BMJ Open

Cardiovascular risk factor mediation of the effects of education and Genetic Risk Score on cardiovascular disease: a prospective observational cohort study of the Framingham Heart Study

Katie L Powell,[1] Sebastien R Stephens,[2,3] Alexandre S Stephens [4,5,6]

To cite: Powell KL, Stephens SR, Stephens AS. Cardiovascular risk factor mediation of the effects of education and Genetic Risk Score on cardiovascular disease: a prospective observational cohort study of the Framingham Heart Study. *BMJ Open* 2021;**11**:e045210. doi:10.1136/bmjopen-2020-045210

► Prepublication history and additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-045210).

Received 27 September 2020
Revised 24 November 2020
Accepted 16 December 2020

For numbered affiliations see end of article.

Correspondence to
Dr Alexandre S Stephens; alexandre.stephens@health.nsw.gov.au

ABSTRACT

Objectives Level of education and genetic risk are key predictors of cardiovascular disease (CVD). While several studies have explored the causal mechanisms of education effects, it remains uncertain to what extent genetic risk is mediated by established CVD risk factors. This study sought to investigate this and explored the mediation of education and genetic effects on CVD by established cardiovascular risk factors in the Framingham Heart Study (FHS).

Design Prospective observational cohort study.

Participants 7017 participants from the FHS.

Setting Community-based cohort of adults in Framingham, Massachusetts, USA.

Primary outcome measure Incident CVD. The total effects of education and genetic predisposition using a 63-variant genetic risk score (GRS) on CVD, as well as those mediated by established CVD risk factors, were assessed via mediation analysis based on the counterfactual framework using Cox proportional hazards regression models.

Results Over a median follow-up time of 12.0 years, 1091 participants experienced a CVD event. Education and GRS displayed significant associations with CVD after adjustment for age and sex and the established risk factors smoking, total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), body mass index, systolic blood pressure (SBP) and diabetes. For education effects, smoking, HDL-C and SBP were estimated to mediate 18.8% (95% CI 9.5% to 43%), 11.5% (95% CI 5.7% to 29.0%) and 4.5% (95% CI 1.6% to 13.3%) of the total effect of graduate degree, respectively, with the collective of all risk factors combined mediating 38.5% (95% 24.1% to 64.9%). A much smaller proportion of the effects of GRS were mediated by established risk factors combined (17.6%, 95% CI 2.4% to 35.7%), with HDL-C and TC mediating 11.5% (95% CI 6.2% to 21.5%) and 3.1% (95% CI 0.2% to 8.3%), respectively.

Conclusions Unlike education inequalities, established risk factors mediated only a fraction of GRS effects on CVD. Further research is required to elucidate the underlying causal mechanisms of genetic contributions to CVD.

Strengths and limitations of this study

- This study explored the underlying causal mechanisms of education and genetic contributions to cardiovascular disease using the long running, longitudinal and well characterised Framingham Heart Study.
- Mediation analysis of both education and genetic contributions to cardiovascular disease enabled a contemporaneous comparison of the extent to which each were mediated by established risk factors.
- This study adds to the paucity of existing evidence on established risk factor mediation of genetic contributions to cardiovascular disease.
- The study focused specifically on the role of established risk factors in mediating the effects of education and Genetic Risk Score on cardiovascular disease and further studies exploring the role of other risk factors, such as lifestyle factors and health behaviours, are warranted.

INTRODUCTION

Cardiovascular disease (CVD) remains a major contributor to global burden of disease and mortality.[1 2] Estimates placed CVD as causing 17.8 million deaths worldwide in 2017,[2] and affecting a staggering 422.7 million individuals globally in 2015.[1] In the USA alone, CVD accounted for 647 457 deaths in 2017, ranking as the leading cause of death.[3] Although CVD remains a major contributor to morbidity and mortality, CVD mortality is declining in high-income countries,[1 4] linked to improvements in key modifiable risk factors, such as diet, smoking, blood pressure and cholesterol levels,[4] and general increases in the high-end spectrum of the Sociodemographic Index (SDI).[1]

While the SDI has generally increased over recent decades globally,[5] including high-income nations, marked socioeconomic gradients in CVD risk still exist in many

countries.[6 7] For example, a recent study showed that major cardiovascular events were more common among individuals with low levels of education in a collection of 20 low-income, middle-income and high-income countries, and that education was a far stronger predictor of cardiovascular events than wealth.[7]

Another key predictor of CVD is genetic predisposition. Genetic knowledge of CVD continues to expand as demonstrated by recent, large-scale Genome-Wide Association Studies, which identified more novel genetic variants associated with CVD risk.[8 9] Such studies also highlight the polygenic nature of CVD, demonstrating that many genes, each imparting only a small effect, collectively combine to influence CVD risk through the accumulation of risk alleles.[9 10] Genetic Risk Scores (GRSs) have also demonstrated significant predictive benefit when combined with traditional risk factors for CVD,[9 10] highlighting the complex nature of CVD as a condition defined by genetic, biological and environmental influences.

Taken together, established risk factors, education and genetics play key roles in CVD. What is less known, though, is how these factors interact to influence CVD risk, and, in particular, the extent to which established risk factors, such as cholesterol levels, blood pressure, body mass index (BMI) and smoking, mediate the effects of education and genetic risk on CVD risk, particularly genetic contributions. To this end, we undertook a contemporaneous analysis of the effects of education, GRS and established risk factors on CVD in the Framingham Heart Study (FHS). The main objective of the study was to investigate the extent to which genetic effects on CVD were mediated by established risk factors in comparison to the degree to which the effects of education were mediated by the same risk factors. The established risk factors investigated in the study included total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), cigarette smoking and systolic blood pressure (SPB), as key components of the Framingham Risk Score for Hard Coronary Heart Disease (CHD),[11] and BMI and diabetes status, which have also been implicated as mediators of education effects on CVD.[12] The mediating effects of these six risk factors were considered individually and in combination, and a subset of the risk factors displaying the largest contributions to genetic and education effects on CVD were assessed using causally ordered mediation analysis.

METHODS

Study sample

The study sample included participants from the multigeneration, community-based FHS. The FHS is a long running and well-characterised longitudinal study of free-living adults residing in the Town of Framingham, Massachusetts, USA, which commenced in 1948 via the enrolment of 5209 individuals into the Original Cohort. In 1971, the FHS was expanded by incorporating an Offspring cohort, which included 5124 participants comprising of children and their spouses, of Original Cohort members. In 2002, a third-generation cohort commenced enrolment and includes 4095 children from the Offspring cohort. Comprehensive, in-person examinations were conducted for each participant at baseline with study data collected through a battery of tests, examinations, anthropometric measurements and questionnaires specifically designed for the study. Follow-up examinations occurred every 2–6 years for each cohort. A full description of the FHS, including study design, participant follow-up, examination cycles and phenotype and outcome measures, has been described previously.[13] Data for 5013 offspring cohort and 4078 third-generation cohort participants were available (total N=9091). Of these, 2074 were excluded due to lack of genotype data (did not consent for DNA studies or incomplete genotype data) or missing covariate information. Final analyses included 7017 participants.

Patient and public involvement

Patients and the public were not directly involved in this study.

Study outcome variable: incident CVD events

The primary study outcome variable was incident CVD. In accordance with previously described criteria,[14] incident CVD was defined as a composite of CHD (coronary death, myocardial infarction, coronary insufficiency and angina), cerebrovascular events (including ischaemic stroke, haemorrhagic stroke and transient ischaemic attack), peripheral artery disease (intermittent claudication) and heart failure.

Main study factors: education level and GRS

The main study factors were education and GRS. Level of education was derived from years education or highest degree obtained, which were measured at exams 2 and 8 for the Offspring cohort and exam 1 for the third-generation cohort. Education level was categorised into three levels representing: (1) high school or less; (2) more than high school or bachelor's degree, but less than graduate degree and (3) graduate degree, which broadly equates to ≤12, 13–16 and ≥17 years of education as used previously.[15 16]

Detailed information on FHS genotyping and genotype imputation have been previously described.[8] Imputed genetic data were used to construct a 63-variant GRS for CVD based on previously reported genetic variants (online supplemental table 1).[8] A weighted GRS was constructed by multiplying genetic dose of risk alleles by the natural log of the risk estimate for each variant and summing the products across all variants. Higher scores indicate higher genetic predisposition to CVD. Rare or low frequency variants with a minor allele frequency <5% were excluded from the GRS.

Cardiovascular risk factors and covariates

Information on cardiovascular risk factors and covariates were obtained from examination data as described

previously.[13 17] Study covariates included age in years and sex (female or male), which were obtained at baseline examination. TC and HDL-C levels were determined on fasting blood samples and measured in units of mg/dL; these were converted to units of mmol/L in the study. Cigarette smoking (smoking) was determined by self-report and defined as smoking regularly in the year prior to baseline examination. Diabetes was defined as a fasting plasma glucose ≥126 mg/dL or use of glucose-lowering medication.[18] SPB was measured using mercury sphygmomanometer in subjects seated for at least 5 min. BMI was derived from height and weight information, measured at baseline examination, as weight in kilograms (kg) divided by height in metres squared (m^2).

Statistical analysis

Characteristics and risk factor measurements of the study participants were analysed descriptively via the calculation of mean and SD for continuous variables, counts and percentages for categorical data, and median and IQR for year's follow-up data. Cox proportional hazards regression models were used to explore associations between the main study factors and CVD while adjusting for covariates. Times to incident events (or censoring) were based on follow-up times from baseline examination. Unadjusted and adjusted associations between the main study factors, education level and GRS (1-SD increase), and CVD were assessed in models including the study factor only (unadjusted models) and models additionally adjusting for the six cardiovascular risk factors, which are the explicit focus of the study, and age and sex (adjusted models). Risk factors and covariates were included in models as baseline variables and were not time varying. To account for family structure (and correlations) in the data, frailty terms clustering on family were included in Cox regression models. Proportional hazards assumptions were assessed graphically and statistically using scaled Schoenfeld residuals. Multivariable linear regressions (for continuous risk factors) and binary logistic regressions (for binary risk factors) were used to explore the associations between each of the main study factors and cardiovascular risk factors. In statistical models, age (years), GRS (weighted risk score), TC (mmol/L), HDL-C (mmol/L), BMI (kg/m^2) and systolic blood pressure (SBP) (mm Hg) were included as mean centred, standardised continuous variables (ie, expressed in units of SDs away from means of 0). Therefore, a 1-unit increase in the value of these continuous variables in statistical models corresponds to a 1-SD increase. Sex (male or female), education (three level as described in main study factors above), smoking (yes or no) and diabetes (yes or no) were included in statistical models as categorical variables. Measures of associations derived from statistical models are displayed as HRs, ORs or standardised mean differences with 95% CIs.

For mediation analysis, two separate approaches based on the counterfactual framework were applied. The first approach was based on the methods for estimating causal effects described by VanderWeele for Cox proportional hazards regression.[19] Using this approach, the total effect of an exposure on outcome (CVD) was decomposed into the natural direct effect (NDE) and the natural indirect effect (NIE), where the NIE is interpreted as the effect of the exposure on outcome mediated by another risk factor (ie, the mediator). The proportion of the effects of education or GRS mediated by established cardiovascular risk factors were also calculated. CIs for the NDE, NIE and the proportion mediated were obtained via bootstrap resampling with 1000 replications. Separate analyses were undertaken to assess the mediation effects of each risk factor individually. In these analyses, all other risk factors, which were not treated as the mediator variable, were included in models as covariates. Analyses were also undertaken investigating the mediation effects of all six risk factors combined. We used the product coefficient method to calculate indirect effects for risk factors individually and the difference in coefficient method to investigate all six risk factors combined.[12 19] It is important to note that assessing each risk factor individually and combined, while valid approaches (eg, Carter *et al*[12]), fails to take into account any relationships that may exist between the mediators if any sit of the causal pathway of others. Sensitivity analyses evaluating how much unmeasured mediator outcome confounding would explain away the observed NDEs and NIEs were also undertaken.[20]

A second approach to mediation was applied based on the methods espoused in Cho and Huang.[21] Briefly, using this approach, path-specific effects (PSEs) of the exposure on the outcome in the form of transformed survival time using Cox proportional hazards regression models through multiple, causally ordered mediators was assessed. This method caters for mediation effects to be assessed in causally ordered pathways based on hypothesised causal relationships between mediators and, thus, accommodates for the complexity that can arise if mediators sit on the causal pathway of others, which is a main advantage of the method. The approach decomposes the total effects into PSEs relating the associations between the exposure and outcome through direct effects and effects through the mediators. The results of this approach have intuitive interpretations whereby the magnitude and direction (sign) of the PSEs are related to the effect on survival (eg, a positive PSE is associated with increased survival). The notation used to describe the PSEs in this study are represented by deltas (Δ). As an example, in a hypothetical three-mediator model (mediators M_1, M_2 and M_3) relating an exposure (S) to an outcome (Y): $\Delta_{S\rightarrow Y}$ represents the effect of S on Y not mediated through any of the mediators; $\Delta_{S\rightarrow M_1 Y}$ represents the effect of S on Y mediated through M_1 and possibly M_2 and M_3. $\Delta_{S\rightarrow M_2 Y}$ represents the effect of S on Y mediated through M_2 and possibly M_3; and $\Delta_{S\rightarrow M_3\rightarrow Y}$ represents the effect of S on Y mediated through M_3. This approach was used to support the findings of decomposing total effects into NDEs and NIEs without the rare outcome assumption, while also catering for multimediator settings. Separate multimediator models were used to investigate GRS

and education effects on CVD. The risk factors selected for inclusion as mediators were based on the findings of the individual Cox proportional hazards regression mediation analyses in combination with existing literature. For GRS contributions to CVD, a three-mediator model including BMI, HDL-C and TC in hypothesised causal order (figure 1A)[22 23] was undertaken. For education effects on CVD, a three-mediator model including Smoking, HDL-C and SBP in hypothesised causal order (figure 1B)[22 24] was undertaken.

Phenotype and genotype data extraction were performed in the University of Sydney's High-Performance Computing environment. Subsequent data manipulation, including preparation of the final dataset for analysis and statistical analyses were carried out in R Studio with R V.3.6.0 (https://www.r-project.org) using the survival package for Cox proportional hazards regression analyses. The R code from Cho and Huang[21] was adapted for causally ordered multimediator analysis.

RESULTS

Study cohort characteristics

Table 1 displays the characteristics of the study sample for all participants combined and by education level. Of the 7017 participants, just over half (52.7%) were female, the average age at baseline examination was 38 years and 1716 (24.5%), 4145 (59.0%) and 1156 (16.5%) had education of high school or less, more than high school or bachelor's, or graduate degree level, respectively. The mean GRS was 3.5 and varied little by education level. There was a total of 1091 CVD events, with a median follow-up of 12.0 years. Mean or percentage values of cardiovascular risk factors and the percentage of participants with CVD events were lower in those with higher education (table 1).

Unadjusted effects of education level and GRS on CVD

Higher levels of education were associated with decreased risk of CVD with HRs of 0.67 (95% CI 0.59 to 0.76) and 0.44 (95% CI 0.36 to 0.54) for more than high school or bachelor's degree and graduate degree, respectively, relative to high school or less. A 1-SD increase in the accumulation of genetic risk alleles was associated with increased risk of CVD (HR 1.15, 95% CI 1.08 to 1.22).

Adjusted effects of education level and GRS on CVD

Figure 2 displays the adjusted effects of education level and GRS on CVD risk. For education, effects were markedly attenuated with adjustment for cardiovascular risk factors and study covariates, with education of more than high school or bachelor's degree no longer reaching statistical significance (the 95% CI includes the null value of 1), and graduate degree showing an adjusted

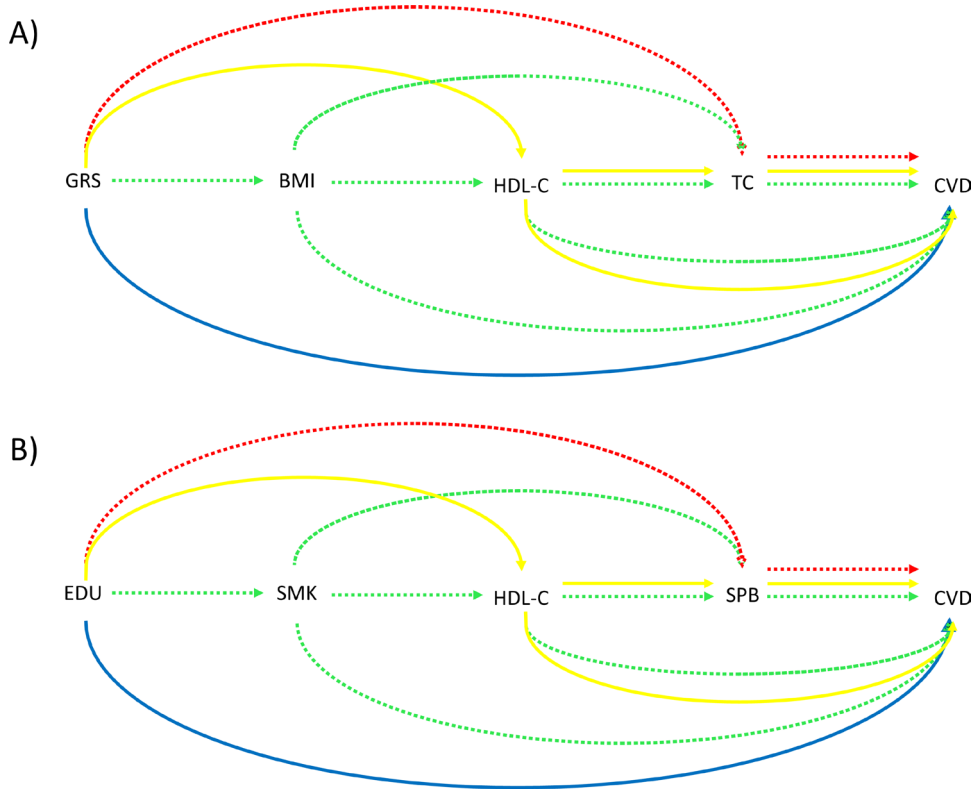

Figure 1 Causal diagrams of education and genetic contributions to cardiovascular disease (CVD). (A) Causal pathways relating Genetic Risk Score (GRS), the causally order mediators body mass index (BMI), high-density lipoprotein cholesterol (HDL-C) and total cholesterol (TC) and CVD. (B) Causal pathways relating graduate degree education (EDU), the causally order mediators smoking (SMK), HDL-C and systolic blood pressure (SPB) and CVD. The four path-specific effects are denoted by different colours: Green: $\Delta_{S\to M_1 Y}$; Yellow: $\Delta_{S\to M_2 Y}$; Red: $\Delta_{S\to M_3\to Y}$; and Blue: $\Delta_{S\to Y}$.

Table 1 Descriptive characteristics of study cohort participants by education level

Characteristic	Total	High school or less	More than high school or bachelor's	Graduate degree
No of participants	7017	1716	4145	1156
Age, years	38.0 (9.7)	39.8 (9.6)	37.3 (9.5)	37.5 (9.9)
Females, N (%)	3700 (52.7)	912 (53.1)	2223 (53.6)	565 (48.9)
Total cholesterol, mmol/L	5.4 (1.2)	5.8 (1.2)	5.3 (1.1)	5.3 (1.0)
HDL cholesterol, mmol/L	1.4 (0.4)	1.3 (0.4)	1.4 (0.4)	1.4 (0.4)
BMI, kg/m^2	28.2 (5.7)	29.4 (5.8)	28.0 (5.7)	27.4 (5.3)
SBP, mm Hg	118.5 (14.9)	121.6 (15.7)	117.6 (14.5)	116.7 (14.4)
GRS	3.5 (0.3)	3.4 (0.4)	3.5 (0.3)	3.5 (0.4)
Smoking, N (%)	3596 (51.2)	1127 (65.7)	2044 (49.3)	425 (36.8)
Diabetes, N (%)	656 (9.3)	276 (16.1)	311 (7.5)	69 (6.0)
CVD				
No of events, N (%)	1091 (15.5)	446 (26.0)	532 (12.8)	113 (9.8)
Follow-up (years), median (IQR)	12.0 (29.5)	13.1 (29.4)	11.7 (28.9)	12.0 (29.9)

Data are presented as means (SD) unless specified otherwise.
BMI, body mass index; CVD, cardiovascular disease; GRS, Genetic Risk Score; HDL, high density lipoprotein; SBP, systolic blood pressure.

HR (aHR) of 0.72 (95% CI 0.58 to 89). In contrast, the estimated effect of a 1-SD increase in GRS remained largely unchanged (figure 2). Among the cardiovascular risk factors, strong effects for smoking (aHR 1.50, 95% CI 1.31 to 1.71), diabetes mellitus (1.48, 95% CI 1.28 to 1.71) and HDL-C (0.75, 95% CI 0.69 to 0.81) were seen. TC, BMI and SBP were also associated with CVD (figure 2).

Effects of GRS and education on established cardiovascular risk factors

Figure 3 displays the effects of education and GRS on established cardiovascular risk factors. While GRS was either not or only marginally associated with differences in risk factors, higher levels of education were associated with lower BMI, SBP and TC, decreased likelihood of smoking and diabetes, and higher HDL-C (figure 3).

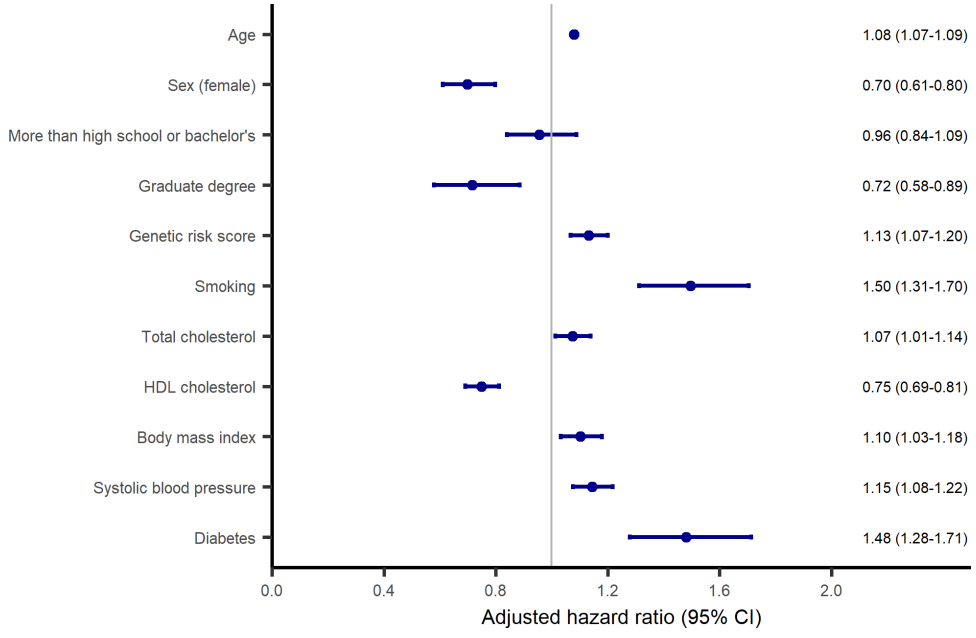

Figure 2 Adjusted[a] associations between explanatory variables and incident cardiovascular disease. Estimates represent Cox proportional hazards regression model adjusted HRs with 95% CIs. [a]Adjusted for all explanatory variables displayed in figure. HDL, high-density lipoprotein.

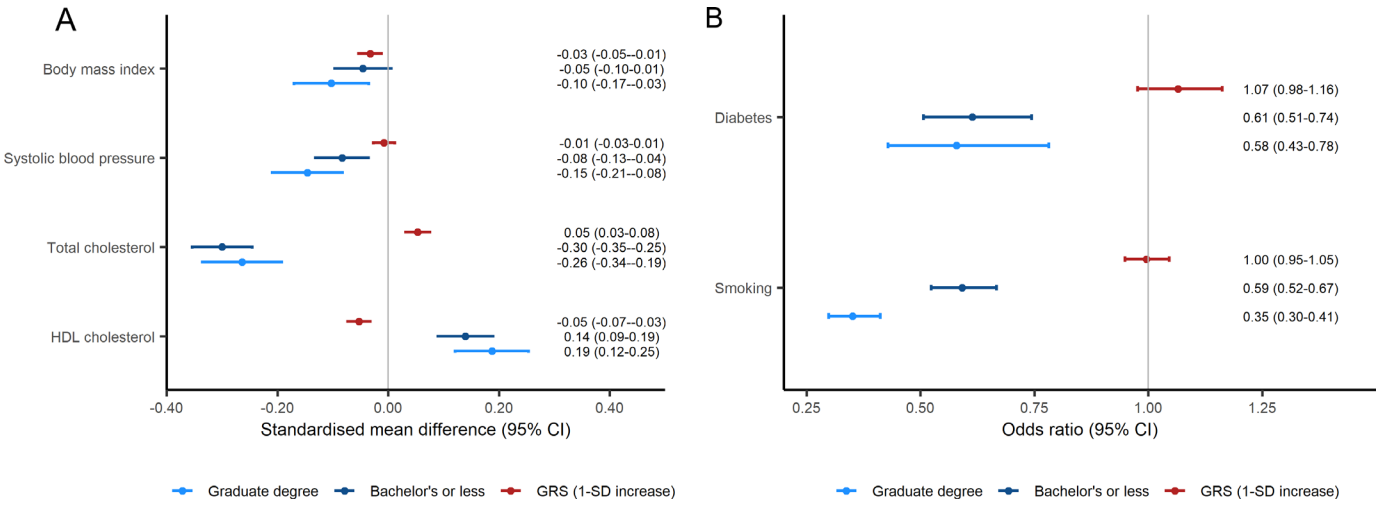

Figure 3 Effects of education and Genetic Risk Score (GRS) on cardiovascular disease risk factors. For continuous risk factors (A), estimated effects represent adjusted[a] standardised mean differences in risk factor with 95% CIs for specified level of education (relative to high school or less) or increase GRS. For binary risk factors (B), estimates represent adjusted[a] ORs of the relationship between main study factor and binary risk factor with 95% CIs. The estimated effects for GRS represent the change in standardised mean difference or odds of risk factor for each 1-SD increase in accumulation of risk alleles. [a]Adjusted for age, sex, smoking, total cholesterol, HDL cholesterol, body mass index, systolic blood pressure and diabetes unless treated as the outcome variable. HDL, high-density lipoprotein.

Mediation of education and GRS effects on CVD by established cardiovascular risk factors

Table 2 displays the total effects, NDEs and NIEs of education (graduate degree) and GRS on CVD in analyses assessing mediation by each established cardiovascular risk factor individually and combined. GRS was associated with increased rate of CVD, due to both direct and indirect effects, with mediation by HDL-C and TC (in a model jointly exploring their effects, HDL-C and TC mediated 15% of the effect of GRS) and BMI (table 2).

Table 2 Estimated adjusted HRs for total effects, natural direct effects and natural indirect effects and proportion mediated through established risk factors for the association between genetic risk score and cardiovascular disease (CVD), and graduate degree and CVD

	Adjusted* HR (95% CI)			Proportion mediated
Risk factor	Total effect	Natural direct effect	Natural indirect effect	(%; 95% CI)
Genetic Risk Score				
BMI	1.13 (1.06 to 1.21)	1.13 (1.07 to 1.21)	1.00 (0.99 to 1.00)	–2.58 (–7.28 to –0.44)
SBP	1.13 (1.07 to 1.21)	1.13 (1.07 to 1.20)	1.00 (1.00 to 1.00)	–0.85 (-4.15 to 1.38)
HDL Cholesterol	1.15 (1.08 to 1.22)	1.13 (1.07 to 1.20)	1.02 (1.01 to 1.02)	11.53 (6.20 to 21.51)
Total cholesterol	1.14 (1.08 to 1.21)	1.13 (1.07 to 1.21)	1.00 (1.00 to 1.01)	3.14 (0.21 to 8.25)
Diabetes	1.14 (1.07 to 1.21)	1.13 (1.07 to 1.21)	1.00 (1.00 to 1.00)	0.84 (-0.34 to 3.37)
Smoking	1.13 (1.06 to 1.20)	1.13 (1.07 to 1.20)	1.00 (1.00 to 1.00)	–0.27 (-4.60 to 3.67)
All ERFs	1.16 (1.10 to 1.24)	1.13 (1.06 to 1.20)	1.03 (1.00 to 1.05)	17.56 (2.39 to 35.70)
Graduate degree				
BMI	0.71 (0.57 to 0.88)	0.72 (0.58 to 0.89)	0.99 (0.98 to 1.00)	2.36 (0.53 to 9.44)
SBP	0.70 (0.57 to 0.88)	0.71 (0.57 to 0.88)	0.98 (0.97 to 0.99)	4.51 (1.62 to 13.28)
HDL Cholesterol	0.70 (0.57 to 0.85)	0.72 (0.58 to 0.88)	0.95 (0.92 to 0.97)	11.51 (5.67 to 29.04)
Total Cholesterol	0.71 (0.56 to 0.87)	0.72 (0.57 to 0.88)	0.98 (0.96 to 1.00)	4.56 (0.50 to 14.94)
Diabetes	0.70 (0.58 to 0.86)	0.71 (0.59 to 0.86)	0.99 (0.98 to 1.00)	2.82 (0.96 to 8.72)
Smoking	0.68 (0.54 to 0.83)	0.71 (0.57 to 0.88)	0.91 (0.87 to 0.94)	18.77 (9.45 to 42.50)
All ERFs	0.58 (0.46 to 0.70)	0.71 (0.57 to 0.88)	0.81 (0.76 to 0.86)	38.48 (24.13 to 64.86)

*Analyses adjusted for age and sex and also included all other cardiovascular risk and study factors, which were not the subject of mediation analysis, as explanatory covariates.
BMI, body mass index; ERFs, established risk factors; HDL, high density lipoprotein; ; SPB, systolic blood pressure.

Table 3 Estimated path-specific effects in the form of transformed survival time relating the main study factors, genetic risk score and graduate degree education, with cardiovascular disease via cardiovascular risk factor mediators

Path-specific effect		95% CI limit			
		Estimate	Lower	Upper	P value
Genetic risk score					
$\Delta_{S\rightarrow Y}$	GRS → CVD	−0.121	−0.179	−0.063	<0.001
$\Delta_{S\rightarrow M_3\rightarrow Y}$	GRS → TC → CVD	−0.005	−0.009	−0.002	0.005
$\Delta_{S\rightarrow M_2 Y}$	GRS → HDL-C → TC → CVD	−0.014	−0.021	−0.007	<0.001
$\Delta_{S\rightarrow M_1 Y}$	GRS → BMI → HDL-C → TC → CVD	0.003	−0.001	0.007	0.178
Graduate degree					
$\Delta_{S\rightarrow Y}$	EDU → CVD	0.323	0.111	0.536	0.003
$\Delta_{S\rightarrow M_3\rightarrow Y}$	EDU → SPB → CVD	0.020	0.007	0.032	0.002
$\Delta_{S\rightarrow M_2 Y}$	EDU → HDL-C → SPB → CVD	0.054	0.030	0.078	<0.001
$\Delta_{S\rightarrow M_1 Y}$	EDU → SMK → HDL-C → SPB → CVD	0.099	0.064	0.135	<0.001

BMI, body mass index; CVD, cardiovascular disease; EDU, graduate degree education; GRS, Genetic Risk Score; HDL-C, high-density lipoprotein cholesterol; SMK, cigarette smoking; SPB, systolic blood pressure; TC, total cholesterol.

For education, graduate degree was associated with decreased risk of CVD, due to both direct and indirect effects, with mediation by established cardiovascular risk factors. Smoking, HDL-C, TC and SBP were the strongest mediators of graduate degree effects on CVD (table 2). A model exploring the effects of established cardiovascular risk factors combined indicated they collectively mediated 38.5% of the association between graduate degree and CVD (table 2). Sensitivity analyses of unmeasured mediator-exposure-outcome confounding for NDEs and NIEs indicated that fairly substantial confounding would be required to explain away the NDEs, but smaller levels of confounding could explain away the NIEs (online supplemental table 2).

The results of causally ordered multimediator analyses are displayed in table 3. The effect of GRS on transformed survival time independent of BMI, HDL-C and TC was significant ($\Delta_{S\rightarrow Y}$ = −0.121, 95% CI −0.179 to −0.063, p<0.001). The effect of GRS on transformed survival time mediated through BMI, and possibly HDL-C and TC was not significant ($\Delta_{S\rightarrow M_1 Y}$ = 0.003, 95% CI −0.001 to 0.007, p=0.178). The effect of GRS mediated by HDL-C and possibly TC was significant ($\Delta_{S\rightarrow M_2 Y}$ = −0.014, 95% CI −0.021 to −0.007, p ≤0.001) as was the effect mediated through TC ($\Delta_{S\rightarrow M_3\rightarrow Y}$ = −0.005,−0.009 to −0.002, p-value=0.005). For education, graduate degree was associated with increased transformed survival time independent of smoking, HDL-C and SBP ($\Delta_{S\rightarrow Y}$ = 0.323, 95% CI 0.111 to 0.536, p=0.003). The effect of graduate degree on transformed survival time mediated by decreased likelihood of smoking and possibly HDL-C and SBP was significant ($\Delta_{S\rightarrow M_1 Y}$ = 0.099, 95% CI 0.064 to 0.135, p<0.001), as was the effect mediated through HDL-C and possibly SBP ($\Delta_{S\rightarrow M_2 Y}$ = 0.054, 95% CI 0.030 to 0.078, p ≤0.001). The effect of graduate degree only mediated by SBP was also significant ($\Delta_{S\rightarrow M_3\rightarrow Y}$ = 0.020, 95% CI 0.007 to 0.032, p=0.002).

DISCUSSION

This study explored the effects of education and GRS on CVD while seeking to decompose total effects into those that were direct and those that were mediated by established cardiovascular risk factors. The findings showed independent effects for education and GRS on CVD with adjustment for established risk factors and study covariates. However, these effects were attenuated relative to unadjusted estimates. Decomposition of total effects into indirect or PSEs demonstrated considerable mediation by established risk factors for the effects of education, but not for GRS, where the effects (>80%) appeared largely independent of established risk factors. Collectively, these findings demonstrate the importance of established risk factors in mediating the effects of education on CVD and highlight that the mechanisms underlying the genetic contributions to CVD, in contrast to education, are largely independent of established risk factors.

The exploration of the effects of education on CVD revealed associations through direct and indirect effects. Although a considerable proportion of the total effect of education was explained by established risk factors, more than half remained unexplained and warrants further investigation. Our results showed stronger mediating effects for smoking and HDL-C, with smaller effects for other established risk factors through individual causal mediation analyses. These findings were supported by causally ordered mediation analysis, which showed significant PSEs involving smoking, HDL-C and SBP as causally ordered mediators. Collectively, these findings are consistent with those of several other studies.[12 22] For

example, Carter *et al*[12] demonstrated, using both observational and mendelian randomisation methods, mediating effects for smoking behaviour (19% of total effect for the observational method), SBP (11%) and BMI (15%) in a large study of European ancestry participants. The total mediating effect of all three risk factors combined in this study was 42%.[12] An earlier study in Dutch participants showed mediating effects for smoking (27%) and several other risk factors, including hypertension (5.3%) and hypercholesterolaemia (3.5%), with the collective of established risk factors in this study accounting for 57% of the total effect of education on CHD.[22] Our findings are consistent with these studies and adds to the growing body of knowledge on the effects of education on CVD as a confirmatory and repeatable research contribution in an independent study cohort.

Independent effects for GRS on CVD were also observed. In contrast to education, GRS effects changed little with adjustment for established risk factors, indicating GRS effects operated largely independently of them, which is consistent with previous studies exploring the genomic prediction of CVD based on GRSs.[8 9] Causal mediation analysis revealed mediation by HDL-C and, to a lesser degree, TC. Notably, mediation by BMI was also observed in individual mediation analysis, but in the opposite direction to HDL-C and TC through its negative association with GRS. Similar findings were found when these three factors were assessed via causally ordered mediation analysis, although there was no longer any evidence for mediation by BMI. In contrast to education effects on CVD, prior research exploring mediation effects of established risk factors on the effect of GRS on CVD is scarce. Fritz *et al*[25] explored the mediating effects of several established risk factors, including apoA1, apoB, SBP and diabetes mellitus, on the effect of a 50-single nucleotide polymorphism GRS on CHD. This study demonstrated that only a fraction (<20%) of the genetic effect was explained by established risk factors,[25] and our study provides evidence to support these findings.

Exploring the mediating effects of education on CVD risk by established risk factors, while previously assessed in several other studies, concurrently with GRS enables the genetic contributions to CVD, and the lack of mediation by established risk factors, to be contextualised. In particular, while approximately 40% of the total effect of education on CVD was mediated by the established risk factors assessed in the study, the proportion mediated was less than half that for genetic contributions. The study findings highlight the complexity of using genetic risk information, and the challenge faced by clinicians and healthcare professionals to manage patients with high genetic risk. Within this paradigm, while acknowledging that education inequalities in health and disease remains a wicked problem (eg, how do you intervene on education?), in the context of CVD, doctors and healthcare professionals can equip themselves with the knowledge that a significant proportion of education effects are mediated by established risk factors, and accordingly, prescribe relevant treatments, such as lifestyle modification, exercise and pharmacological therapies.[12 22 26] Importantly, there is evidence that education effects on CVD are mediated through factors other than those assessed in this study. Lifestyle factors, health behaviours and other risk factors such as diet, alcohol consumption, physical activity and health literacy have all been demonstrated to mediate education contributions to CVD.[16 22 27–30] Admittedly, these factors are likely interrelated, and may sit on the causal pathways of each other (as indicated by some studies, eg, Carter *et al*[12] and Nordahl *et al*[29]), but they collectively represent important factors that may be more readily intervened than education itself.

The case for genetic risk, which operates largely independently of established risk factors and where there is lack of knowledge on the underlying mechanisms of increased risk, is not as straight forward. This could create a dilemma for clinicians, who may become aware that patients are at increased risk of CVD, but have little information about the potential underlying causes.[26] However, as per Abraham *et al*,[10] which showed that modifiable risk factors displayed large effects on cumulative CHD risk in individuals of high genetic predisposition, improving lifestyles may compensate for increased genetic risk. A recent editorial by Tada *et al*[26] also nicely summarises the latest evidence for addressing increased genetic risk, proposing several strategies. The first strategy is to prescribe statins, which was demonstrated to have both relative and absolute benefits in individuals of high genetic risk compared with those with lower risk.[31] The second approach is to promote healthier lifestyles, with evidence indicating that adherence to healthy lifestyles is associated with marked reductions in risk compared with non-adhering individuals.[32] However, the need for more research to better understand the causal mechanisms of increased genetic risk is warranted.

The study had several strengths and limitations. Study findings were based on the analysis of the FHS, a long running and well characterised longitudinal cohort study with extensive follow-up. The study benefited from the availability of rich information on numerous established cardiovascular risk factors, including many which were instrumental in establishing current risk algorithms for CVD used in clinical practice.[13] There was also comprehensive genetic data available based on genotyping arrays and imputation. Two methods of mediation analysis were applied, including a novel multi-mediator method, which is not dependent on the rare outcome assumption,[21] and sensitivity analyses provided information on the size of unmeasured confounding required to explain away observed NDEs and NIEs. Although the final sample size included over 7000 participants and 1000 CVD events, replication of mediation analyses in other, independent cohorts is warranted. Assessment of other covariates and risk factors as mediating variables for the effects of education and GRS for CVD would also be beneficial, especially given the findings of sensitivity analyses, which demonstrated that unmeasured confounders with rather small

effect sizes could easily explain away the observed indirect effects. It is a limitation of the approach used to assess the mediation effects of each cardiovascular risk factor individually and combined that it does not take into consideration possible interrelationships between risk factors if any sit on the causal pathway of others. Although the characteristics of those excluded from the study due to missing data were similar to the final study population (data not shown), it is not known how their inclusion would have impacted the study findings. It is a limitation that study risk factors and covariates were not included in Cox regression models as time-varying variables.

CONCLUSIONS

Exploration of the effects of education and GRS on CVD revealed important mediation by established cardiovascular risk factors. While a considerable proportion of the total effect of education was mediated by established risk factors, this was not true of GRS. Collectively, the findings highlight the relevance of established risk factors in mediating the effects of education on CVD and that further research is required to elucidate the underlying causal mechanisms for the genetic contributions to CVD.

Author affiliations
[1]Faculty of Health Sciences and Medicine, Bond University, Gold Coast, Queensland, Australia
[2]Orthopaedics, Gold Coast Hospital and Health Service, Southport, Queensland, Australia
[3]School of Medicine, Griffith University Faculty of Health, Gold Coast, Queensland, Australia
[4]Clinical Governance, Northern NSW Local Health District, Lismore, New South Wales, Australia
[5]School of Public Health, The University of Sydney Faculty of Medicine and Health, Sydney, New South Wales, Australia
[6]School of Health and Human Sciences, Southern Cross University, Lismore, New South Wales, Australia

Contributors ASS contributed to study conception and design, data and statistical analysis, interpretation of study findings and drafting and critical review of the manuscript. KLP contributed to data and statistical analysis, interpretation of study findings and drafted the manuscript. SRS contributed to the interpretation of study findings and drafting and critical review of the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval Ethical approval was obtained from The University of Sydney Human Research Ethics Committee (2018/146), and approval to access FHS data via the Database of Genotypes and Phenotypes (Authorised Access) was obtained from the National Heart, Lung, and Blood Institute Data Access Committee.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. Data analysed as part of this study were obtained from the Framingham Heart Study under authorised access granted by the National Heart, Lung and Blood Institute Data Access Committee. Applications for access to this data are to be made to the National Heart, Lung, and Blood Institute Data Access Committee as per guidance provided at: https://dbgap.ncbi.nlm.nih.gov/aa/wga.cgi?page=login.

ORCID iD
Alexandre S Stephens http://orcid.org/0000-0003-2465-8117

REFERENCES

1 Roth GA, Johnson C, Abajobir A, *et al*. Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015. *J Am Coll Cardiol* 2017;70:1–25.
2 Mensah GA, Roth GA, Fuster V. The Global Burden of Cardiovascular Diseases and Risk Factors: 2020 and Beyond. *J Am Coll Cardiol* 2019;74:2529–32.
3 Heron M. Deaths: leading causes for, 2017. Available: https://www.cdc.gov/nchs/data/nvsr/nvsr68/nvsr68_06-508.pdf [Accessed 1 Mar 2020].
4 Jagannathan R, Patel SA, Ali MK, *et al*. Global updates on cardiovascular disease mortality trends and Attribution of traditional risk factors. *Curr Diab Rep* 2019;19:44.
5 Global burden of disease study 2015 (GBD 2015). socio-demographic index (SDI) 1980–2015 | GHDx. Available: http://ghdx.healthdata.org/record/ihme-data/gbd-2015-socio-demographic-index-sdi-1980%E2%80%932015 [Accessed 1 Mar 2020].
6 Stringhini S, Carmeli C, Jokela M, *et al*. Socioeconomic status and the 25 × 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1·7 million men and women. *Lancet* 2017;389:1229–37.
7 Rosengren A, Smyth A, Rangarajan S, *et al*. Socioeconomic status and risk of cardiovascular disease in 20 low-income, middle-income, and high-income countries: the prospective urban rural epidemiologic (pure) study. *Lancet Glob Health* 2019;7:e748–60.
8 Nikpay M, Goel A, Won H-H, *et al*. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015;47:1121–30.
9 Inouye M, Abraham G, Nelson CP, *et al*. Genomic risk prediction of coronary artery disease in 480,000 adults: implications for primary prevention. *J Am Coll Cardiol* 2018;72:1883–93.
10 Abraham G, Havulinna AS, Bhalala OG, *et al*. Genomic prediction of coronary heart disease. *Eur Heart J* 2016;37:3267–78.
11 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;285:2486–97.
12 Carter AR, Gill D, Davies NM, *et al*. Understanding the consequences of education inequality on cardiovascular disease: Mendelian randomisation study. *BMJ* 2019;365:l1855.
13 Tsao CW, Vasan RS. Cohort profile: the Framingham heart study (FHS): overview of milestones in cardiovascular epidemiology. *Int J Epidemiol* 2015;44:1800–13.
14 D'Agostino RB, Vasan RS, Pencina MJ, *et al*. General cardiovascular risk profile for use in primary care: the Framingham heart study. *Circulation* 2008;117:743–53.
15 Loucks EB, Lynch JW, Pilote L, *et al*. Life-Course socioeconomic position and incidence of coronary heart disease: the Framingham offspring study. *Am J Epidemiol* 2009;169:829–36.
16 Loucks EB, Gilman SE, Howe CJ, *et al*. Education and coronary heart disease risk: potential mechanisms such as literacy, perceived constraints, and depressive symptoms. *Health Educ Behav* 2015;42:370–9.
17 Thanassoulis G, Peloso GM, Pencina MJ, *et al*. A genetic risk score is associated with incident cardiovascular disease and coronary artery calcium: the Framingham heart study. *Circ Cardiovasc Genet* 2012;5:113–21.
18 Abraham TM, Pencina KM, Pencina MJ, *et al*. Trends in diabetes incidence: the Framingham heart study. *Diabetes Care* 2015;38:482–7.

19 VanderWeele TJ. Causal mediation analysis with survival data. *Epidemiology* 2011;22:582–5.
20 Smith LH, VanderWeele TJ. Mediational E-values: approximate sensitivity analysis for unmeasured Mediator-Outcome confounding. *Epidemiology* 2019;30:835–7.
21 Cho S-H, Huang Y-T. Mediation analysis with causally ordered mediators using COX proportional hazards model. *Stat Med* 2019;38:1566–81.
22 Kershaw KN, Droomers M, Robinson WR, *et al*. Quantifying the contributions of behavioral and biological risk factors to socioeconomic disparities in coronary heart disease incidence: the MORGEN study. *Eur J Epidemiol* 2013;28:807–14.
23 Holmes MV, Lange LA, Palmer T, *et al*. Causal effects of body mass index on cardiometabolic traits and events: a Mendelian randomization analysis. *Am J Hum Genet* 2014;94:198–208.
24 Gepner AD, Piper ME, Johnson HM, *et al*. Effects of smoking and smoking cessation on lipids and lipoproteins: outcomes from a randomized clinical trial. *Am Heart J* 2011;161:145–51.
25 Fritz J, Shiffman D, Melander O, *et al*. Metabolic mediators of the effects of family history and genetic risk score on coronary heart Disease-Findings from the Malmö diet and cancer study. *J Am Heart Assoc* 2017;6:e005254.
26 Tada H, Takamura M, Kawashiri M-A. What is the mechanism of genetic contributions to the development of atherosclerosis? *Atherosclerosis* 2020;307:72–4.
27 Marmot MG, Shipley MJ, Hemingway H, *et al*. Biological and behavioural explanations of social inequalities in coronary heart disease: the Whitehall II study. *Diabetologia* 2008;51:1980–8.
28 Méjean C, Droomers M, van der Schouw YT, *et al*. The contribution of diet and lifestyle to socioeconomic inequalities in cardiovascular morbidity and mortality. *Int J Cardiol* 2013;168:5190–5.
29 Nordahl H, Rod NH, Frederiksen BL, *et al*. Education and risk of coronary heart disease: assessment of mediation by behavioral risk factors using the additive hazards model. *Eur J Epidemiol* 2013;28:149–57.
30 Dégano IR, Marrugat J, Grau M, *et al*. The association between education and cardiovascular disease incidence is mediated by hypertension, diabetes, and body mass index. *Sci Rep* 2017;7:12370.
31 Mega JL, Stitziel NO, Smith JG, *et al*. Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. *Lancet* 2015;385:2264–71.
32 Khera AV, Emdin CA, Drake I, *et al*. Genetic risk, adherence to a healthy lifestyle, and coronary disease. *N Engl J Med* 2016;375:2349–58.