dementia, the twin partner was also invited for an identical clinical work-up [8]. We considered 936 participants not recruited in SATSA, OCTO or GENDER with genotype information and at least one FHS risk factor available. We excluded 169 subjects missing at least one FHS risk factor or BMI (N=81), or with previous CHD event (N=99). Hence, 767 individuals were included in our analyses.

TWINGENE

Twins born before 1958 and who has participated in a telephone screening between 1998 and 2002 were re-contacted between April 2004 and December 2008. Health and medication data were collected from self-reported questionnaires, and a blood

sampling kit was mailed to the subject who then contacted a local health care center for blood sampling and a health check-up. A total of 10,946 twins had at least one FHS risk factor available. All dizygotic twins and one random monozygotic twin per pair were genotyped resulting in 9,023 twins with genotype information available and not recruited in the other twin studies previously described. Further, we excluded 1,426 individuals with missing FHS risk factors or BMI (N=894), or with previous CHD event (N=578) resulting in 7,597 subjects included in the present analyses.

ULSAM

Subjects born between 1920 and 1924 in Uppsala, Sweden were invited to participate at age 50 in this longitudinal cohort study, which was started in 1970{Byberg, 1998 #11}. Subjects were reinvestigated at the ages of 60, 70, 77, 82 and 88 years. Information collected includes a medical questionnaire, blood pressure and anthropometric measurements, glucose tolerance test and 24-hour ambulatory blood pressure. Blood samples for DNA extraction and determination of established cardiovascular risk factors were available from the investigation at 70 years of age (1991-1994). We considered 1,116 participants with genotype information and at least one FHS risk factor available. We excluded 33 subjects with missing FHS risk factors or BMI, and 102 with previous CHD event resulting in 981 individuals included in the present analyses.

WTCCC

The WTCCC samples {, 2007 #13} included in this paper comprise two studies that were used as controls (1958 Birth Cohort and UK Blood Services Controls) and 2,000 coronary heart disease (CHD) cases. The 1958 Birth Cohort includes all births in England, Wales and Scotland, during one week in 1958. In a biomedical examination at ages 44-45, 9,377 cohort members were visited at home and 7,692 of them donated blood samples. DNA samples were extracted from 1,500 cell lines of self-reported white ethnicity and representative proportions of gender and geographical regions were selected for use as controls. The UK Blood Services Controls consisted of 1,500 individuals selected from a sample of blood donors recruited as part of WTCCC. Samples were selected based on sex and geographical region (to reproduce the distribution of the samples of the 1958 Birth Cohort). CHD cases recruitment was carried out on a national basis in the UK through a direct approach to the public via (1) the media; and (2) mailing all general practices (family physicians) with information about the study. After QC and imputation as described below, 1,972 CHD cases and 2,891 controls were analyzed in order to assess weights for the MGRS and to develop the polygene score. Associations between SNPs and CHD were assessed using additive logistic models adjusted for age, sex and geographical region.

CHD assessment

Information regarding incident CHD in the Swedish epidemiological studies (SATSA, OCTO-Twin, GENDER, HARMONY, TWINGENE, ULSAM) was collected from the Swedish National In-Patient Register and the Cause of Death Register, and was

defined as hospitalization or death with any of the following primary diagnoses: acute myocardial infarction and unstable angina (ICD-10: I20.0, I21, I22; ICD-9: 410, 411B; ICD-8: 410, 411 and surgical codes: FNG02, FNG05, FNC, FND, FNE). These diagnoses were defined according to the primary diagnosis or the primary cause of death, as recorded in the Patient and Cause of Death Registers. The Patient Register includes hospitalized cases, as well as outpatient visits, but not visits to the primary care. The positive predictive value (i.e. validity) of the myocardial infarction diagnosis in the Swedish Patient register has been demonstrated to be 95% when only primary diagnoses are considered [9]. We recorded the follow-up time from the baseline examination to the date of the first CHD event (as defined above), emigration from Sweden, death or end of follow-up (Dec 31, 2010), whichever occurred first. CHD cases in the WTCCC study had a validated history of either myocardial infarction or coronary revascularization (coronary artery bypass surgery or percutaneous coronary angioplasty) before their 66th birthday.

Genotyping

Blood samples from participants enrolled in SATSA, OCTO-twin, GENDER, HARMONY and ULSAM were genotyped using the CardioMetabochip, a custom Illumina iSelect genotyping array designed to analyze cost-effectively ≈185,000 SNPs identified through GWA study meta-analyses of cardiovascular and metabolic traits. TwinGene participants were genotyped with Illumina Human OmniExpress (≈700,000 SNPs) and WTCCC samples using Affymetrix 500K (≈500,000 SNPs). Genotyping, except for WTCCC, was performed at the Uppsala University SNP Technology Platform (www.genotyping.se). All the samples underwent the same quality control (QC), and imputation of polymorphic HapMap CEU SNPs (release 22) was performed using IMPUTE2 [10].

Imputation and SNP quality control

All samples (including the WTCCC samples which were used for calculating weights) underwent the same quality control and imputation procedures. Specifically, we used these inclusion criteria:

- Minor allele frequency greater than 0.01.
- Hardy-Weinberg Equilibrium test P-value greater than $1*10^{-7}$.
- Genotype rate per SNP greater than 95%.
- Genotype rate per individual greater than 97%.
- Reported sex equal to sex identified using the X chromosome.

SNPs that were not directly genotyped were imputed using IMPUTE2 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html) and HapMap (http://hapmap.ncbi.nlm.nih.gov) CEU build 36 release 22 as reference panel

In samples genotyped with CardioMetabochip we observed a total of twenty-three SNPs with low imputation quality scores (information metric <0.4 from IMPUTE2). When calculating the genetic risk scores in individuals from these samples, we set the number of alleles equal to the average number of risk alleles in TWINGENE. This

was done because poor imputation quality was observed only in the CardioMetabochip samples, and we did not want to exclude these SNPs when creating genetic scores for all the individuals in the study. This is likely to drive our findings towards the null due to the random error introduced, but not to cause any systematic bias to our results.

Finally, within each score, we pruned SNPs to be non-correlated ($r^2 < 0.2$) within a distance of 250kb, and summed the number of risk alleles. For imputed SNPs, where the exact number of risk alleles was not available, we used the posterior genotype probability from IMPUTE2 to calculate a proxy for this quantity. The proxy for the number of risk alleles in imputed SNPs ranges between 0 and 2 and corresponds to the sum of the risk alleles weighted by the corresponding posterior probability.

SNP selection procedures

We selected *a priori* 119 traits that are related to CHD in a direct or broader sense (**Supplementary Table I**) from the NGHRI catalog up to the 14^h January 2013. Within the selected studies, we manually checked ethnicities reported in the catalog and we excluded studies containing one of the following words in the 'Initial_Sample_Size' variable: *Japanese, Chinese, African, Filipino, Micronesian, Indian, Asian, Korean, Pima Indians, Kosraean.* If the study reported one of the aforementioned ethnicities together with the words *Caucasian* or *European*, we manually checked the article to identify whether the number of Caucasian subjects was higher than the number of non-Caucasians. When this was the case, we chose to include the study. We further excluded studies where the number of subjects in the initial sample size and in the replication sample size combined was lower than 10,000. Furthermore, we excluded single nucleotide polymorphisms (SNPs) with missing P-values or P-values $>5*10^{-8}$. We manually checked that the SNP names were formally annotated according to the dbSNP annotation (i.e. with an rs-id). Deviations from this format were manually curated. Some studies, where more than one phenotype was investigated, reported the exact phenotype in the 'p_Value__text_' variable. We extracted this information and assigned the appropriate phenotype. When possible, missing risk alleles (RA) or missing allele frequencies were extracted manually from the original paper; if not possible, we excluded the study. We checked whether the reported SNPs were aligned to the plus strand, and whether the reported reference allele (the other allele is not reported in the catalog) and allele frequency were reasonable when compared with HapMap. To do this, we used BioMart (http://hapmap.ncbi.nlm.nih.gov/biomart/martview/3e9af05f2a67fe20a2302eb8c3a8a5b3) to annotate each SNP using HapMap release 27 and the following algorithm:

- If the reported alleles in HapMap were one of the combinations "C/T","A/G","A/C","G/T" (or their complement) and the reference allele in the catalog corresponded to one of the two alleles in HapMap; then the reference allele was considered to be correctly reported.
- If the alleles in HapMap were one of the combinations "C/T","A/G","A/C","G/T" and the allele in the catalog corresponded to the complementary strand alleles, then we assumed them to be correct, but on the minus strand (if not reported specifically).

- If the alleles in HapMap are one of these combination: "A/T","C/G" and the allele in the catalog corresponded to one of the two alleles in HapMap with similar allele frequency (±0.10 acceptance interval); then the allele annotation was considered to be correct.
- If the alleles in HapMap were one of the combinations "A/T","C/G" and allele frequencies were between 0.40 and 0.60, then a manual check of the article was required.
- All other combinations and differences in allele frequencies ±0.10 were manually checked against the original paper. If we could not come to a definitive conclusion, we excluded the paper.

Further, we checked whether the same SNP was reported on more than one occasion for different phenotypes with different risk alleles; this was the case for 8 SNPs. We kept the reference allele corresponding to the phenotype that was had stronger association with CHD.

Statistical Analysis

We tested the association between MGRS and incident CHD using a Cox proportional hazard model, adjusting for FHS risk factors and study of origin using age as timescale. The proportionality of hazards was confirmed both by using a statistical test and visual inspection of the Schoenfeld residuals distribution. We evaluated the clinical utility of MGRS in terms of discrimination, reclassification and calibration based on the 10-year individual risk of CHD. For these analyses, we excluded the TwinGene study due to short follow-up (median follow-up: 3.9 years, IQR: 3.4 – 5.1), and utilized the remaining 5 cohorts (median follow-up: 10.1 years, IQR: 5.3 - 15.0, N of CHD events during 10 years = 388).

To estimate the improvement in discrimination, we compared the C-indexes, for a model with and without MGRS. Risk reclassification was evaluated with the net reclassification improvement [11] (categories of <10%, 10% to 20%, >20%, as discussed in ATP-III [12]). Calibration, the comparison between the predicted and observed number of events, was assessed with the Grønnesby and Borgan goodness-of-fit-test, using the implementation proposed by May and Hosmer[13]. We studied the association of each MGRS with established cardiovascular risk factors (BMI, HDL-cholesterol, systolic blood pressure, total cholesterol, smoking, type 2 diabetes) and individual risk at 10 years using a linear or logistic regression adjusted by age, sex and study of origin. Further adjustment for lipid-lowering treatment was done for HDL-cholesterol and total cholesterol outcomes. The proportion of total variance explained by each trait-specific MGRS for association with the corresponding traits was calculated as the improvement in the R^2 over the basic model.

The number of event-free life years saved per 1,000 people screened was calculated using the approach proposed by Rapsomaniki and colleagues [14]. This method requires only information about the risk threshold for assigning an individual to treatment (e.g. 20% of 10-year risk of CHD according to ATP-III) and the treatment effect (e.g. 20% reduction in cardiovascular risk for statin treatment[15]). The number of event-free life years saved was calculated as difference between the net benefits of a model

including only established risk factors and established risk factors + CHD-specific genetic risk score. Confidence intervals were estimated through bootstrapping.

In our analysis, we randomly considered only one twin per pair when monozygous twins were concordant. For discordant monozygous pairs, we selected the twin with CHD, as this genome should be considered to be consistent with experiencing CHD. To account for the correlated structure of the data resulting from considering both members of dizygous twins pairs, we used robust variance estimates (Huber-White method for clustered samples[16]). All statistical analyses were performed using R statistical package (version 2.13) or Plink version 1.07[17].

Calculation of event-free-life-years

Event-free-life-years (EFLY) are calculated as described in the paper by Rapsomaniki and colleagues [14]. Briefly, suppose that *S(u)* is the probability of surviving to time *u* under no treatment. *S(u, θ) is the* probability of survival to time *u* using a treatment with hazard ratio *θ* (this can obtained form clinical trials, e.g. 0.8 for statins treatment). The benefit of using the treatment can be calculated as:

$$B(T) = \int_0^T [S(u,\theta) - S(u)]du$$

where *T* is the time-horizon, e.g. 10-years.

The cost of using the treatment:

$$C(T) = k\int_0^T [S(u,\theta)]du$$

where *k* is the cost of treatment per year relative to the value of one EFLY.

In practice, *k* can be calculated using the risk cutpoint (*c*) for prescribing the treatment, as specified by current guidelines; e.g. 20% risk for statins prescription. Therefore, assuming that *c* is an optimal cutpoint, in the sense that benefits equal costs when the risk is equal to *c, k* can be calculated as*:*

$$k = 1 - \theta\frac{1 - e^{-\lambda T}}{1 - e^{-\lambda\theta T}}$$

where

$$\lambda = -(\log(1-c)/T)$$

Then the net benefit (NB) can be calculated as:

$$NB(T) = P[B(T) - C(T)]$$

where *P* is the proportion of those treated among the screened.

Finally, the EFLY gained by using a new prediction model (NB_2) instead of the established prediction model (NB_1) is:

$$EFLY\_gained = NB_2(T) - NB_1(T)$$

REFERENCES

1. Lichtenstein P, De Faire U, Floderus B, Svartengren M, Svedberg P, Pedersen NL. The swedish twin registry: A unique resource for clinical, epidemiological and genetic studies. *J Intern Med*. 2002;252:184-205
2. Byberg L, Siegbahn A, Berglund L, McKeigue P, Reneland R, Lithell H. Plasminogen activator inhibitor-1 activity is independently related to both insulin sensitivity and serum triglycerides in 70-year-old men. *Arterioscler Thromb Vasc Biol*. 1998;18:258-264
3. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97:1837-1847
4. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447:661-678
5. Pedersen NL, Friberg L, Floderus-Myrhed B, McClearn GE, Plomin R. Swedish early separated twins: Identification and characterization. *Acta geneticae medicae et gemellologiae*. 1984;33:243-250
6. Finkel D, Reynolds CA, McArdle JJ, Hamagami F, Pedersen NL. Genetic variance in processing speed drives variation in aging of spatial and memory abilities. *Developmental psychology*. 2009;45:820-834
7. McClearn GE, Johansson B, Berg S, Pedersen NL, Ahern F, Petrill SA, Plomin R. Substantial genetic influence on cognitive abilities in twins 80 or more years old. *Science*. 1997;276:1560-1563
8. Gold CH, Malmberg B, McClearn GE, Pedersen NL, Berg S. Gender and health: A study of older unlike-sex twins. *The journals of gerontology. Series B, Psychological sciences and social sciences*. 2002;57:S168-176
9. Ludvigsson JF, Andersson E, Ekbom A, Feychting M, Kim JL, Reuterwall C, Heurgren M, Olausson PO. External review and validation of the swedish national inpatient register. *BMC Public Health*. 2011;11:450
10. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*. 2009;5:e1000529
11. 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. *Stat Med*. 2008;27:157-172; discussion 207-112
12. National Cholesterol Education Program Expert Panel on Detection E, Treatment of High Blood Cholesterol in A. Third report of the national cholesterol education program (ncep) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel iii) final report. *Circulation*. 2002;106:3143-3421
13. May S, Hosmer DW. A simplified method of calculating an overall goodness-of-fit test for the cox proportional hazards model. *Lifetime Data Anal*. 1998;4:109-120
14. Rapsomaniki E, White IR, Wood AM, Thompson SG, Emerging Risk Factors C. A framework for quantifying net benefits of alternative prognostic models. *Stat Med*. 2012;31:114-130
15. Cholesterol Treatment Trialists C, Mihaylova B, Emberson J, Blackwell L, Keech A, Simes J, Barnes EH, Voysey M, Gray A, Collins R, Baigent C. The

effects of lowering ldl cholesterol with statin therapy in people at low risk of vascular disease: Meta-analysis of individual data from 27 randomised trials. *Lancet*. 2012;380:581-590
16. White H. Maximum-likelihood estimation of mis-specified models. *Econometrica*. 1982;50:1-25
17. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. Plink: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559-575