

Metabolic Mediators of the Effects of Family History and Genetic Risk Score on Coronary Heart Disease—Findings From the Malmö Diet and Cancer Study

Josef Fritz, MSc; Dov Shiffman, PhD; Olle Melander, MD, PhD; Hayato Tada, MD, PhD; Hanno Ulmer, PhD

Background—Family history of coronary heart disease (CHD) as well as genetic predisposition to CHD assessed by a genetic risk score (GRS) are predictors of CHD risk. It is, however, uncertain to what extent these risk predictors are mediated by major metabolic pathways.

Methods and Results—Total effects of self-reported family history and a 50-variant GRS (GRS50), as well as effects mediated by apolipoprotein B and A-I (apoB, apoA-I), blood pressure, and diabetes mellitus, on incidence of CHD were estimated in 23 595 participants of the Malmö Diet and Cancer study (a prospective, population-based study). During a median follow-up of 14.4 years, 2213 participants experienced a first CHD event. Family history of CHD and GRS50 (highest versus other quintiles) were associated with incident CHD, with hazard ratios of 1.52 (95% CI: 1.39–1.65) and 1.53 (95% CI: 1.39–1.68), respectively, after adjusting for age, sex, and smoking status. Small proportions of the family history effect were mediated by metabolic risk factors: 8.3% (95% CI: 5.8–11.7%) by the apoB pathway, 1.7% (95% CI: 0.2–3.4%) by apoA-I, 8.5% (95% CI: 5.9–12.0%) by blood pressure, and 1.5% (95% CI: –0.8 to 3.8%) by diabetes mellitus. Similarly, small proportions of GRS50 were mediated: 8.1% (95% CI: 5.5–11.8%) by apoB, 1.2% (95% CI: 0.5–3.0%) by apoA-I, 4.2% (95% CI: 1.3–7.5%) by blood pressure, and –0.9% (95% CI: –3.7 to 1.6%) by diabetes mellitus.

Conclusions—A fraction of the CHD risk associated with family history or with GRS50 is mediated through elevated blood lipids and hypertension, but not through diabetes mellitus. However, a major part ($\geq 80\%$) of the genetic effect operates independently of established metabolic risk factor pathways. (*J Am Heart Assoc.* 2017;6:e005254 DOI: 10.1161/JAHA.116.005254).

Key Words: coronary heart disease • epidemiology • family history • genetic association • risk factor

Family history of coronary heart disease (CHD)^{1–3} as well as genetic predisposition to CHD directly assessed by genotyping and expressed as genetic risk score (GRS) for CHD^{4–7} are strong predictors of CHD risk. Among others,

Barrett-Connor and Khaw already showed that family history predicts cardiovascular risk independently of established metabolic risk factors.^{8–13} Similarly, it was later shown that GRSs are associated with CHD independently of established risk factors.^{5,14,15} However, a portion of the risk may well be explained or mediated through metabolic risk factor pathways. The fraction of risk associated with family history and GRS that is mediated through metabolic risk factor pathways has not been adequately reported in the medical literature.

Previous studies demonstrated that family history of CHD or GRS was significantly associated with CHD after adjusting for risk factors such as blood lipids, hypertension, and diabetes mellitus (DM).^{5,8–15} However, metabolic risk factors may not be confounders of genetic risk, but potential intermediate variables or mediators of the causal pathway between exposure (family history of CHD or GRS) and outcome (CHD incidence). A metabolically unhealthy person with a family history of CHD may suffer from hypertension or hypercholesterolemia; thus that person's predisposition to CHD is mediated through these risk factors. In contrast, a metabolically healthy person may have an increased genetic

From the Department of Medical Statistics, Informatics and Health Economics, Medical University of Innsbruck, Austria (J.F., H.U.); Quest Diagnostics, San Juan Capistrano, CA (D.S.); Department of Clinical Sciences, Lund University, Malmö, Sweden (O.M.); Department of Internal Medicine, Skåne University Hospital, Malmö, Sweden (O.M.); Department of Cardiovascular and Internal Medicine, Kanazawa University Graduate School of Medicine, Kanazawa, Japan (H.T.).

Accompanying Data S1 and Tables S1 through S6 are available at <http://jaha.ahajournals.org/content/6/3/e005254/DC1/embed/inline-supplementary-material-1.pdf>

Correspondence to: Hanno Ulmer, PhD, Department of Medical Statistics, Informatics and Health Economics, Medical University of Innsbruck, Schoepfstrasse 41, 6020 Innsbruck, Austria. E-mail: hanno.ulmer@i-med.ac.at

Received December 12, 2016; accepted February 9, 2017.

© 2017 The Authors. Published on behalf of the American Heart Association, Inc., by Wiley Blackwell. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

risk without elevated risk factor levels. The first case is an example of an indirect effect mediated by intermediate variables, while the second case is an example of a direct effect of the exposure on the outcome. The sum of the indirect and direct effects is the total effect. In contrast to adjustment for risk factors, statistical mediation analysis quantifies the magnitudes of total, direct, and indirect effects.

In the present study, we estimated the total effects of family history and of GRS, as well as the direct effect and indirect effects mediated by elevated blood lipids, hypertension, and DM on the incidence of CHD. Analyses were performed in the Malmö Diet and Cancer (MDC) study, a prospective, population-based study providing the most up-to-date information on genetic CHD risk for 23 595 men and women. For this purpose, we applied recently developed methods^{16,17} in the framework of causal inference.

Methods

Data Source and Study Population

The MDC study is a community-based, prospective observational study of 30 447 participants drawn from about 230 000 residents of Malmö, Sweden. Men and women aged 45 to 73 years were invited to participate and were enrolled between 1991 and 1996. Details of the MDC study design have been previously reported.^{18,19}

Participants completed a baseline examination that included a blood draw and a questionnaire with assessment of cardiovascular risk factors, including cigarette smoking, family history of myocardial infarction, lipid-lowering therapy, antihypertension therapy, and antidiabetic therapy.

Of note, apolipoprotein A-I (apoA-I) and apolipoprotein B (apoB) plasma levels were measured rather than high-density lipoprotein (HDL) and low-density lipoprotein (LDL) cholesterol levels. DM was self-reported based on a physician's diagnosis of DM or use of antidiabetic medication. Self-reported family history was based on the response of participants to a questionnaire about whether their mother, father, or sibling had a history of myocardial infarction.

Genotypes of the MDC study participants were determined using a multiplex method that combines polymerase chain reaction, allele-specific oligonucleotide ligation assays, and hybridization to oligonucleotides coupled to Luminex[®] 100TM xMAPTM microspheres (Luminex, Austin, TX).²⁰ Subsequently, a genetic risk score accounting for 50 CHD-related single nucleotide polymorphisms (SNPs) 50-variant genetic risk score (GRS50) was calculated, as explained in more detail in the work of Tada et al.²¹ In short, each study participant received a score equal to the sum of the number of risk alleles for each SNP weighted by the log of the odds ratio reported with the SNP in the original report.

The primary end point of the study was time to first occurrence of CHD. Incident CHD was defined as coronary revascularization, fatal or nonfatal myocardial infarction, or death attributable to ischemic heart disease. Events were identified by comparison with 3 registers—the Swedish Hospital Discharge Register, the Swedish Cause of Death Register, and the Swedish Coronary Angiography and Angioplasty Registry (SCAAR)—via ICD-9 codes 410, 412, and 414, ICD-10 codes I21, I22, I23, and I24, procedure codes 3065, 3066, 3068, 3080, 3092, 3105, 3127, and 3158 (the Op6 system), and FN (the KKÅ97 system).²²

After we excluded participants who were not genotyped, had prevalent CHD, or had missing information, 23 595 participants remained eligible for analyses in the current investigation. Details have been reported elsewhere.²¹

The MDC study was approved by the ethics committee at Lund University and performed in accordance with the 1964 Declaration of Helsinki and later amendments or comparable ethical standards. All participants provided written informed consent.

Statistical Analysis

Study participants' characteristics and risk factor measurements were analyzed descriptively (mean and SD for continuous variables, and counts and percentages for categorical data). Associations of metabolic risk factors with incidence of CHD were evaluated in a multivariable Cox proportional hazards models adjusting for participants' age, sex, smoking, and use of antihypertensives.

Mediation analysis was performed using the natural effect model proposed by Lange et al,^{16,17} based on the counterfactual framework. This approach offers a tool to decompose the total effect of a given exposure into a natural direct effect and a natural indirect effect through 1 or several mediators. The natural direct effect here is the effect one would expect if the sole difference between 2 individuals is the exposure (ie, family history of CHD or GRS50) and all risk factors are kept at the value they would naturally take. Detailed definitions of these effects have been described elsewhere.^{23,24} Lange et al's method is designed for categorical exposures; therefore, we classified GRS50 into a low-risk score category (quintile 1), an intermediate-risk score category (quintiles 2–4), and a high-risk score category (quintile 5). Mediation analysis was performed comparing the GRS50 risk categories high versus low and high versus intermediate and low combined. Figure shows the underlying model of our analyses depicted as a directed acyclic graph. We assessed the mediating role of apoA-I, apoB, systolic blood pressure and hypertension treatment, and prevalent DM on the association between family history or GRS50 and CHD incidence, in the presence of the mediator-outcome confounding variables age,

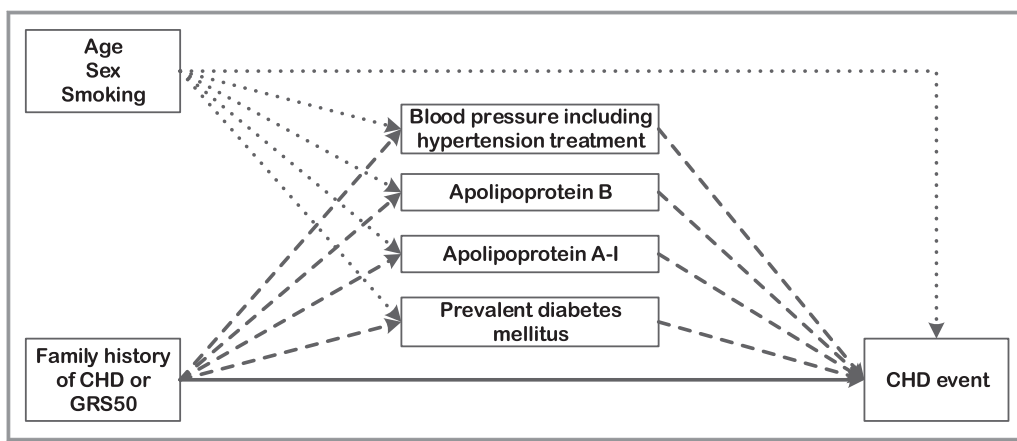


Figure. Causal diagram of the relations among genetic risk of CHD (exposure), metabolic risk factors of CHD (mediators), and CHD incidence (outcome) with measured confounders; metabolic risk factors are intermediate variables on the causal pathway between exposure (family history or GRS50) and outcome (CHD incidence). The direct effect is represented by the solid arrow and indirect effects are represented by dashed arrows; confounder pathways are depicted as dotted arrows. CHD indicates coronary heart disease; GRS50, 50-variant genetic risk score.

sex, and smoking. For this purpose, systolic blood pressure and the information on hypertension treatment were aggregated into 1 combined variable: In a Cox model corrected for all confounders, the effect of hypertension treatment was estimated to be equivalent to an increase of 27.5 mm Hg in systolic blood pressure; therefore, for all participants using hypertension treatment, the systolic blood pressure value was increased by an extra 27.5 mm Hg and this modified value was used for all further mediation analyses. Similar models were calculated regarding family history as exposure and GRS50 as mediator and vice versa.

Point estimates of the natural direct and natural indirect effects for the outcome CHD incidence were obtained by means of weighted regression models of the outcome on the exposure, the confounders, and additional counterfactual variables. Additional counterfactual variables taking the opposite (“counterfactual”) of the original exposure value were introduced in an extended set of the original data. Weights were derived from multivariable models regressing the mediators on the confounders and the exposure of interest (family history of CHD and GRS50, respectively). We used both Cox proportional hazards models quantifying relative risk in terms of hazard ratios (HRs), and additive hazards models with time-independent effects²⁵ quantifying absolute risk in terms of additional number of CHD events per 100 000 person-years. In Data S1 we give annotated R code of our implementation of the statistical modeling procedure. The assumption of proportional hazards for the Cox models was visually checked by inspecting log(-log) survival plots for all relevant variables. We also checked the assumption that each mediator is independent of the others, conditional on exposure and confounders in separate models.¹⁷

For HRs, the contribution of direct and indirect effects to the total effect of family history of CHD and GRS50 was calculated on the $\ln(\text{HR})$ scale, since HRs are additive on this scale. Ninety-five percent CIs were computed by bootstrapping using 2000 replications.

Possible effect modification by sex and age was examined by adding the respective multiplicative terms to the models. Because there was a significant interaction for indirect effects of apoA-I (sex, $P=0.025$) and apoB (sex, $P=0.026$, and age, $P<0.001$), we repeated the previously described analyses separately for men and women, as well as for individuals <50 years of age.

It is worth noting that our approach is not equivalent to Mendelian randomization.^{26,27} Neither self-reported family history nor GRS qualify as instrumental variables, because of the direct effects operating independently of metabolic risk factors.²⁸ A null or negative indirect effect in our setting does not mean that the variable concerned is not a risk factor for CHD.

Effects were considered statistically significant if null effects were not included in the 95% CI. All analyses were conducted in R, version 3.2.0.²⁹

Results

Study Population

Demographic and risk factor characteristics of the 23 595 MDC study participants are shown in Table 1 (stratified by self-reported family history and GRS50 categories). Mean age was 58.0 years, and 8973 (38.0%) of participants were male. By June 30, 2009, 2213 first CHD events were observed during a median follow-up of 14.4 years. CHD incidence was

Table 1. Baseline Characteristics According to Self-Reported Family History of CHD and GRS50 Categories

	Family History		GRS50 Score		
	No (N=14 807)	Yes (N=8788)	Low (N=4719)	Intermediate (N=14 157)	High (N=4719)
Age, y	57.8 (7.7)	58.4 (7.7)	58.2 (7.8)	58.0 (7.7)	57.9 (7.7)
Men, n (%)	5963 (40.3)	3010 (34.3)	1829 (38.8)	5342 (37.7)	1802 (38.2)
Body mass index, kg/m ²	25.7 (4.0)	25.8 (4.0)	25.7 (4.0)	25.7 (4.0)	25.8 (4.0)
Systolic blood pressure, mm Hg	140.5 (20.2)	142.1 (20.0)	140.8 (20.1)	141.0 (20.1)	141.8 (20.1)
Use of antihypertensives, n (%)	2307 (15.6)	1746 (19.9)	825 (17.5)	2417 (17.1)	811 (17.2)
Apolipoprotein A-I, g/L	1.569 (0.281)	1.574 (0.279)	1.573 (0.279)	1.571 (0.283)	1.565 (0.275)
Apolipoprotein B, g/L	1.059 (0.259)	1.086 (0.262)	1.051 (0.260)	1.068 (0.261)	1.089 (0.259)
Prevalent diabetes mellitus, n (%)	566 (3.8)	364 (4.1)	178 (3.8)	575 (4.1)	177 (3.8)
Current smoker, n (%)	4268 (28.8)	2354 (26.8)	1316 (27.9)	3999 (28.2)	1307 (27.7)
Self-reported family history of CHD, n (%)	0 (0.0)	8788 (100.0)	1552 (32.9)	5314 (37.5)	1922 (40.7)
GRS50	−0.045 (0.99)	0.075 (1.00)	−1.37 (0.42)	−0.02 (0.46)	1.43 (0.50)
Incident CHD event, n (%)	1215 (8.2)	998 (11.4)	318 (6.7)	1303 (9.2)	592 (12.5)
Median (mean) Follow-up	13.87 (14.45)	13.65 (14.37)	14.47 (13.97)	14.41 (13.80)	14.39 (13.56)
Incidence rate from CHD per 100 000 person-years	591.8	832.1	482.3	667.1	925.2

Data are presented as mean (SD) unless indicated. CHD indicates coronary heart disease; GRS50, 50-variant genetic risk score.

markedly higher in participants with family history and those in the high GRS50 category. Participants with low GRS50 or without family history of CHD had lower levels of apoB and systolic blood pressure than those with high GRS50 or with family history of CHD. Those with family history of CHD were more likely to have prevalent DM than those without family history, although the differences between the groups were small.

All metabolic risk factors were associated with risk of incident CHD: the HRs per 1 SD increment were 1.28 (95% CI, 1.23–1.34) for systolic blood pressure, 1.32 (95% CI, 1.26–1.37) for ApoB, and 0.78 (95% CI, 0.74–0.82) for ApoA-I; the HR for prevalent DM was 2.28 (95% CI, 1.99–2.61).

Mediation Analysis: Total, Direct, and Indirect Effects of Family History of CHD

Family history of CHD was associated with incident CHD with an HR of 1.52 (95% CI, 1.39–1.65), after adjusting for age, sex, and smoking status. A fraction of this risk (20.0%, 95% CI, 14.8–26.4%) could be attributed to indirect pathways mediated by established metabolic risk factors. Specifically, 8.3% (95% CI, 5.8–11.7%) of the total effect was mediated through the apoB pathway, 1.7% (95% CI, 0.2–3.4%) through apoA-I, and 8.5% (95% CI, 5.9–12.0%) through systolic blood pressure. The indirect effect through DM (1.5% [95% CI, −0.8% to 3.8%]) did not reach statistical significance. In absolute terms, family history of CHD was associated with 220 additional events of CHD per 100 000 person-years at risk, of which 52 could be

attributed to metabolic risk factor pathways; the most relevant risk factors apo-B and systolic blood pressure added 20 and 24 additional events, respectively (Table 2).

When comparing men and women in stratified analyses, 20.8% (95% CI, 14.1–31.0%) of family history was mediated through metabolic risk factors in women versus 16.3% (95% CI, 8.3–27.6%) in men (Table S1). The effects of family history (total HR: 1.81 [95% CI, 1.31–2.53]; proportion mediated: 23.5% [95% CI, 10.5–51.6%]) were generally higher in individuals below 50 years of age as compared to the overall population (Table S2).

Mediation Analysis: Total, Direct, and Indirect Effects of GRS50

GRS50 (highest versus other quintiles) was associated with incident CHD with an HR of 1.53 (95% CI, 1.39–1.68), after adjusting for age, sex, and smoking status. Only 12.6% (95% CI, 7.3–19.1%) of this risk could be attributed to indirect pathways mediated by established metabolic risk factors. Most of this attributable risk was mediated through the apoB pathway (8.1%, 95% CI, 5.5–11.8%), 1.2% (95% CI, −0.5% to 3.0%) through apoA-I, and 4.2% (95% CI, 1.3–7.5%) through systolic blood pressure. No statistically significant mediation was observed for DM (−0.9% [95% CI, −3.7% to 1.6%]). In absolute terms, GRS50 was associated with 281 additional events of CHD per 100 000 person-years at risk, of which 33 could be attributed to metabolic risk factor pathways; the most relevant risk factors apo-B and systolic blood pressure added 20 and 12 additional events, respectively (Table 3).

Table 2. Total, Direct, and Indirect Effects of Family History on Incident CHD With Metabolic Mediators, Adjusted for Age at Baseline, Sex, and Smoking Status

Effects	Family History (Yes vs No)		
	Hazard Ratio (95% CI)	Proportion Explained (%) (95% CI)*	Additional Incident CHD Cases Per 100 000 Person-Years at Risk†
Total effect	1.52 (1.39–1.65)	100.0%	269.7
Direct effect‡	1.40 (1.28–1.52)	80.0% (73.6–85.2%)	220.2
Indirect effect, combined	1.09 (1.07–1.11)	20.0% (14.8–26.4%)	52.5
Indirect effect, through systolic blood pressure and hypertension treatment	1.04 (1.03–1.05)	8.5% (5.9–12.0%)	23.5
Indirect effect, through apoA-I	1.01 (1.00–1.01)	1.7% (0.2–3.4%)	5.1
Indirect effect, through apoB	1.04 (1.03–1.05)	8.3% (5.8–11.7%)	19.8
Indirect effect, through diabetes mellitus	1.01 (1.00–1.02)	1.5% (–0.8% to 3.8%)	4.1

apoA-I indicates apolipoprotein A-I; apoB, apolipoprotein B; CHD, coronary heart disease.

*On ln(HR) scale.

†Estimates from additive hazards models with time-independent effects.

‡Effect of family history not mediated by the 4 analyzed risk factors.

When comparing only the highest versus the lowest quintile, the total effect was more pronounced, with an HR of 2.01 (95% CI, 1.76–2.30). However, proportions mediated were similar (Table 3).

In sex-specific analyses, 17.7% (95% CI, 9.0–30.9%) of GRS50 (highest versus other quintiles) was mediated through metabolic risk factors in women versus 7.4% (95% CI, –0.1% to 15.5%) in men (Table S3). The effects of GRS50 (total HR: 2.01 [95% CI, 1.40–2.78]) were similar regarding proportion

mediated (10.1% [95% CI, –1.3% to 28.8%]) in individuals below 50 years of age as compared to the overall population (Table S4).

Relationship of Family History of CHD With GRS50

An analysis of the relationship between family history of CHD and GRS50 showed that of the total HR associated with family

Table 3. Total, Direct, and Indirect Effects of GRS50 on Incident CHD With Metabolic Mediators, Adjusted for Age at Baseline, Sex, and Smoking Status

Effects	GRS50 (High vs Low)			GRS50 (High vs Low/Intermediate)		
	Hazard Ratio (95% CI)	Proportion Explained (%) (95% CI)*	Additional Incident CHD Cases Per 100 000 Person-Years at Risk†	Hazard Ratio (95% CI)	Proportion Explained (%) (95% CI)*	Additional Incident CHD Cases Per 100 000 Person-Years at Risk†
Total effect	2.01 (1.76–2.30)	100.0%	469.3	1.53 (1.39–1.68)	100.0%	314.2
Direct effect‡	1.87 (1.64–2.14)	89.3% (84.0–94.2%)	424.7	1.45 (1.32–1.59)	87.4% (80.9–92.7%)	281.2
Indirect effect, combined	1.08 (1.04–1.11)	10.7% (5.8–16.0%)	44.6	1.06 (1.03–1.08)	12.6% (7.3–19.1%)	33.0
Indirect effect, through systolic blood pressure and hypertension treatment	1.02 (1.01–1.04)	3.5% (1.0–5.9%)	16.2	1.02 (1.01–1.03)	4.2% (1.3–7.5%)	11.9
Indirect effect, through apoA-I	1.01 (1.00–1.02)	1.1% (–0.2% to 2.6%)	5.3	1.01 (1.00–1.01)	1.2% (–0.5% to 3.0%)	3.6
Indirect effect, through apoB	1.04 (1.03–1.06)	6.0% (3.7–8.6%)	22.3	1.04 (1.02–1.05)	8.1% (5.5–11.8%)	19.9
Indirect effect, through diabetes mellitus	1.00 (0.99–1.02)	0.2% (–1.6% to 2.7%)	0.7	1.00 (0.98–1.01)	–0.9% (–3.7% to 1.6%)	–2.4

apoA-I indicates apolipoprotein A-I; apoB, apolipoprotein B; CHD, coronary heart disease; GRS50, 50-variant genetic risk score.

*On ln(HR) scale.

†Estimates from additive hazards models with time-independent effects.

‡Effect of GRS50 not mediated by the 4 analyzed risk factors.

history (1.52; 95% CI, 1.39–1.65) only 6.2% (95% CI, 4.3–8.7%) was mediated through GRS50 (Table S5), and of the total HR associated with GRS50 (1.53, 95% CI, 1.39–1.68; highest versus other quintiles), 4.2% (95% CI, 2.5–6.6%) was mediated by family history (Table S6).

Discussion

The main purpose of the present work was to explore whether the previously identified associations between family history or GRS50 and CHD incidence are mediated through major metabolic risk factors. Specifically, we sought to decompose the total effect of family history and GRS50, respectively, on CHD risk into a direct part and distinct indirect parts via different risk factor pathways. Our findings demonstrate that some of the risk associated with family history as well as some of the risk conferred by GRS50 is mediated through known metabolic risk factors, specifically apolipoproteins and blood pressure, but we did not find evidence that DM mediates these effects. However, a major part ($\geq 80\%$) of the family history effect and of the GRS50 effect operates independently of established metabolic risk factor pathways.

Recently, genetic risk scores were added to cardiovascular risk prediction, reflecting a partially different aspect of genetic information than reported family history.^{6,21,30} Through genome-wide association studies, identification of CHD-related SNPs has progressed rapidly. Consequently, CHD-related SNPs have been aggregated in genetic risk scores, providing simple predictive measures of the risk of developing CHD.^{4–7} The number of known CHD-related SNPs used in GRSs has expanded from just 13 in 2010⁴ to 25 in 2011,³¹ and 46 in 2013.⁵ The 50-SNP GRS used for this study was introduced in 2016 by Tada et al²¹ and represents the most comprehensive GRS at time of writing.

Attempts to relate CHD-predicting SNPs to metabolic risk factors have been made in several studies.^{5,14,15,31} The CARDIoGRAMplusC4D Consortium found in their study of 46 CHD-predicting SNPs, a subset of the GRS50 used here, that 12 SNPs were associated with lipid traits, 5 with blood pressure, and none with DM, leaving the major part of SNPs operating independently of these 3 risk factors.¹⁵ Ganna et al came to similar conclusions that “most of the CHD loci are not involved in pathways perturbing currently known risk factors.”⁵ However, their methods did not allow quantification of the proportions mediated. Our findings, that about 9% of the total effect of the GRS50 was mediated by apolipoproteins, 4% by blood pressure, and none by DM, confirm their heuristic findings.

Other studies investigated the role of family history, either self-reported or with validated parental event records, as a predictor for CHD.^{1–3,8–13} Sesso et al¹ examined the association of self-reported family history of myocardial infarction

with risk of offspring cardiovascular disease using data from the Physicians' Health Study and the Women's Health Study. They adjusted for cardiovascular risk factors, but did not perform a mediation analysis. However, proportions mediated through cardiovascular risk factors can be estimated from their published results as a rough approximate measure. The results of Sesso et al¹ were quantitatively similar to ours, specifically, the proportions mediated as converted were in a comparable range. Lloyd-Jones et al¹³ examined the association between parental cardiovascular disease and CHD risk using data from the Framingham Heart Study where parental events were validated. In contrast to Sesso et al¹ and our results, Lloyd-Jones et al¹³ gave higher estimates of total effects, and the proportions mediated as estimated from these published results were higher as well. It is possible that the effects of family history and mediated proportions may be diluted because of the self-reporting of family history. In line with our subgroup analysis of MDC study participants aged below 50 years (Table S2), participants' age may constitute another reason for differences in effect estimates, because family history more strongly predicts onset of CHD at a younger age.^{21,32} Thus, the higher proportion mediated through risk factors in the Framingham Heart Study, which investigated a younger study population, seems plausible. Furthermore, the knowledge of an existing genetic risk may alter a person's health behavior and may thus lead to reduced risk factor levels³³ and consequently to smaller proportions mediated. Finally, errors in the measurement of the mediators may also lead to an underestimation of the indirect effects.³⁴

Mediation analysis applied on the relationship between family history and GRS50 suggests that these 2 measures of genetic risk are partly independent of each other. The proportions mediated by metabolic risk factors were stronger for family history than for GRS50 (20.0% versus 12.6%), mainly because blood pressure mediation was larger for family history than for GRS (8.5% versus 4.2%). The search for candidate genes for CHD is still ongoing;³⁰ thus, the addition of new genetic variants to GRSs can alter the difference between the mediated fractions of family history and GRS50.

Our study has several strengths and some potential limitations. Major strengths are the prospective study design, the large sample size, and the length of follow-up in the MDC study. In addition, we were able to use the newest available genetic data in CHD so far. The recently developed mediation analysis technique^{16,17} allowed for the first time quantification of the mediated effects separately for each single metabolic factor. Limitations are that the study was conducted in Swedish middle-aged individuals; hence, the generalizability to other ethnicities or other age groups is uncertain. Although genetic risk assessment could be useful in the young,²¹ this study population did not include individuals younger than 45 years and thus no risk estimates for this age group are

possible. Overrepresentation of women in the MDC study cohort may influence the somewhat different effect estimates between men and women. A recently published study showed age- and sex-related differences regarding metabolic mediation.³⁵ LDL cholesterol and HDL cholesterol levels were not available for our study population; therefore, we used the available apoA-I and apoB plasma levels as covariates in our established risk factors model. Although apoB is incorporated in a few lipoproteins in addition to LDL and apoA-I does not completely represent HDL, it has been shown that apolipoproteins have equally strong predictive abilities for future CHD events as LDL cholesterol does.^{36–38}

In conclusion, a fraction of the CHD risk associated with family history or with GRS50 is mediated through elevated blood lipids and hypertension, but not through DM. However, the major part of the genetic effect operates independently from the established metabolic risk factors, confirming the importance of the assessment of family history and genetic predisposition to CHD. Metabolically healthy individuals with genetic predisposition form an important group of individuals at risk for CHD, providing a major challenge for primary prevention. Therefore, intensified preventive measures in addition to risk factor surveillance and treatment may be of benefit in genetically predisposed individuals.

Sources of Funding

The Malmö Diet and Cancer study was made possible by grants from the Swedish Cancer Society, the Swedish Medical Research Council, the Swedish Dairy Association, the Albert Pålsson and Gunnar Nilsson Foundations, and the Malmö City Council. Melander is supported by the European Research Council (StG-282255); the Swedish Heart and Lung Foundation, Swedish Research Council; the Novo Nordisk Foundation; the Skåne University Hospital donation funds; the Medical Faculty, Lund University; the Governmental funding of clinical research within the National Health Services; the Albert Pålsson Research Foundation, Region Skåne; the King Gustav V and Queen Victoria Foundation; and the Marianne and Marcus Wallenberg Foundation. Melander is the recipient of an investigator-initiated grant from Quest Diagnostics.

Disclosures

Shiffman is an employee of Quest Diagnostics. The remaining authors have no disclosures to report.

References

- Sesso HD, Lee IM, Gaziano JM, Rexrode KM, Glynn RJ, Buring JE. Maternal and paternal history of myocardial infarction and risk of cardiovascular disease in men and women. *Circulation*. 2001;104:393–398.
- Nielsen M, Andersson C, Gerds TA, Andersen PK, Jensen TB, Køber L, Gislason G, Torp-Pedersen C. Familial clustering of myocardial infarction in first-degree relatives: a nationwide study. *Eur Heart J*. 2013;34:1198–1203.
- Weijmans M, van der Graaf Y, Reitsma JB, Visseren FLJ. Paternal or maternal history of cardiovascular disease and the risk of cardiovascular disease in offspring. A systematic review and meta-analysis. *Int J Cardiol*. 2015;179:409–416.
- Ripatti S, Tikkanen E, Orho-Melander M, Havulinna AS, Silander K, Sharma A, Guiducci C, Perola M, Jula A, Sinisalo J, Lokki M-L, Nieminen MS, Melander O, Salomaa V, Peltonen L, Kathiresan S. A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. *Lancet*. 2010;376:1393–1400.
- Ganna A, Magnusson PKE, Pedersen NL, de Faire U, Reilly M, Arnlöv J, Sundström J, Hamsten A, Ingelsson E. Multilocus genetic risk scores for coronary heart disease prediction. *Arterioscler Thromb Vasc Biol*. 2013;33:2267–2272.
- Mega JL, Stitzel NO, Smith JG, Chasman DI, Caulfield MJ, Devlin JJ, Nordio F, Hyde CL, Cannon CP, Sacks FM, Poulter NR, Sever PS, Ridker PM, Braunwald E, Melander O, Kathiresan S, Sabatine MS. Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. *Lancet*. 2015;385:2264–2271.
- Kathiresan S, Srivastava D. Genetics of human cardiovascular disease. *Cell*. 2012;148:1242–1257.
- Barrett-Connor E, Khaw K. Family history of heart attack as an independent predictor of death due to cardiovascular disease. *Circulation*. 1984;69:1065–1069.
- Shea S, Ottman R, Gabrieli C, Stein Z, Nichols A. Family history as an independent risk factor for coronary artery disease. *J Am Coll Cardiol*. 1984;4:793–801.
- Colditz GA, Stampfer MJ, Willett WC, Rosner B, Speizer FE, Hennekens CH. A prospective study of parental history of myocardial infarction and coronary heart disease in women. *Am J Epidemiol*. 1986;123:48–58.
- Colditz GA, Rimm EB, Giovannucci E, Stampfer MJ, Rosner B, Willett WC. A prospective study of parental history of myocardial infarction and coronary artery disease in men. *Am J Cardiol*. 1991;67:933–938.
- Jousilahti P, Puska P, Vartiainen E, Pekkanen J, Tuomilehto J. Parental history of premature coronary heart disease: an independent risk factor of myocardial infarction. *J Clin Epidemiol*. 1996;49:497–503.
- Lloyd-Jones DM, Nam B-H, D'Agostino RB, Levy D, Murabito JM, Wang TJ, Wilson PWF, O'Donnell CJ. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. *JAMA*. 2004;291:2204–2211.
- Angelakopoulou A, Shah T, Sofat R, Shah S, Berry DJ, Cooper J, Palmen J, Tzoulaki I, Wong A, Jefferis BJ, Maniatis N, Drenos F, Gigante B, Hardy R, Laxton RC, Leander K, Motterle A, Simpson IA, Smeeth L, Thomson A, Verzilli C, Kuh D, Ireland H, Deanfield J, Caulfield M, Wallace C, Samani N, Munroe PB, Lathrop M, Fowkes FGR, Marmot M, Whincup PH, Whittaker JC, de Faire U, Kivimäki M, Kumari M, Hyppönen E, Power C, Humphries SE, Talmud PJ, Price J, Morris RW, Ye S, Casas JP, Hingorani AD. 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.
- CARDIOGRAMplusC4D Consortium, Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, Ingelsson E, Saleheen D, Erdmann J, Goldstein BA, Stirrups K, König IR, Cazier JB, Johansson A, Hall AS, Lee JY, Willer CJ, Chambers JC, Esko T, Folkersen L, Goel A, Grundberg E, Havulinna AS, Ho WK, Hopewell JC, Eriksson N, Kleber ME, Kristiansson K, Lundmark P, Lyytikäinen LP, Rafelt S, Shungin D, Strawbridge RJ, Thorleifsson G, Tikkanen E, Van Zuydam N, Voight BF, Waite LL, Zhang W, Ziegler A, Absher D, Altschuler D, Balmforth AJ, Barroso I, Braund PS, Burgdorf C, Claudi-Boehm S, Cox D, Dimitriou M, Do R, DIAGRAM Consortium; CARDIOGENICS Consortium, Doney AS, El Mokhtari N, Eriksson P, Fischer K, Fontanillas P, Franco-Cereceda A, Gigante B, Groop L, Gustafsson S, Hager J, Hallmans G, Han BG, Hunt SE, Kang HM, Illig T, Kessler T, Knowles JW, Kolovou G, Kuusisto J, Langenberg C, Langford C, Leander K, Lokki ML, Lundmark A, McCarthy MI, Meisinger C, Melander O, Mihailov E, Maouche S, Morris AD, Müller-Nurasyid M; MuTHER Consortium, Nikus K, Peden JF, Rayner NW, Rasheed A, Rosinger S, Rubin D, Rumpf MP, Schäfer A, Sivananthan M, Song C, Stewart AF, Tan ST, Thorgeirsson G, van der Schoot CE, Wagner PJ; Wellcome Trust Case Control Consortium, Wells GA, Wild PS, Yang TP, Amouyel P, Arveiler D, Basart H, Boehnke M, Boerwinkle E, Brambilla P, Cambien F, Cupples AL, de Faire U, Dehghan A, Diemert P, Epstein SE, Evans A, Ferrario MM, Ferrières J, Gauguier D, Go AS, Goodall AH, Gudnason V, Hazen SL, Holm H, Iribarren C, Jiang Y, Kähönen M, Kee F, Kim HS, Klopp N, Koenig W, Kratzer W, Kuulasmaa K, Laakso M, Laaksonen R, Lee JY, Lind L, Ouwehand WH, Parish S, Park JE, Pedersen NL, Peters A, Quertermous T, Rader DJ, Salomaa V, Schadt E, Shah SH, Sinisalo J, Stark K, Stefansson K, Tréguët DA, Virtamo J, Wallentin L, Wareham N, Zimmermann ME, Nieminen MS, Hengstenberg C, Sandhu MS,

- Pastinen T, Syvänen AC, Hovingh GK, Dedoussis G, Franks PW, Lehtimäki T, Metspalu A, Zalloua PA, Siegbahn A, Schreiber S, Ripatti S, Blankenberg SS, Perola M, Clarke R, Boehm BO, O'Donnell C, Reilly MP, März W, Collins R, Kathiresan S, Hamsten A, Kooner JS, Thorsteinsdóttir U, Danesh J, Palmer CN, Roberts R, Watkins H, Schunkert H, Samani NJ. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet.* 2013;45:25–33.
16. Lange T, Vansteelandt S, Bekaert M. A simple unified approach for estimating natural direct and indirect effects. *Am J Epidemiol.* 2012;176:190–195.
 17. Lange T, Rasmussen M, Thygesen LC. Assessing natural direct and indirect effects through multiple pathways. *Am J Epidemiol.* 2014;179:513–518.
 18. Berglund G, Elmståhl S, Janzon L, Larsson SA. The Malmö Diet and Cancer Study. Design and feasibility. *J Intern Med.* 1993;233:45–51.
 19. Smith J, Platonov P, Hedblad B, Engström G, Melander O. Atrial fibrillation in the Malmö Diet and Cancer study: a study of occurrence, risk factors and diagnostic validity. *Eur J Epidemiol.* 2010;25:95–102.
 20. Shiffman D, O'Meara ES, Bare LA, Rowland CM, Louie JZ, Arellano AR, Lumley T, Rice K, Jakubova O, Luke MM, Young BA, Malloy MJ, Kane JP, Ellis SG, Tracy RP, Devlin JJ, Psaty BM. Association of gene variants with incident myocardial infarction in the Cardiovascular Health Study. *Arterioscler Thromb Vasc Biol.* 2008;28:173–179.
 21. Tada H, Melander O, Louie JZ, Catanese JJ, Rowland CM, Devlin JJ, Kathiresan S, Shiffman D. Risk prediction by genetic risk scores for coronary heart disease is independent of self-reported family history. *Eur Heart J.* 2016;37:561–567.
 22. Ludvigsson JF, Andersson E, Ekblom A, Feychting M, Kim J-L, Reuterwall C, Heurgren M, Olausson PO. External review and validation of the Swedish national inpatient register. *BMC Public Health.* 2011;11:450.
 23. Lange T, Hansen JV. Direct and indirect effects in a survival context. *Epidemiology.* 2011;22:575–581.
 24. Pearl J. Direct and indirect effects. In: Proceedings of the Seventeenth Conference on Uncertainty and Artificial Intelligence. San Francisco: Morgan Kaufmann; 2001:411–420.
 25. Aalen O. A model for non-parametric regression analysis of counting processes. In: Klonecki W, Kozek A, Rosinski J, eds. *Lecture Notes in Statistics, Mathematical Statistics and Probability Theory.* New York: Springer; 1980:1–25.
 26. Didelez V, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. *Stat Methods Med Res.* 2007;16:309–330.
 27. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res.* 2015;1–26.
 28. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol.* 2000;29:722–729.
 29. R Core Team. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing; 2013.
 30. Schunkert H. Family or SNPs: what counts for hereditary risk of coronary artery disease? *Eur Heart J.* 2016;37:568–571.
 31. Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AF, Barbalic M, Gieger C, Absher D, Aherrahrou Z, Allayee H, Altshuler D, Anand SS, Andersen K, Anderson JL, Ardissino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, Buyschaert I, Cardiogenics, Carlquist JF, Chen L, Cichon S, Codd V, Davies RW, Dedoussis G, Dehghan A, Demissie S, Devaney JM, Diemert P, Do R, Doering A, Eifert S, Mokhtari NE, Ellis SG, Elosua R, Engert JC, Epstein SE, de Faire U, Fischer M, Folsom AR, Freyer J, Gigante B, Girelli D, Gretarsdóttir S, Gudnason V, Gulcher JR, Halperin E, Hammond N, Hazen SL, Hofman A, Horne BD, Illig T, Iribarren C, Jones GT, Jukema JW, Kaiser MA, Kaplan LM, Kastelein JJ, Khaw KT, Knowles JW, Kolovou G, Kong A, Laaksonen R, Lambrechts D, Leander K, Lettre G, Li M, Lieb W, Loley C, Lotery AJ, Mannucci PM, Maouche S, Martinelli N, McKeown PP, Meisinger C, Meitinger T, Melander O, Merlini PA, Mooser V, Morgan T, Mühleisen TW, Muhlestein JB, Münzel T, Musunuru K, Nahrstaedt J, Nelson CP, Nöthen MM, Olivieri O, Patel RS, Patterson CC, Peters A, Peyvandi F, Qu L, Quyyumi AA, Rader DJ, Rallidis LS, Rice C, Rosendaal FR, Rubin D, Salomaa V, Sampietro ML, Sandhu MS, Schadt E, Schäfer A, Schillert A, Schreiber S, Schrezenmeier J, Schwartz SM, Siscovick DS, Sivananthan M, Sivapalaratnam S, Smith A, Smith TB, Snoop JD, Soranzo N, Spertus JA, Stark K, Stirrups K, Stoll M, Tang WH, Tennstedt S, Thorgeirsson G, Thorleifsson G, Tomaszewski M, Uitterlinden AG, van Rij AM, Voight BF, Wareham NJ, Wells GA, Wichmann HE, Wild PS, Willenborg C, Wittman JC, Wright BJ, Ye S, Zeller T, Ziegler A, Cambien F, Goodall AH, Cupples LA, Quertermous T, März W, Hengstenberg C, Blankenberg S, Ouwehand WH, Hall AS, Deloukas P, Thompson JR, Stefansson K, Roberts R, Thorsteinsdóttir U, O'Donnell CJ, McPherson R, Erdmann J. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet.* 2011;43:333–338.
 32. Gränsbo K, Almgren P, Nilsson PM, Hedblad B, Engström G, Melander O. Risk factor exposure in individuals free from cardiovascular disease differs according to age at first myocardial infarction. *Eur Heart J.* 2016;37:1977–1981.
 33. Kullo IJ, Jouni H, Austin EE, Brown S-A, Kruisselbrink TM, Isseh IN, Haddad RA, Marroush TS, Shameer K, Olson JE, Broeckel U, Green RC, Schaid DJ, Montori VM, Bailey KR. Incorporating a genetic risk score into coronary heart disease risk estimates: effect on LDL cholesterol levels (the MIGENES clinical trial). *Circulation.* 2016;133:1181–1188.
 34. VanderWeele TJ. *Explanation in Causal Inference: Methods for Mediation and Interaction.* 1st ed. New York: Oxford University Press; 2015.
 35. Fritz J, Edlinger M, Kelleher C, Strohmaier S, Nagel G, Concin H, Ruttman E, Hochleitner M, Ulmer H. Mediation analysis of the relationship between sex, cardiovascular risk factors and mortality from coronary heart disease: findings from the population-based VHM&PP cohort. *Atherosclerosis.* 2015;243:86–92.
 36. Pischon T, Girman CJ, Sacks FM, Rifai N, Stampfer MJ, Rimm EB. Non-high-density lipoprotein cholesterol and apolipoprotein B in the prediction of coronary heart disease in men. *Circulation.* 2005;112:3375–3383.
 37. Ingelsson E, Schaefer EJ, Contois JH, McNamara JR, Sullivan L, Keyes MJ, Pencina MJ, Schoonmaker C, Wilson PWF, D'Agostino RB, Vasan RS. Clinical utility of different lipid measures for prediction of coronary heart disease in men and women. *JAMA.* 2007;298:776–785.
 38. Sondermeijer BM, Rana JS, Arsenault BJ, Shah PK, Kastelein JJP, Wareham NJ, Boekholdt SM, Khaw K-T. Non-HDL cholesterol vs. apo B for risk of coronary heart disease in healthy individuals: the EPIC-Norfolk prospective population study. *Eur J Clin Invest.* 2013;43:1009–1015.

Supplemental Material

Data S1.

In this appendix, we provide detailed annotated R code of how we implemented the natural effects model, the statistical procedure we used for our analysis, including the calculation of bootstrapping confidence intervals, exemplarily for the variable family history. This method was introduced by Lange et al. in 2014 (Lange T, Rasmussen M, Thygesen LC. Assessing natural direct and indirect effects through multiple pathways. *Am J Epidemiol.* 2014;179:513–8).

The data frame `myData` contains the following variables. Data originates from the MDCS cohort.

<code>ID</code>	Subject identifier
<code>age</code>	Age at baseline examination
<code>sex</code>	Sex (1=male, 2=female)
<code>smoking</code>	Indicator variable (0/1) of smoking status
<code>FH</code>	Reported family history of CHD (0=no, 1=yes)
<code>ApoA</code>	ApoA-I level at baseline examination
<code>ApoB</code>	ApoB level at baseline examination
<code>DM</code>	Prevalent diabetes mellitus at baseline examination (0=no, 1=yes)
<code>BP</code>	Systolic blood pressure at baseline examination
<code>time</code>	CHD event time or censor time (years after baseline examination)
<code>event</code>	event indicator (0= No CHD event, 1=CHD event)

Our aim is to quantify how much of the effect of family history of CHD is mediated through the apoA-I pathway, the apoB pathway, the DM pathway, and the BP pathway, in the presence of the mediator-outcome confounders age, sex, and smoking. More specifically, we want to obtain estimates of natural direct, natural indirect, and total effects conditional on the confounder variables. We refer to the work of Pearl for definitions of these effects (Pearl J. *Causality: Models, reasoning and inference.* Cambridge University Press. 2009. 2nd edn.).

In a first step, we regress each mediator on the exposure variable `FH`, while conditioning on the covariates `age`, `sex`, and `smoking`. For the continuous mediators, we choose linear regressions, for the dichotomous mediator `DM` a logistic regression. We also save the residual variances of the linear models since they will be needed for the computation of weights later on. Since being convenient for the weight calculation, also a new variable `FHtemp` is created.

```
myData$FHtemp <- myData$FH
fitApoA <- lm(ApoA ~ FHtemp+age+sex+smoking, data=myData)
fitApoAvariance <- summary(fitApoA)$sigma^2
fitApoB <- lm(ApoB ~ FHtemp+age+sex+smoking, data=myData)
fitApoBvariance <- summary(fitApoB)$sigma^2
fitBP <- lm(BP ~ FHtemp+age+sex+smoking, data=myData)
fitBPvariance <- summary(fitBP)$sigma^2
fitDM <- glm(DM ~ FHtemp+age+sex+smoking, data=myData,
family="binomial")
```

To assess if the mediators do indeed correspond to distinct non-intertwined causal pathways, these models can be extended to include also the respective other mediator variables. For the variable `ApoA`, for example, the respective model is

```
fitApoA_check <- lm(ApoA ~ FHtemp+age+sex+smoking+ApoB+DM+BP,
data=myData)
```

Small parameter estimates for the additional variables `ApoB`, `DM`, and `BP` as well as non-significant p-values, obtained for example by

```
anova(fitApoA_check)
```

are an indication of non-intertwined causal pathways. Analogous analyses were done for the remaining three mediators.

Next, we construct a new, expanded dataset, denoted `myDataExpanded`, with new auxiliary exposure variables `FHInd1`, `FHInd2`, `FHInd3`, and `FHInd4`, each corresponding to the value of the exposure relative to the respective indirect path. We replicate each observation

16 times and assign FHInd1, FHInd2, FHInd3, and FHInd4 the values 0 or 1 in all possible combinations of these ($2^4=16$ combinations). This can be achieved by the following code:

```
levelsOfFH <- unique(myData$FH)
ds1 <- myData
ds2 <- myData
ds1$FHInd1 <- levelsOfFH[1]
ds2$FHInd1 <- levelsOfFH[2]
tempds <- rbind(ds1,ds2)
ds1 <- tempds
ds2 <- tempds
ds1$FHInd2 <- levelsOfFH[1]
ds2$FHInd2 <- levelsOfFH[2]
tempds <- rbind(ds1,ds2)
ds1 <- tempds
ds2 <- tempds
ds1$FHInd3 <- levelsOfFH[1]
ds2$FHInd3 <- levelsOfFH[2]
tempds <- rbind(ds1,ds2)
ds1 <- tempds
ds2 <- tempds
ds1$FHInd4 <- levelsOfFH[1]
ds2$FHInd4 <- levelsOfFH[2]
myDataExpanded <- rbind(ds1,ds2)
```

Now, it comes to the computation of the weights. The weights consist of a part corresponding to the ApoA-I level, a part corresponding to the ApoB level, a part corresponding to the DM status, and a part corresponding to systolic blood pressure. These parts are calculated consecutively via the following code (using the `predict` function) and are then combined.

The weight part corresponding to the apoA-I level:


```

myDataExpanded$FHtemp <- myDataExpanded$FH
tempDir1 <-
dnorm(myDataExpanded$ApoA, mean=predict (fitApoA, type="response", newda
ta=myDataExpanded), sd=sqrt (fitApoAvariance))
myDataExpanded$FHtemp <- myDataExpanded$FHInd1
tempIndir1 <-
dnorm(myDataExpanded$ApoA, mean=predict (fitApoA, type="response", newda
ta=myDataExpanded), sd=sqrt (fitApoAvariance))
myDataExpanded$weight1 <- tempIndir1/tempDir1

```

The weight part corresponding to the apoB level:

```

myDataExpanded$FHtemp <- myDataExpanded$FH
tempDir2 <-
dnorm(myDataExpanded$ApoB, mean=predict (fitApoB, type="response", newda
ta=myDataExpanded), sd=sqrt (fitApoBvariance))
myDataExpanded$FHtemp <- myDataExpanded$FHInd2
tempIndir2 <-
dnorm(myDataExpanded$ApoB, mean=predict (fitApoB, type="response", newda
ta=myDataExpanded), sd=sqrt (fitApoBvariance))
myDataExpanded$weight2 <- tempIndir2/tempDir2

```

The weight part corresponding to the DM status:

```

myDataExpanded$FHtemp <- myDataExpanded$FH
temp <- predict (fitDM, type="response", newdata=myDataExpanded)
tempDir3 <- ifelse (myDataExpanded$DM==1, temp, 1-temp)
myDataExpanded$FHtemp <- myDataExpanded$FHInd3
temp <- predict (fitDM, type="response", newdata= myDataExpanded)
tempIndir3 <- ifelse (myDataExpanded$DM==1, temp, 1-temp)
myDataExpanded$weight3 <- tempIndir3/tempDir3

```

The weight part corresponding to blood pressure:

```

myDataExpanded$FHtemp <- myDataExpanded$FH
tempDir4 <-
dnorm(myDataExpanded$BP,mean=predict(fitBP,type="response",newdata=
myDataExpanded),sd=sqrt(fitBPvariance))
myDataExpanded$FHtemp <- myDataExpanded$FHInd4
tempIndir4 <-
dnorm(myDataExpanded$BP,mean=predict(fitBP,type="response",newdata=m
yDataExpanded),sd=sqrt(fitBPvariance))
myDataExpanded$weight4 <- tempIndir4/tempDir4

```

The final weight is obtained by multiplying over the four parts:

```

myDataExpanded$weight <-
myDataExpanded$weight1*myDataExpanded$weight2*myDataExpanded$weight3
*myDataExpanded$weight4

```

The stability of the weights can be evaluated by drawing a histogram:

```
hist(myDataExpanded$weight)
```

Finally, natural direct and natural indirect effects can be computed using a natural effects model, for example a Cox proportional hazards model (function `coxph` of the R package `survival`) or an additive hazards model (function `aalen` of the R package `timereg`), applied on the weighted dataset just created.

```

library(survival)
coxph(Surv(time,event) ~
FH+FHInd1+FHInd2+FHInd3+FHInd4+cluster(ID)+age+sex+smoking,
data=myDataExpanded,weights=myDataExpanded$weight)
library(timereg)
aalen(Surv(time,event) ~
const(FH)+const(FHInd1)+const(FHInd2)+const(FHInd3)+const(FHInd4)+co
nst(age)+const(sex)+const(smoking),data=myDataExpanded,
clusters=myDataExpanded$ID, weights=myDataExpanded$weight, robust=T)

```

The parameter estimates for `FH` of these models yield estimates of the natural direct effect, and the parameter estimates for `FHInd1`, `FHInd2`, `FHInd3`, and `FHInd4` yield estimates of the natural indirect effects via apoA-I, apoB, DM, and blood pressure pathways (conditional on confounder variables age, sex, and smoking status).

Instead of relying on the robust standard errors as output by the functions `coxph/aalen`, we use bootstrapping to obtain 95% bootstrap percentile confidence intervals. This is done by encapsulating the previous code in a function, and then applying the function `boot` of the R package `boot`.

The following R code encapsulates the previous commands in the function `NaturalEffectsModel` for a Cox proportional hazards model as the natural effects model.

```
NaturalEffectsModel <- function(ds0,i) {  
  library(survival)  
  ds <- ds0[i,]  
  #Mediator models  
  ds$FHtemp <- ds$FH  
  fitApoA <- lm(ApoA ~ FHtemp+age+sex+smoking, data=ds)  
  fitApoAvariance <- summary(fitApoA)$sigma^2  
  fitApoB <- lm(ApoB ~ FHtemp+age+sex+smoking, data=ds)  
  fitApoBvariance <- summary(fitApoB)$sigma^2  
  fitBP <- lm(BP ~ FHtemp+age+sex+smoking, data=ds)  
  fitBPvariance <- summary(fitBP)$sigma^2  
  fitDM <- glm(DM ~ FHtemp+age+sex+smoking, data=ds,  
  family="binomial")  
  #Expansion of dataset  
  levelsOfFH <- unique(ds$FH)  
  ds1 <- ds  
  ds2 <- ds  
  ds1$FHInd1 <- levelsOfFH[1]
```

```

ds2$FHInd1 <- levelsOfFH[2]

tempds <- rbind(ds1,ds2)

ds1 <- tempds

ds2 <- tempds

ds1$FHInd2 <- levelsOfFH[1]
ds2$FHInd2 <- levelsOfFH[2]
tempds <- rbind(ds1,ds2)

ds1 <- tempds

ds2 <- tempds

ds1$FHInd3 <- levelsOfFH[1]
ds2$FHInd3 <- levelsOfFH[2]
tempds <- rbind(ds1,ds2)

ds1 <- tempds

ds2 <- tempds

ds1$FHInd4 <- levelsOfFH[1]
ds2$FHInd4 <- levelsOfFH[2]

myDataExpanded <- rbind(ds1,ds2)

#Weight corresponding to apoA-I
myDataExpanded$FHtemp <- myDataExpanded$FH
tempDir1 <-
dnorm(myDataExpanded$ApoA,mean=predict (fitApoA,type="response",
newdata=myDataExpanded),sd=sqrt (fitApoAvariance))

myDataExpanded$FHtemp <- myDataExpanded$FHInd1
tempIndir1 <-
dnorm(myDataExpanded$ApoA,mean=predict (fitApoA,type="response",
newdata=myDataExpanded),sd=sqrt (fitApoAvariance))

myDataExpanded$weight1 <- tempIndir1/tempDir1

#Weight corresponding to apoB
myDataExpanded$FHtemp <- myDataExpanded$FH

```



```

tempDir2 <-
dnorm(myDataExpanded$ApoB,mean=predict (fitApoB,type="response",
newdata=myDataExpanded),sd=sqrt (fitApoBvariance))
myDataExpanded$FHtemp <- myDataExpanded$FHInd2
tempIndir2 <-
dnorm(myDataExpanded$ApoB,mean=predict (fitApoB,type="response",
newdata=myDataExpanded),sd=sqrt (fitApoBvariance))
myDataExpanded$weight2 <- tempIndir2/tempDir2
#Weight corresponding to DM status
myDataExpanded$FHtemp <- myDataExpanded$FH
temp <- predict (fitDM,type="response",newdata=myDataExpanded)
tempDir3 <- ifelse(myDataExpanded$DM==1,temp,1-temp)
myDataExpanded$FHtemp <- myDataExpanded$FHInd3
temp <- predict (fitDM,type="response",newdata= myDataExpanded)
tempIndir3 <- ifelse(myDataExpanded$DM==1,temp,1-temp)
myDataExpanded$weight3 <- tempIndir3/tempDir3
#Weight corresponding to blood pressure
myDataExpanded$FHtemp <- myDataExpanded$FH
tempDir4 <-
dnorm(myDataExpanded$BP,mean=predict (fitBP,type="response",newd
ata= myDataExpanded),sd=sqrt (fitBPvariance))
myDataExpanded$FHtemp <- myDataExpanded$FHInd4
tempIndir4 <-
dnorm(myDataExpanded$BP,mean=predict (fitBP,type="response",newd
ata= myDataExpanded),sd=sqrt (fitBPvariance))
myDataExpanded$weight4 <- tempIndir4/tempDir4
#Final weight

```

```

myDataExpanded$weight <-
myDataExpanded$weight1*myDataExpanded$weight2*myDataExpanded$we
ight3*myDataExpanded$weight4
#Final natural effects model
fitNatEffMod <- coxph(Surv(time,event) ~
FH+FHInd1+FHInd2+FHInd3+FHInd4+cluster(ID)+age+sex+smoking,
data=myDataExpanded,weights=myDataExpanded$weight)
#Calculation of natural direct, natural indirect, and total
effects as well as percentages
coef <- fitNatEffMod$coefficients
HRTE <- exp(coef[1]+coef[2]+coef[3]+coef[4]+coef[5])
HRNDE <- exp(coef[1])
HRNIE <- exp(coef[2]+coef[3]+coef[4]+coef[5])
HRNIE1 <- exp(coef[2])
HRNIE2 <- exp(coef[3])
HRNIE3 <- exp(coef[4])
HRNIE4 <- exp(coef[5])
HRTEpct <- log(HRTE)/log(HRTE)*100
HRNDEpct <- log(HRNDE)/log(HRTE)*100
HRNIEpct <- log(HRNIE)/log(HRTE)*100
HRNIE1pct <- log(HRNIE1)/log(HRTE)*100
HRNIE2pct <- log(HRNIE2)/log(HRTE)*100
HRNIE3pct <- log(HRNIE3)/log(HRTE)*100
HRNIE4pct <- log(HRNIE4)/log(HRTE)*100
return(c(HRTE,HRNDE,HRNIE,HRNIE1,HRNIE2,HRNIE3,HRNIE4,HRNDEpct,
HRNIEpct,HRNIE1pct,HRNIE2pct,HRNIE3pct,HRNIE4pct))
}

```

Bootstrap percentile confidence intervals can then be obtained by the function `bootstrap` as defined as follows:

```
bootstrap <- function(pathname,numit){
  results <- boot(myData,NaturalEffectsModel,R=numit,stype="i")
  TE <-boot.ci(results, type = c("perc"),index=1)
  NDE<-boot.ci(results, type = c("perc"),index=2)
  NIE<-boot.ci(results, type = c("perc"),index=3)
  NIE1<-boot.ci(results, type = c("perc"),index=4)
  NIE2<-boot.ci(results, type = c("perc"),index=5)
  NIE3<-boot.ci(results, type = c("perc"),index=6)
  NIE4<-boot.ci(results, type = c("perc"),index=7)
  NDEpct<-boot.ci(results, type = c("perc"),index=8)
  NIEpct<-boot.ci(results, type = c("perc"),index=9)
  NIE1pct<-boot.ci(results, type = c("perc"),index=10)
  NIE2pct<-boot.ci(results, type = c("perc"),index=11)
  NIE3pct<-boot.ci(results, type = c("perc"),index=12)
  NIE4pct<-boot.ci(results, type = c("perc"),index=13)
  sink(pathname,append=FALSE,split=FALSE)
  cat("\n\nNumber of bootstrapping iterations: ",numit)
  cat("\n\nTotal effect: ", round(TE$t0,2),"
  [",round(TE$percent[4],2),"-",round(TE$percent[5],2),"],
  ",round(100,2)," ["",round(100,2),"-",round(100,2),""]\n", sep="")
  cat("Direct effect: ", round(NDE$t0,2),"
  [",round(NDE$percent[4],2),"-",round(NDE$percent[5],2),"],
  ",round(NDEpct$t0,2)," ["",round(NDEpct$percent[4],2),"-
  ",round(NDEpct$percent[5],2),""]\n", sep="")
  cat("Indirect effect: ", round(NIE$t0,2),"
  [",round(NIE$percent[4],2),"-",round(NIE$percent[5],2),"],
```

```

", round(NIEpct$t0, 2), " [" , round(NIEpct$percent[4], 2), "-
", round(NIEpct$percent[5], 2), "]" \n", sep="")

  cat("Indirect effect (Mediator 1): ", round(NIE1$t0, 2), "
[" , round(NIE1$percent[4], 2), "- ", round(NIE1$percent[5], 2), "],
", round(NIE1pct$t0, 2), " [" , round(NIE1pct$percent[4], 2), "-
", round(NIE1pct$percent[5], 2), "]" \n", sep="")

  cat("Indirect effect (Mediator 2): ", round(NIE2$t0, 2), "
[" , round(NIE2$percent[4], 2), "- ", round(NIE2$percent[5], 2), "],
", round(NIE2pct$t0, 2), " [" , round(NIE2pct$percent[4], 2), "-
", round(NIE2pct$percent[5], 2), "]" \n", sep="")

  cat("Indirect effect (Mediator 3): ", round(NIE3$t0, 2), "
[" , round(NIE3$percent[4], 2), "- ", round(NIE3$percent[5], 2), "],
", round(NIE3pct$t0, 2), " [" , round(NIE3pct$percent[4], 2), "-
", round(NIE3pct$percent[5], 2), "]" \n", sep="")

  cat("Indirect effect (Mediator 4): ", round(NIE4$t0, 2), "
[" , round(NIE4$percent[4], 2), "- ", round(NIE4$percent[5], 2), "],
", round(NIE4pct$t0, 2), " [" , round(NIE4pct$percent[4], 2), "-
", round(NIE4pct$percent[5], 2), "]" \n", sep="")

sink()

return()

}

```

The function `bootstrap` performs the bootstrapping with the number of iterations specified by the parameter `numit`, and writes the results of the natural effects model for family history of CHD in a text file in the folder `pathname`.

For our data, we get the following output which is also the basis of Table 2 of the main text:

```

Number of bootstrapping iterations: 2000
Total effect: 1.52 [1.39-1.65], 100 [100-100]
Direct effect: 1.4 [1.28-1.52], 79.96 [73.58-85.21]

```


Indirect effect: 1.09 [1.07-1.11], 20.04 [14.79-26.42]

Indirect effect (Mediator 1): 1.01 [1-1.01], 1.7 [0.2-3.39]

Indirect effect (Mediator 2): 1.04 [1.03-1.05], 8.32 [5.84-11.7]

Indirect effect (Mediator 3): 1.01 [1-1.02], 1.49 [-0.78-3.76]

Indirect effect (Mediator 4): 1.04 [1.03-1.05], 8.53 [5.94-11.93]

Table S1. Total, direct, and indirect effects of family history on incident CHD with metabolic mediators, adjusted for age at baseline and smoking status, stratified by sex

Effects	Men (N=8,973)		Women (N=14,622)	
	Hazard Ratio (95% CI)	Proportion explained (%) (95% CI)†	Hazard Ratio (95% CI)	Proportion explained (%) (95% CI)†
Total effect	1.42 (1.27 to 1.59)	100.0%	1.67 (1.45 to 1.92)	100.0%
Direct effect*	1.34 (1.20 to 1.49)	83.7% (72.4% to 91.7%)	1.50 (1.30 to 1.72)	79.2% (69.0% to 85.9%)
Indirect effect, combined	1.06 (1.03 to 1.09)	16.3% (8.3% to 27.6%)	1.11 (1.08 to 1.15)	20.8% (14.1% to 31.0%)
Indirect effect, through systolic blood pressure and hypertension treatment	1.02 (1.01 to 1.04)	6.4% (2.2% to 12.0%)	1.05 (1.03 to 1.06)	8.8% (5.6% to 13.8%)
Indirect effect, through apoA-I	1.01 (1.00 to 1.02)	2.2% (-0.3% to 5.3%)	1.01 (1.00 to 1.02)	1.7% (-0.1% to 3.9%)
Indirect effect, through apoB	1.03 (1.01 to 1.04)	7.5% (3.8% to 12.8%)	1.04 (1.03 to 1.06)	7.7% (4.8% to 12.4%)
Indirect effect, through diabetes mellitus	1.00 (0.99 to 1.01)	0.2% (-3.5% to 3.6%)	1.01 (1.00 to 1.03)	2.6% (-0.6% to 6.1%)

CI – confidence interval; CHD – coronary heart disease

*Effect of family history not mediated by the four analyzed risk factors.

†On ln(HR) scale.

Table S2. Total, direct, and indirect effects of family history on incident CHD with metabolic mediators, adjusted for age at baseline, sex, and smoking status, in a subgroup of individuals <50 years of age

Effects	<50 years of age (N=4,684)	
	Hazard Ratio (95% CI)	Proportion explained (%) (95% CI)†
Total effect	1.81 (1.31 to 2.53)	100.0%
Direct effect*	1.58 (1.15 to 2.19)	76.5% (48.4% to 89.5%)
Indirect effect, combined	1.15 (1.07 to 1.24)	23.5% (10.5% to 51.6%)
Indirect effect, through systolic blood pressure and hypertension treatment	1.05 (1.02 to 1.09)	7.9% (2.8% to 19.9%)
Indirect effect, through apoA-I	1.02 (1.00 to 1.04)	3.0 (0.4% to 9.5%)
Indirect effect, through apoB	1.08 (1.03 to 1.13)	12.4% (4.7% to 30.0%)
Indirect effect, through diabetes mellitus	1.00 (0.97 to 1.03)	0.2% (-6.1% to 5.4%)

CI – confidence interval; CHD – coronary heart disease

*Effect of family history not mediated by the four analyzed risk factors.

†On ln(HR) scale.

Table S3. Total, direct, and indirect effects of GRS50 (high vs. low/intermediate) on incident CHD with metabolic mediators, adjusted for age at baseline, and smoking status, stratified by sex

Effects	Men (N=8,973)		Women (N=14,622)	
	Hazard Ratio (95% CI)	Proportion explained (%) (95% CI)†	Hazard Ratio (95% CI)	Proportion explained (%) (95% CI)†
Total effect	1.52 (1.35 to 1.70)	100.0%	1.56 (1.32 to 1.83)	100.0%
Direct effect*	1.47 (1.31 to 1.65)	92.6% (84.53% to 100.1%)	1.44 (1.23 to 1.68)	82.3% (69.1% to 91.0%)
Indirect effect, combined	1.03 (1.00 to 1.06)	7.4% (-0.1% to 15.5%)	1.08 (1.04 to 1.13)	17.7% (9.0% to 30.9%)
Indirect effect, through systolic blood pressure and hypertension treatment	1.02 (1.00 to 1.04)	4.9% (0.5% to 10.0%)	1.02 (1.00 to 1.04)	3.7% (-0.1% to 8.3%)
Indirect effect, through apoA-I	1.00 (0.99 to 1.01)	-0.9% (-3.8% to 1.7%)	1.01 (1.00 to 1.02)	2.9% (0.3% to 6.2%)
Indirect effect, through apoB	1.02 (1.01 to 1.04)	5.8% (2.5% to 10.3%)	1.05 (1.03 to 1.07)	10.0% (5.5% to 17.0%)
Indirect effect, through diabetes mellitus	0.99 (0.98 to 1.00)	-2.5% (-6.6% to 0.8%)	1.00 (0.99 to 1.03)	1.1% (-3.6% to 6.0%)

CI – confidence interval; CHD – coronary heart disease

*Effect of GRS50 not mediated by the four analyzed risk factors.

†On ln(HR) scale.

Table S4. Total, direct, and indirect effects of GRS50 (high vs. low/intermediate) on incident CHD with metabolic mediators, adjusted for age at baseline, sex, and smoking status, in a subgroup of individuals <50 years of age

Effects	<50 years of age (N=4,684)	
	Hazard Ratio (95% CI)	Proportion explained (%) (95% CI)†
Total effect	2.01 (1.40 to 2.78)	100.0%
Direct effect*	1.87 (1.30 to 2.59)	89.9% (71.2% to 101.3%)
Indirect effect, combined	1.07 (0.99 to 1.17)	10.1% (-1.3% to 28.8%)
Indirect effect, through systolic blood pressure and hypertension treatment	1.00 (0.97 to 1.03)	-0.2% (-6.4% to 5.0%)
Indirect effect, through apoA-I	1.00 (0.97 to 1.02)	-0.4% (-4.4% to 3.1%)
Indirect effect, through apoB	1.07 (1.02 to 1.14)	10.3% (3.1% to 24.9%)
Indirect effect, through diabetes mellitus	1.00 (0.97 to 1.04)	0.4% (-5.4% to 7.2%)

CI – confidence interval; CHD – coronary heart disease

*Effect of GRS50 not mediated by the four analyzed risk factors.

†On ln(HR) scale.

Table S5. Total, direct, and indirect effects of family history on incident CHD with mediator GRS50, adjusted for age at baseline, sex, and smoking status

Effects	Family history (yes vs. no)		
	Hazard Ratio (95% CI)	Proportion explained (%) (95% CI)†	Additional incident CHD cases per 100,000 person-years at risk‡
Total effect	1.52 (1.40 to 1.65)	100.0%	273.2
Direct effect*	1.48 (1.37 to 1.61)	93.8% (91.4% to 95.7%)	255.7
Indirect effect through GRS50	1.03 (1.02 to 1.03)	6.2% (4.3% to 8.7%)	17.6

CI – confidence interval; CHD – coronary heart disease

*Effect of family history not mediated by GRS50.

†On ln(HR) scale.

‡Estimates from additive hazards models with time-independent effects.

Table S6. Total, direct, and indirect effects of GRS50 on incident CHD with mediator family history, adjusted for age at baseline, sex, and smoking status

Effects	GRS50 (high vs. low)			GRS50 (high vs. low/intermediate)		
	Hazard Ratio (95% CI)	Proportion explained (%) (95% CI) [†]	Additional incident CHD cases per 100,000 person-years at risk [‡]	Hazard Ratio (95% CI)	Proportion explained (%) (95% CI) [†]	Additional incident CHD cases per 100,000 person-years at risk [‡]
Total effect	2.00 (1.75 to 2.29)	100.0%	470.1	1.53 (1.39 to 1.68)	100.0%	315.6
Direct effect*	1.94 (1.69 to 2.21)	95.0% (92.2% to 96.9%)	447.1	1.50 (1.36 to 1.65)	95.8% (93.4% to 97.4%)	304.0
Indirect effect through family history	1.04 (1.02 to 1.05)	5.0% (3.1% to 7.8%)	23.0	1.02 (1.01 to 1.03)	4.2% (2.5% to 6.6%)	11.6

CI – confidence interval; CHD – coronary heart disease

*Effect of GRS50 not mediated by family history.

[†]On ln(HR) scale.

[‡]Estimates from additive hazards models with time-independent effects.