

## Methodology

# Longitudinal effects of lipid indices on incident cardiovascular diseases adjusting for time-varying confounding using marginal structural models: 25 years follow-up of two US cohort studies



Fatemeh Koohi<sup>a,b</sup>, Davood Khalili<sup>c</sup>, Hamid Soori<sup>d</sup>, Maryam Nazemipour<sup>e</sup>,  
 Mohammad Ali Mansournia<sup>f,\*</sup>

<sup>a</sup> Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>b</sup> Obesity Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>c</sup> Prevention of Metabolic Disorders Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>d</sup> Safety Promotion and Injury Prevention Research center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>e</sup> Osteoporosis Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran

<sup>f</sup> Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

## ARTICLE INFO

## Keywords:

Lipids  
 Marginal structural models  
 Inverse probability-of-exposure weighting  
 Time-varying confounding  
 Cardiovascular disease  
 Coronary heart disease  
 Stroke

## ABSTRACT

**Background:** This study assesses the effect of blood lipid indices and lipid ratios on cardiovascular diseases (CVDs) using inverse probability-of-exposure weighted estimation of marginal structural models (MSMs).

**Methods:** A pooled dataset of two US representative cohort studies, including 16736 participants aged 42–84 years with complete information at baseline, was used. The effect of each lipid index, including low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), triglyceride (TG), ratios of TC/HDL-C, LDL-C/HDL-C, and TG/HDL-C on coronary heart disease (CHD) and stroke were estimated using weighted Cox regression.

**Results:** There were 1638 cases of CHD and 1017 cases of stroke during a median follow-up of 17.1 years (interquartile range: 8.5 to 25.7). Compared to optimal levels, the risk of CVD outcomes increased substantially in high levels of TC, LDL-C, TC/HDL-C, and LDL-C/HDL-C. If everyone had always had high levels of TC ( $\geq 240$  mg/dL), risk of CHD would have been 2.15 times higher, and risk of stroke 1.35 times higher than if they had always had optimal levels ( $< 200$  mg/dL). Moreover, if all participants had been kept at very high ( $\geq 190$  mg/dL) levels of LDL-C, risk of CHD would have been 2.62 times higher and risk of stroke would have been 1.92 times higher than if all participants had been kept at optimal levels, respectively. Our results suggest that high levels of HDL-C may be protective for CHD, but not for stroke. There was also no evidence of an adverse effect of high triglyceride levels on stroke.

**Conclusions:** Using MSM, this study highlights the effect of TC and LDL-C on CVD, with a stronger effect on CHD than on stroke. There was no evidence for a protective effect of high levels of HDL-C on stroke. Besides, triglyceride was not found to affect stroke.

## Introduction

Cardiovascular diseases (CVDs), mainly coronary heart disease (CHD) and stroke, remain the leading cause of premature death and a

major cause of disability globally [1]. Previous epidemiological studies have suggested that unfavorable lipid indices levels may play an essential role in the occurrence and progression of cardiovascular events [2]. Most observational studies have investigated the effects of lipid

**Abbreviations:** CVD, Cardiovascular Disease; CHD, Coronary Heart Disease; CKD, Chronic Kidney Disease; LDL-C, Low-Density Lipoprotein Cholesterol; HDL-C, High-Density Lipoprotein Cholesterol; TG, Triglyceride; MSM, Marginal Structural Model; NCEP, The National Cholesterol Education Program; NHLBI, National Heart, Lung, and Blood Institute; ARIC, Atherosclerosis Risk In Communities; MESA, Multi-Ethnic Study of Atherosclerosis; MET, The metabolic equivalent of task; IQR, Interquartile Range; CI, Confidence Interval; HR, Hazard Ratio.

\* Corresponding author.

E-mail address: [mansournia\\_m@sina.tums.ac.ir](mailto:mansournia_m@sina.tums.ac.ir) (M.A. Mansournia).

<https://doi.org/10.1016/j.gloepi.2022.100075>

Received 11 November 2021; Received in revised form 14 May 2022; Accepted 16 May 2022

Available online 23 May 2022

2590-1133/© 2022 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

indices on CVD using the information at baseline only. However, it is essential to consider subsequent changes in lipid profile and other CVD risk factors (e.g., hypertension, diabetes, smoking, drinking, obesity, and lipid-lowering drugs) to estimate a more precise effect [3].

In longitudinal studies of the effect of a time-varying exposure, some time-varying confounders may be affected by the prior exposure, so that they have the dual role of confounder and mediator in the causal pathway. In the presence of these variables, standard statistical models such as the time-dependent Cox regression model, which adjusts for all baseline and time-varying confounders, may provide biased estimates of the effect of exposure [4–7] due to introducing collider-stratification bias [8–10] and over-adjustment bias [11–13]. Time-varying confounding affected by prior exposure is possible in estimating the effects of lipid indices on CVD, e.g., hypertension and body mass index as independent risk factors for CVD may be affected by preceding lipid levels, and also predict the subsequent lipid levels.

Marginal structural models (MSMs), developed by Robins et al. [14,15], are an alternative approach to estimating a time-varying exposure's causal effect, appropriately adjusting for time-varying confounders affected by prior exposure. The parameters of MSMs can be calculated using inverse probability-of-exposure weighting [16–20]. To our knowledge, there has been no published paper on the effects of lipid indices using marginal structural models. Using pooled data from two large prospective cohort studies from the US population, we sought to estimate the effects of lipid indices on CVDs, including CHD and stroke, separately in the presence of time-varying confounders using a marginal structural Cox model.

## Methods

### Study population

The present study was based on pooled data from two large prospective cohort studies sponsored by the National Heart, Lung, and Blood Institute (NHLBI), the ARIC (Atherosclerosis Risk In Communities) study [21] and the MESA study (Multi-Ethnic Study of Atherosclerosis) [22]. The data harmonization methodology of Maelstrom Research was used for pooling data [23]. The design details of each study have been described elsewhere [21,22]. Briefly, the ARIC study [21] is an ongoing prospective cohort study of 15,792 individuals aged 45–64 years recruited from four US communities (Washington County, MD; Forsyth County, NC; Jackson, MS; and suburban Minneapolis, MN). After the baseline examination between 1987 and 1989, participants were re-examined three times at triennial intervals; after approximately eleven years, a subset of participants have completed further visits, the most recent in 2016–17. The MESA study [22] is an ongoing prospective cohort study that includes 6,814 White and Black adults recruited from six US communities (Baltimore City and Baltimore County, MD; Chicago, IL; Forsyth County, NC; Los Angeles County, CA; New York, NY; and St. Paul, MN). Participants were aged 45–84 years and free of clinical CVD at baseline in 2000–2002. After the baseline examination, there have been five additional follow-up visits at biennial intervals, the most recent ongoing in 2016–18. In both studies, all participants (or proxies) were contacted annually by telephone to identify new hospitalizations and medical diagnoses. The studies' websites contain details of all the available data through a fully searchable data dictionary [24,25].

The ARIC and MESA studies were performed following the ethical standards in the Declaration of Helsinki. The Institutional Review Boards approved the ARIC study for each ARIC study affiliated institution, and the MESA study protocol was approved by the Institutional Review Board at each site. All participants provided written informed consent, and all methods were carried out following relevant guidelines and regulations.

The present analyses were restricted to participants without known CVD and with complete data for each lipid component and confounders

at baseline. Due to the large interval between the fourth and fifth visits in the ARIC study, data from the first four visits of both cohorts were used in the current study. The flowchart of the participant selection process and sample size for each lipid component is shown in Fig. 1.

### Exposures

The primary exposures of interest in the current analysis were serum lipid concentrations, including TC, LDL-C, HDL-C, TG, TC/HDL-C, LDL-C/HDL-C, and TG/HDL-C ratios. Clinical data collection for the ARIC and MESA studies has been reported on their websites (available at [https://sites.cscs.unc.edu/aric/desc\\_pub](https://sites.cscs.unc.edu/aric/desc_pub) and <https://www.mesa-nhlbi.org/>). Briefly, blood samples were collected after 12-h fasting using a standardized venipuncture procedure at each visit in both studies. EDTA plasma samples were aliquoted on ice and stored at  $-70^{\circ}\text{C}$  until analysis. Total cholesterol and triglyceride were measured using standard enzymatic processes (Roche Diagnostics). The HDL-C level was measured using the cholesterol oxidase method after precipitation of non-HDL-C with magnesium/dextran. The LDL-C was calculated using the Friedewald formula from the total cholesterol, HDL-C, and triglyceride values  $<400$  mg/dL.

The National Cholesterol Education Program (NCEP) guidelines were used to define categories of all lipid components at baseline and follow-up visits [26]. Each TC measurement was categorized into three groups: optimal ( $<200$  mg/dL), borderline high (200–239 mg/dL), and high ( $\geq 240$  mg/dL). LDL-C measurements were categorized into five groups: optimal ( $<100$  mg/dL), near-optimal (100–129 mg/dL), borderline high (130–159 mg/dL), high (160–189 mg/dL), and very high ( $\geq 190$  mg/dL). HDL-C measurements were categorized into four groups: low ( $<40$  mg/dL), optimal (40–59 mg/dL), high (60–79 mg/dL), and very high ( $\geq 80$  mg/dL). Triglyceride values were categorized into three groups: optimal ( $<150$  mg/dL), borderline high (150–199 mg/dL), and high ( $\geq 200$ –499 mg/dL). All lipid categories were determined by NCEP 2001 guidelines [27]; the very high category of HDL-C was determined based on previous studies suggesting a possible increased risk for CHD and total mortality events in participants with HDL-C  $>80$  mg/dL [28,29]. Finally, quartiles of each lipid ratio were used to define their categories at each visit.

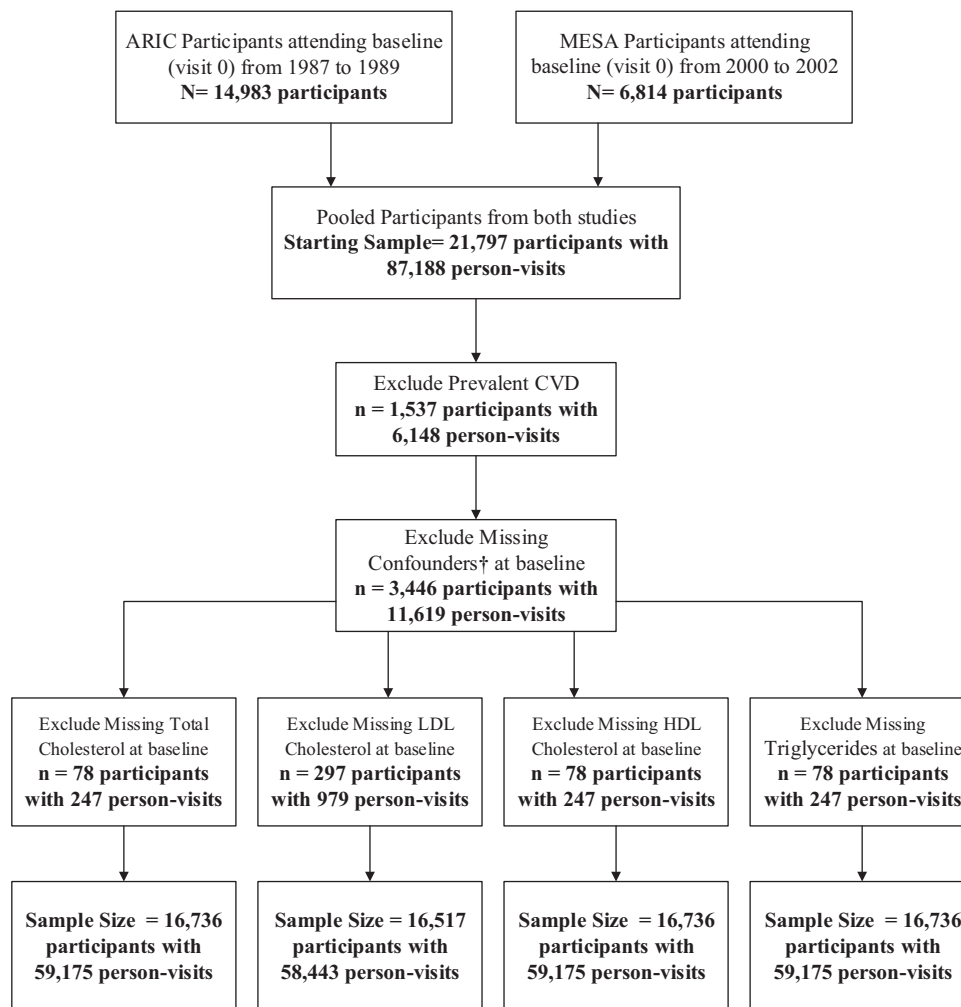
### Outcomes

The primary outcomes of interest for our analysis were incident CHD, defined as myocardial infarction or CHD death, and stroke, defined as definite or probable hospitalized ischemic or hemorrhagic stroke events. Diagnosis of myocardial infarction generally required at least two of the following criteria: chest pain, electrocardiographic abnormalities consistent with myocardial infarction, and elevated cardiac biomarkers. Diagnosis of stroke generally required a persistent central neurologic deficit lasting  $>24$  hours. Events were ascertained by following each participant at 9–12 months intervals using telephone calls. All events were adjudicated from medical records and death certificates for endpoint classification and assignment of incidence dates by a review committee in both studies. In this study, outcome events confirmed by 31 December 2014 were included.

### Confounders

Details of interview and data collection procedures for each study have been described on their websites. Briefly, both studies used standardized protocols and similar standard and validated methods to measure demographic characteristics and CVD risk factors at each study visit.

The potential confounders available only at baseline (visit 0) included age, sex, race (white, African American, Hispanic, and Asian), education (less than high school, high school, some college, and college +), family income (less than \$16000, \$16000–24999, \$25000–34999,



† **Confounders:** age, sex, education, race, family income, health insurance, family history of CVD, physical activity, calorie intake, waist/hip ratio, hypertension, diabetes, use of antihypertension drugs, use of lipid-lowering drug, smoking status, drinking status.

Fig. 1. Flow chart showing the derivation of final analysis samples from 21,797 participants in ARIC and MESA studies.

\$35000–49999, more than \$50000), family history of CVD, health insurance, calorie intake (Kcal), and physical activity (the metabolic equivalent of task (MET)-minute/week).

The time-varying confounders, measured at all visits including baseline, were lipid-lowering drugs use, hypertension (systolic blood pressure  $\geq 90$  or diastolic blood pressure  $\geq 140$  mmHg or taking any medication for high blood pressure) [30], antihypertensive drugs use, diabetes mellitus (fasting blood glucose  $\geq 126$  mg/dl or taking any medication for diabetes) [31], body mass index (categorized into three groups:  $<25$ ,  $\geq 25$  to  $<30$ ,  $\geq 30$  kg/m<sup>2</sup>) [32], waist-to-hip ratio ( $\geq 0.9$  in men and  $\geq 0.85$  in women) [33], smoking status (defined as never, former, and current smoking cigarettes) and drinking status (defined as never, former, and current drinking alcoholic beverages). Baseline and time-varying confounders included HDL-C and triglyceride for the exposure TC and LDL-C; triglyceride and LDL-C for HDL-C; LDL-C and HDL-C for triglyceride; and LDL-C for triglyceride/HDL-c ratio as well. The confounders mentioned above were selected based on previous literature [5,34–41].

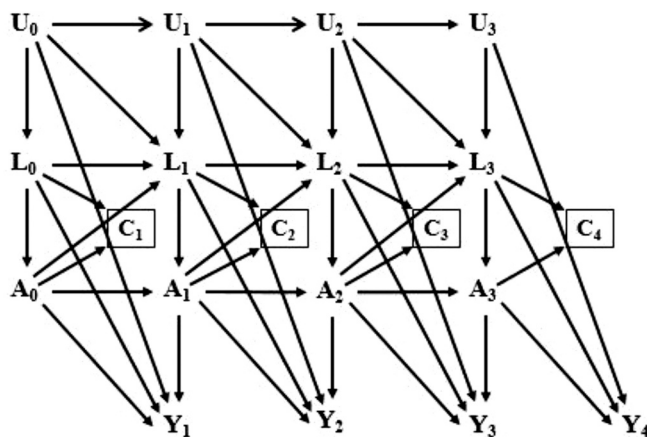
Causal diagram

The causal directed acyclic graph [42–50] in Fig. 2 represents the

assumed causal structure for the effects of each lipid index on CHD and stroke in the study population. Using the notation of Robins et al. [14],  $A(k)$  and  $Y(k+1)$  correspond to a lipid component at visit  $k$  ( $k= 0, 1, 2, 3$ ) and CHD or stroke during the period (visit  $k$ , visit  $k+1$ ), respectively.  $L(k)$  denotes a vector of measured time-varying confounders at visit  $k$ .  $L(0)$  includes the confounders available only at baseline as well. We use overbars to represent a covariate history, for example  $\bar{A}(k)$ ,  $\bar{C}(k)$ , and  $\bar{L}(k)$  are a subject’s observed histories of exposure, censoring, and measured confounders through visit  $k$ , respectively.  $U(k)$  denotes unmeasured causal risk factors at visit  $k$  for  $Y(k+1)$ , and  $C(k+1)$  denotes censoring during the interval (visit  $k$ , visit  $k+1$ ) (1 if censored and 0 otherwise). The square around  $C(k+1)$  indicates that our analyses are restricted to individuals who have not yet been censored. Fig. 2 represents the untestable assumption of no unmeasured confounders for exposure and censoring histories, given data on  $L(k)$ , denoted by the absence of arrows from  $U(k)$  to  $A(k)$  and  $C(k+1)$ . Note that for the simplicity of presentation, the arrows from  $L(0)$  to  $A(1)$ , from  $A(0)$  to  $L(2)$ , etc., were omitted.

Statistical methods

To estimate the effect of lipid indices on CHD and stroke, we fit



**Fig. 2.** Causal diagram representing the assumed causal structure for the effects of lipid indices (A(k)) on CHD or stroke (Y+1) in participants of ARIC and MESA studies across four visits (i.e., k = 0, 1, 2, 3). L(k) includes a vector of measured time-varying confounders at visit k, and L(0) consists of the baseline confounders and the baseline values of time-varying confounders. U(k) represents all unmeasured causal risk factors for Y(k+1) at visit k. C(k+1) denotes censoring during the period (visit k, visit k+1). The square around C(k+1) indicates that our analyses are limited to uncensored subjects.

separate marginal structural Cox models (Cox MSM) for each exposure and outcome (14 models). The parameters of these models were estimated using inverse probability-of-exposure-and-censoring weighting.

For our analyses, we specified the following Cox MSM:

$$\lambda_{T_2}(t|V) = \lambda_{T_0}(t) \exp\left(\sum_{i=1}^{J-1} \beta_i a_i(t) + \beta'_i V\right)$$

where  $\lambda_{T_2}(t|V)$  denotes the counterfactual hazard of CHD or stroke at time t among subjects with the column vector of baseline covariates V (L(0)), if, possibly contrary to fact, they had followed the exposure history  $\bar{a}$ ;  $a_1(t), a_2(t), \dots, a_{J-1}(t)$  are indicators for levels (J) of each lipid component at visit t, and  $\beta_i$  is a column vector of the parameters. Assuming the correct specification of our marginal structure model, the parameters  $\beta_1, \dots, \beta_{J-1}$  have causal interpretations: they represent the causal log-hazard ratio for CHD or stroke, comparing always had an unfavorable level (levels 2, 3, ..., J) of a lipid component with their hazard of CHD or stroke if they always had an optimal (baseline) level.

In real longitudinal data analysis and the absence of censoring, each subject receives only one value of  $\bar{a}(k)$ , the observed exposure history  $\bar{A}(k)$ . Only the actual corresponding outcome Y(k+1) during the period (visit k, visit k+1) is observed. However, one can use inverse probability-of-exposure-and-censoring weighting to consistently estimate MSM parameters under the following four assumptions, most of which are untestable [44]: (1) conditional exchangeability assumes that no unmeasured confounders to adjust for both confounding and selection bias due to loss to follow-up; (2) positivity states that for any observed combination of values of the confounders, every level of exposure can and does occur; (3) consistency states that a subject's counterfactual outcome at visit t under her/his observed value of exposure is equal to her/his observed outcome, which generally requires either a well-defined intervention or an assumption that any intervention used to achieve the exposure change has an equivalent effect; (4) correct specification of models used to estimate weights of exposure and censoring. They are discussed in more detail in the discussion. The stabilized inverse probability-of-exposure weights at visit t ( $SW_{E_i}(t)$ ) are as follows:

$$SW_{E_i}(t) = \prod_{k=0}^t \frac{Pr[A(k) = a_i(k) | \bar{A}(k-1) = \bar{a}_i(k-1), V = v_i, \bar{C}(k) = \bar{0}]}{Pr[A(k) = a_i(k) | \bar{A}(k-1) = \bar{a}_i(k-1), \bar{L}(k) = \bar{l}_i(k), \bar{C}(k) = \bar{0}]}$$

Only the set of baseline confounders V was included in the numerator model, while both concurrent time-varying and baseline confounders were included in the denominator model. Note that V is included in the denominator because L(0) is part of  $\bar{L}(k)$ . We estimated the probabilities in the denominator through pooled multinomial logistic regression models for exposure that include exposure at the previous visit, concurrent time-varying and baseline confounders, and visit. The probabilities in the denominator were estimated using the same regression model but without the time-varying confounders.

The stabilized inverse probability-of-censoring weights at visit t ( $SW_{C_i}(t)$ ) are as follows:

$$SW_{C_i}(t) = \prod_{k=0}^t \frac{Pr[C(k+1) = 0 | \bar{C}(k) = \bar{0}, \bar{A}(k) = \bar{a}_i(k), V = v_i]}{Pr[C(k+1) = 0 | \bar{C}(k) = \bar{0}, \bar{A}(k) = \bar{a}_i(k), \bar{L}(k) = \bar{l}_i(k)]}$$

The probabilities in the numerator and denominator were calculated using pooled multinomial logistic regression models for censoring with the same predictors as exposure models.

To estimate the parameters of Cox MSM, we fit the following pooled logistic regression model with weights  $SW_{T_1}(t) = SW_{E_i}(t) \times SW_{C_i}(t)$ :

$$\text{Logit } Pr[D(t) = 1 | D(t-1) = 0, \bar{A}(t-1), V] = \alpha_0(t) + \sum_{i=1}^{J-1} \alpha_i A_i(t-1) + \alpha_i V$$

Under the identifiability assumptions mentioned above and considering that the risk of an event is less than 10% per person-visit interval [51], the weighted estimates of  $\alpha_1, \alpha_2, \dots, \alpha_{J-1}$  would be approximately consistent for the causal parameters  $\beta_1, \beta_2, \dots, \beta_{J-1}$  of our Cox MSM.

As a sensitivity analysis, we also used lagged time-varying confounders (confounders measured in the previous visit) instead of concurrent confounders to estimate inverse probability-of-exposure weights [52]. Since using weights induces within-subject correlation, we applied cluster-robust standard errors to obtain 95% confidence intervals (CIs) [53–56].

All analyses were performed using Stata Version 14.0 (Stata Corp; College Station, TX) [57] and R statistical software version 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria).

## Results

### Sample characteristics

In total, 16,736 participants aged 42–84 years who had complete data at baseline (visit 0) and with no history of any cardiovascular diseases were included in the study. The main reasons for no inclusion were having a history of cardiovascular diseases and a lack of data on the confounders at baseline (Fig. 1). Baseline and time-varying characteristics of the participants have been summarized in Table 1. At baseline, the mean (SD) age was 56.7 (8.6) years, 54.5% were women, and 66.2% were white. Also, 61.3% had a family history of CVD, 75.4% were centrally obese, and 58.1% were current drinkers. The distribution of exposures and time-varying confounders changed over visits, e.g., the proportion of subjects with hypertension, diabetes, taking antihypertensive drugs, taking lipid-lowering medications, and obesity increased from baseline (visit 0) to visit 3, whereas the proportion of current smoking, current drinking, high levels of TC and LDL-C decreased.

Baseline characteristics of participants who remained in the study and those censored during follow-up are shown in Supplementary Table 1. In summary, for both events of CHD and stroke, the likelihood of being dropped out due to competing causes of death was more in men, blacks, people with an education of less than high school, family history of CVD, hypertension, diabetics, obese people, those taking lipid-lowering medication, people with a high waist-to-hip ratio, current smokers, former alcohol drinkers, and people with high levels of lipids, including HDL-C than others. Also, women, Hispanics, those with higher education, non-smokers, former alcohol drinkers, and people with

**Table 1**  
Baseline and time-varying characteristics of participants across baseline and follow-up visits.

Characteristics <sup>a</sup>	During follow-up			
	At baseline Visit 0 (n = 16,736)	Visit 1 (n = 15,837)	Visit 2 (n = 14,882)	Visit 3 (n = 13,876)
Sex, women (%)	9119 (54.5)	8667 (54.7)	8210 (55.2)	7705 (55.5)
Age, years	56.7 (8.4)	59.1 (8.1)	61.5 (7.8)	64.0 (7.7)
Race				
White/Caucasian	11075 (66.2)	10595 (66.9)	10001 (67.2)	9360 (67.4)
African American/Black	3682 (22.0)	3382 (21.3)	3112 (20.9)	2826 (20.4)
Hispanic/Latino	1250 (7.5)	1168 (7.4)	1097 (7.4)	1048 (7.6)
Asian/Chinese	729 (4.5)	692 (4.4)	672 (4.5)	642 (4.6)
Education				
Less than high school	3260 (19.5)	–	–	–
High school	4697 (28.1)	–	–	–
Some college	5543 (33.1)	–	–	–
College +	3236 (19.3)	–	–	–
Family income (per year)				
Less than \$16,000	3204 (19.1)	–	–	–
\$16,000–24,999	2300 (13.7)	–	–	–
\$25,000–34,999	2743 (16.4)	–	–	–
\$35,000–49,999	3172 (19.0)	–	–	–
More than \$50,000	5317 (31.8)	–	–	–
Health insurance	10657 (63.7)	–	–	–
Family history of CVD	10264 (61.3)	–	–	–
Physical activity (MET-minutes/week)	953.7 (1561.9)	–	–	–
Total energy intake (Kcal)	1601.7 (734.1)	–	–	–
Systolic blood pressure	122.3 (19.5)	121.9 (19.3)	123.5 (19.4)	125.5 (19.6)
Diastolic blood pressure	72.7 (10.7)	71.4 (10.2)	71.1 (10.3)	70.5 (10.2)
Hypertension	5817 (34.8)	5761 (36.4)	6096 (41.0)	6459 (46.6)
Antihypertensive drug	4794 (28.6)	5083 (32.1)	5483 (36.8)	5813 (46.6)
Lipid-lowering drug	3315 (19.8)	3834 (24.2)	4408 (29.6)	4774 (34.4)
Diabetes mellitus	1795 (10.7)	2109 (13.3)	2065 (13.9)	2106 (15.2)
Body mass index, kg/m <sup>2</sup>				
Normal	5526 (33.0)	5000 (31.6)	4353 (29.3)	3861 (27.8)
Overweight	6677 (39.9)	6367 (40.2)	5909 (39.7)	5509 (39.7)
Obese	4530 (27.1)	4468 (28.2)	4620 (31.0)	4506 (32.5)
Waist/hip ratio <sup>b</sup>				
Normal	4133 (24.7)	3878 (24.5)	2854 (19.2)	2338 (16.9)
Abdominal obesity, >0.90 for men and >0.85 for women	12603 (75.3)	11959 (75.5)	12028 (80.8)	11538 (83.2)
Smoking status				
Never	7613 (45.5)	6803 (43.0)	6481 (43.5)	6075 (43.8)
Former	5604 (33.5)	6146 (38.8)	6125 (41.2)	5982 (43.1)
Current	3519 (21.0)	2888 (18.2)	2276 (15.3)	1819 (13.1)

**Table 1 (continued)**

Characteristics <sup>a</sup>	At baseline	During follow-up		
	Visit 0 (n = 16,736)	Visit 1 (n = 15,837)	Visit 2 (n = 14,882)	Visit 3 (n = 13,876)
Drinking status				
Never	3781 (22.6)	3284 (20.7)	3267 (22.0)	2683 (19.3)
Former	3239 (19.4)	3621 (22.9)	3711 (24.9)	4268 (30.8)
Current	9716 (58.0)	8932 (56.4)	7904 (53.1)	6925 (49.9)
Total cholesterol, mg/dL				
Optimal	7533 (45.0)	7705 (48.6)	7465 (50.2)	7555 (54.5)
Borderline high	5937 (35.5)	5587 (35.3)	5184 (34.8)	4610 (33.2)
High	3266 (19.5)	2545 (16.1)	2233 (15.0)	1711 (12.3)
LDL cholesterol, mg/dL				
Optimal	3386 (20.5)	3610 (22.8)	3897 (26.2)	3980 (28.7)
Near optimal	5281 (32.0)	5301 (33.5)	5131 (34.5)	4962 (35.8)
Borderline high	4514 (27.3)	4239 (26.8)	3886 (26.1)	3374 (24.3)
High	2273 (13.8)	1890 (11.9)	1445 (9.7)	1176 (8.5)
Very high	1063 (6.4)	730 (4.6)	494 (3.3)	368 (2.7)
HDL cholesterol, mg/dL				
Low	4106 (24.5)	4144 (26.2)	3450 (23.2)	3410 (24.6)
Optimal	8307 (49.6)	7794 (49.2)	7318 (49.2)	6832 (49.2)
High	3315 (19.8)	3026 (19.1)	3066 (20.6)	2840 (20.5)
Very High	1005 (6.1)	873 (5.5)	1048 (7.0)	794 (5.7)
Triglycerides, mg/dL				
Optimal	12027 (71.9)	11157 (70.5)	10141 (68.2)	9529 (68.7)
Borderline high	2420 (14.4)	2363 (14.9)	2413 (16.2)	2171 (15.7)
High	2289 (13.7)	2317 (14.6)	2328 (15.6)	2176 (15.6)
Total cholesterol / HDL ratio				
Quartile 1, (<3.27)	4177 (25.0)	4030 (25.5)	4278 (28.7)	3953 (28.5)
Quartile 2, (3.27–4.14)	4170 (24.9)	3869 (24.4)	3871 (26.0)	3690 (26.6)
Quartile 3, (4.15–5.24)	4169 (24.9)	3882 (24.5)	3613 (24.3)	3392 (24.5)
Quartile 4, (>5.24)	4217 (25.2)	4057 (25.6)	3120 (21.0)	2841 (20.4)
LDL cholesterol / HDL ratio				
Quartile 1, (<1.91)	4138 (25.1)	4090 (25.8)	4457 (30.1)	4191 (30.2)
Quartile 2, (1.91–2.62)	4101 (24.8)	3936 (24.9)	3932 (26.4)	3723 (26.8)
Quartile 3, (2.62–3.47)	4111 (24.9)	3760 (23.7)	3513 (23.6)	3339 (24.2)
Quartile 4, (>3.47)	4167 (25.2)	3984 (25.2)	2951 (19.9)	2607 (18.8)
Triglycerides / HDL ratio				
Quartile 1, (<1.39)	4184 (25.0)	3544 (22.3)	3450 (23.2)	3105 (22.4)
Quartile 2, (1.39–2.21)	4132 (24.7)	3795 (24.0)	3485 (23.4)	3208 (23.1)
Quartile 3, (2.21–3.67)	4205 (25.1)	4113 (26.0)	3911 (26.3)	3694 (26.6)
Quartile 4, (>3.67)	4212 (25.2)	4385 (27.7)	4036 (27.1)	3869 (27.9)

HDL indicates high-density lipoprotein; LDL, low-density lipoprotein; CVD, cardiovascular disease.

<sup>a</sup> Numbers represent mean (SD) for continuous variables; numbers and percentages correspond to "Yes" for dichotomous variables.

<sup>b</sup> Defined based on World Health Organization cut-off points.

normal total cholesterol and LDL-C levels were more likely to be censored due to loss of follow-up.

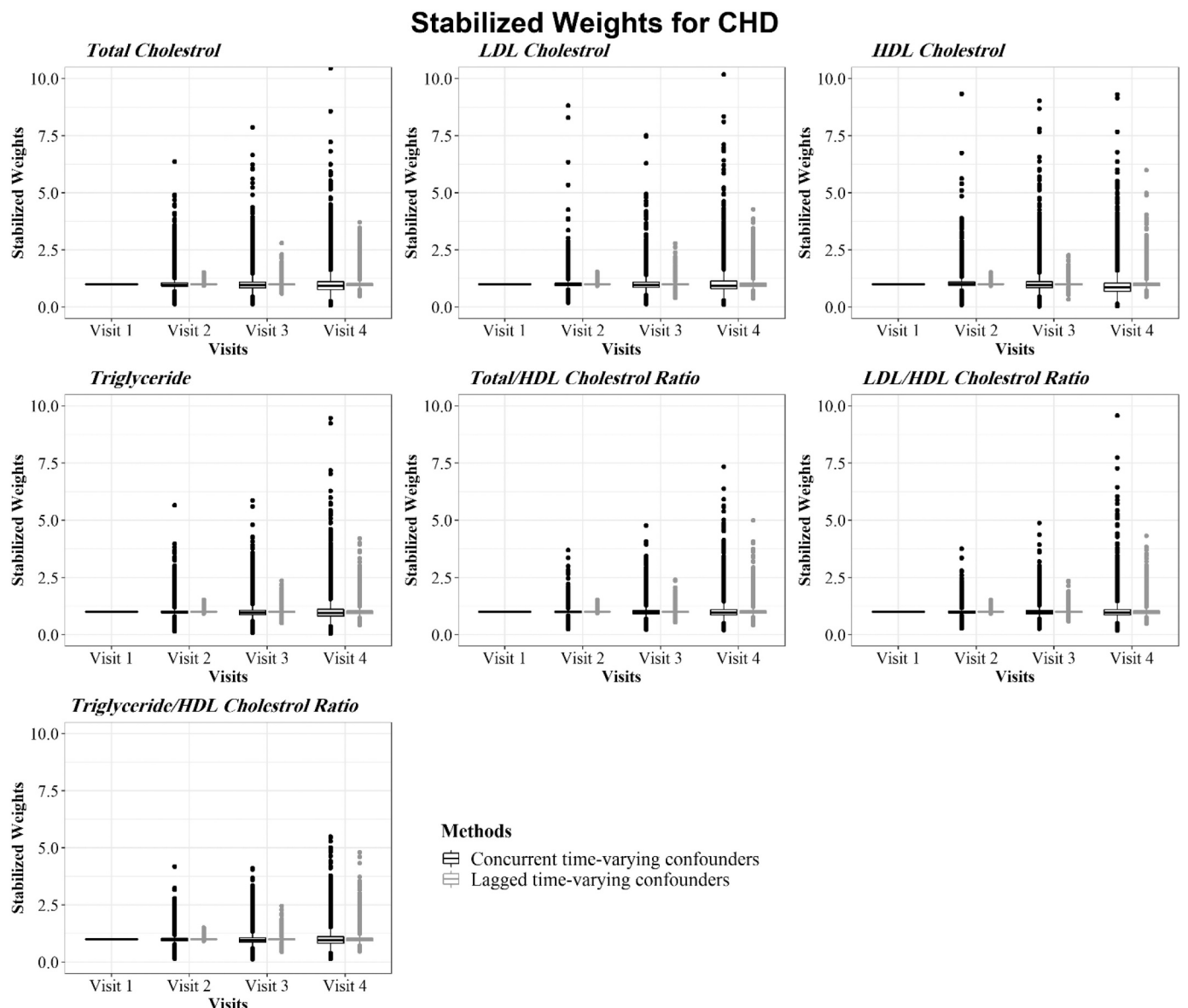
**Marginal structural cox model**

During a median (IQR) follow-up time of 17.1 (8.5 to 25.7) years, there were 1638 (9.8%) incident CHD and 5362 (32.0%) censored; 3616 (21.6%) subjects died from causes other than CHD (competing risk), and 1746 (10.4%) subjects were lost to follow-up. The incidence rate of CHD was 58.6 per 10,000 person-years (95% CI: 55.8–61.5). In addition, during a median (IQR) follow-up time of 18.1 (8.5 to 25.8) years, 1017 (6.1%) new cases of stroke occurred, and 5956 (35.6%) censored; 4038 (24.1%) subjects died from causes other than stroke (competing risk), and 1918 (11.5%) subjects were lost to follow-up. The incidence rate of stroke was 35.8 per 10,000 person-years (95% CI: 33.7–38.1). Among 16736 included subjects, 9736 (58.2%) and 9763 (58.3%), respectively, survived free of CHD and stroke to the end of follow-up.

Model outputs for estimating the final stabilized weights of each lipid

component at all visits for CHD and stroke are presented in Figs 3 and 4, respectively. The stabilized IPTWs used in the Cox MSMs had symmetric distributions centered around the ideal mean of 1 for all lipid indices at all visits, satisfying a necessary condition for correct model specification [53].

Fig. 5 presents MSM estimates of the hazard ratios (HRs) with 95% CIs for the effect of each lipid component on CHD and stroke risk. Based on the results of MSM adjusted for concurrent time-varying confounders, if all participants had been kept at a borderline high (200–239 mg/dL) and high ( $\geq 240$  mg/dL) levels of TC, risk of CHD would have been 1.46 (95% CI, 1.28–1.66), or 2.15 times higher (95% CI, 1.82–2.55), and risk of stroke would have been 1.26 (95% CI, 1.08–1.48), or 1.35 (95% CI, 1.08–1.70) times higher than under a scenario in which all participants had been kept at the optimal levels, respectively. Also, if everyone had always had high (160–189 mg/dL) or very high ( $\geq 190$  mg/dL) levels of LDL-C, risk of CHD would have been 1.85 times (95% CI, 1.51–2.28) and 2.62 times (95% CI, 1.94–3.52) higher, and risk of stroke would have been 1.39 times (95% CI, 1.03–1.85) and 1.92 times (95% CI, 1.21–3.04) higher than if they had always had the optimal levels ( $< 100$  mg/dL), respectively.



**Fig. 3.** Distribution of stabilized weights used in the MSM models for CHD to adjust for concurrent and lagged time-varying confounders.

### Stabilized Weights for Stroke

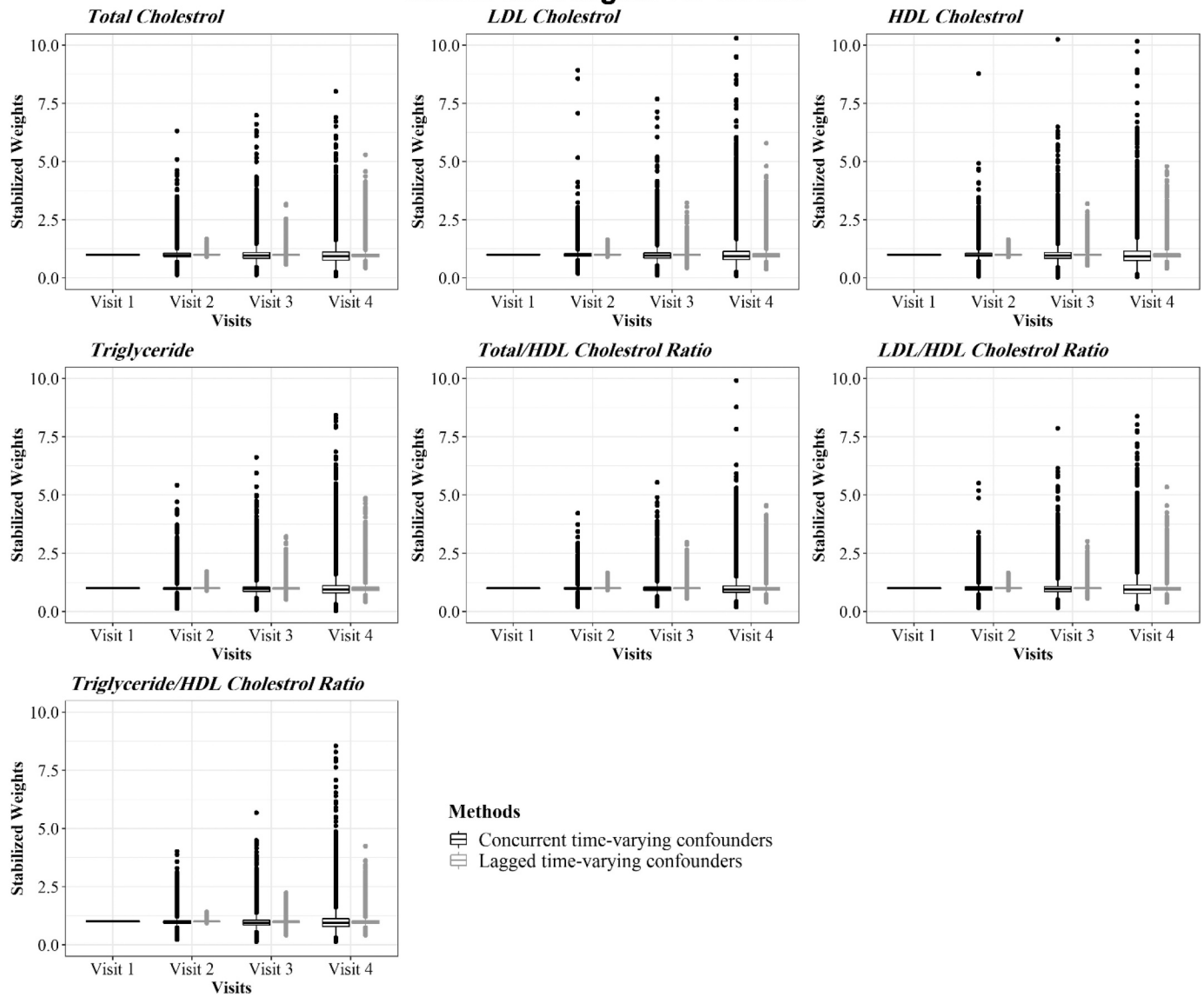


Fig. 4. Distribution of stabilized weights used in the MSM models for stroke to adjust for concurrent and lagged time-varying confounders.

Compared with a scenario in which all participants had been kept at the optimal levels of HDL-C (40–59 mg/dL), the risk of CHD and stroke would have increased by 41% (HR, 1.41; 95% CI, 1.32–1.68), and 29% (HR, 1.29; 95% CI, 1.06–1.58) if all participants had been kept at a low HDL-C level (<40 mg/dL), respectively. If everyone had always had high (60–79 mg/dL) or very high ( $\geq 80$  mg/dL) levels of HDL-C compared with if they had always had the the optimal levels (40–59 mg/dL), the risk of CHD would have decreased by 19% (HR, 0.81; 95% CI, 0.66–0.99) and 33% (HR, 0.67; 95% CI, 0.44–1.01), respectively, while the risk of stroke would have increased by 6% (HR, 1.06; 95% CI, 0.84–1.33) and 18% (HR, 1.18; 95% CI, 0.79–1.76). There was also little evidence of an effect of high levels of triglyceride on stroke (Fig. 5).

If all participants had been kept at the third (4.2–5.2) and fourth ( $\geq 5.2$ ) quintiles of TC/HDL-C ( $\leq 3.3$ ), risk of CHD would have been 1.52 (95% CI, 1.26–1.84), or 2.37 times higher (95% CI, 1.97–2.85), and risk of stroke would have been 1.45 (95% CI, 1.16–1.83), or 1.58 (95% CI, 1.25–1.99) times higher than under a scenario in which all participants had been kept at the first quintile, respectively. Besides, if everyone had been kept at the third (2.6–3.5) and fourth ( $\geq 3.5$ ) quintiles of LDL-C/HDL-C, risk of CHD would have been 1.40 times (95% CI, 1.15–1.70)

and 2.16 times (95% CI, 1.78–2.61) higher, and risk of stroke would have been 1.36 times (95% CI, 1.06–1.73) and 1.68 times (95% CI, 1.32–2.15) higher than if they had been kept at the first quintile ( $\leq 1.9$ ), respectively. Moreover, if all participants had been kept at the third and fourth quartiles of triglyceride/HDL-C, the risk of CHD would have been higher than if all participants had been kept in the first quartile. At the same time, no association was found between triglyceride/HDL-C and stroke.

Results from sensitivity analyses adjusted for lagged time-varying confounders produced similar findings as our primary analyses, except for borderline high and high levels of triglyceride that demonstrated to be associated with incident CHD. The estimates for borderline high and high levels of triglyceride in the lagged analysis of CHD were slightly higher: hazard ratios were 1.21 (95% CI, 1.02–1.43) and 1.18 (95% CI, 1.00–1.39), respectively (Fig. 5).

### Discussion

This large, pooled cohort study investigated the effect of lipid indices and lipid ratios on CHD and stroke using a Cox MSM. It was revealed that

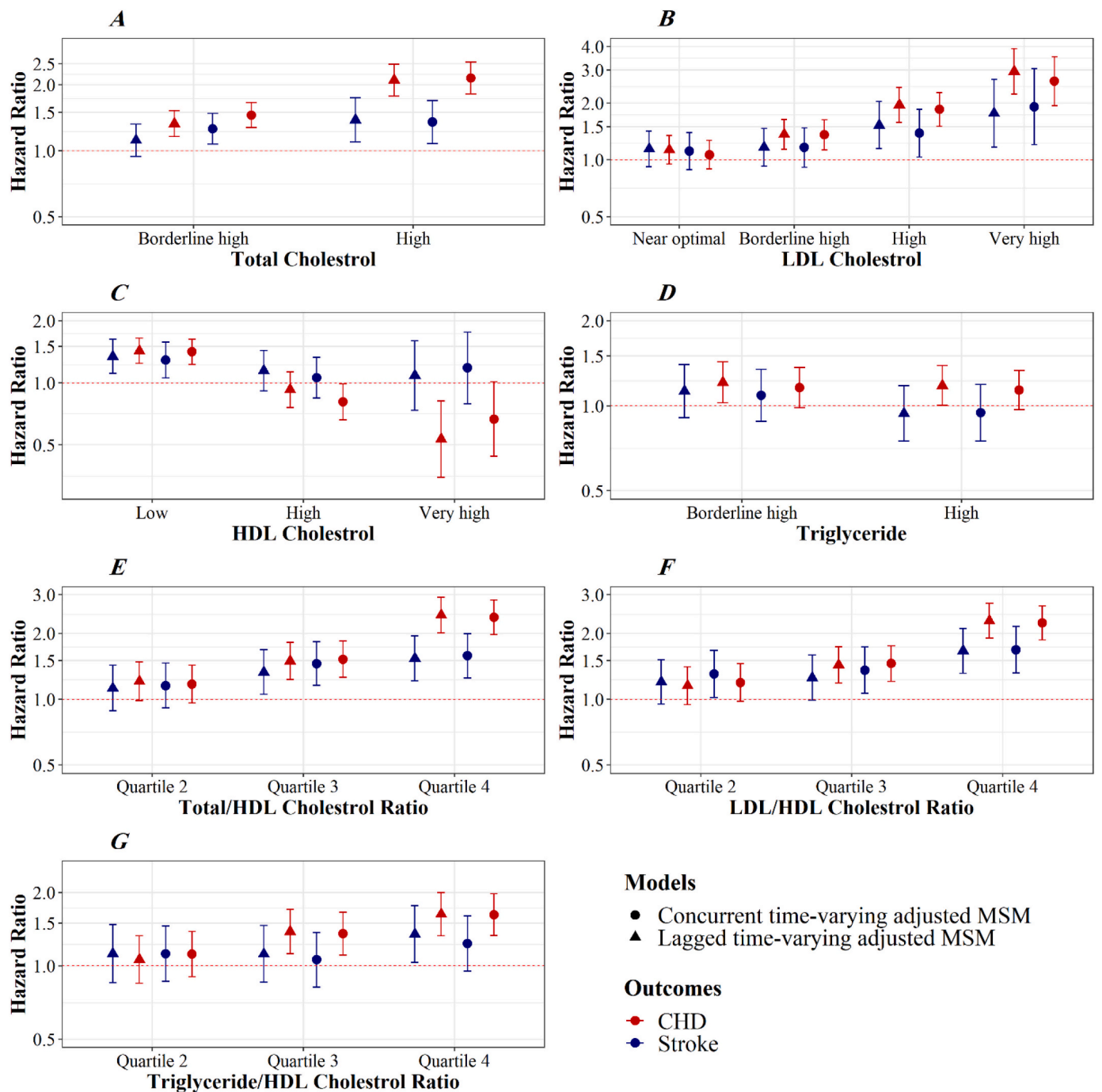


Fig. 5. Hazard ratios and 95% confidence intervals for the effect of each lipid component on the risk of CHD (red) and stroke (blue). Estimates from two MSMs with joint IPTW and IPCW adjusted for concurrent and lagged time-varying confounders. HRs approximated from a pooled logistic regression.

TC and LDL-C play roles in developing CHD and stroke with a weaker effect on stroke than CHD. We also observed no evidence for the protective effect of high levels of HDL-C on stroke. Besides, triglyceride was not found to affect stroke.

Although the association between lipids and CVD has been confirmed in numerous studies, most of them have used standard statistical methods, adjusting only for the baseline confounders [58–60]. We used the Cox marginal structural model, a causal inference model, because standard statistical methods may give biased effect estimates when time-varying confounders, such as lipid-lowering drugs and hypertension, are affected by previous exposure [4,14]. The inverse probability-of-exposure-and-censoring weighted estimators of MSMs

can provide unbiased estimates of the causal effect of a time-varying exposure in this setting by appropriately adjusting for time-varying confounding and selection bias due to censoring [9,53]. To date, only one previous study has examined the effect of LDL-C and HDL-C on CVD and mortality using G-estimation [35]. Besides, since the underlying causes of CHD and stroke may be different, a better understanding of the association of lipids with subtype-specific CVD events rather than focusing on total CVD would be more enlightening for CVD prevention. Therefore, our study is the first that examined the effect of lipid indices and lipid ratios on the risk of CHD and stroke, separately using a method that adjusts for time-varying confounding.

Both MSMs adjusting for concurrent and lagged time-varying



confounders suggested that risk of CHD and stroke would have been higher if everyone had had higher than optimal levels than if everyone had had optimal levels. The risk was higher for CHD than stroke. Our findings are consistent with the evidence obtained from some genetic studies and randomized trials suggesting the effect of LDL-C on both CHD and stroke [61–63]. In a Mendelian randomization study using genetic variants that confer lifelong LDL-C differences, the causal effect of LDL-C was more substantial for CHD than stroke: one mmol/L genetically higher LDL-C increased the risk of CHD by 50%, ischemic stroke by 12%, and cardioembolic stroke by 6% [61]. More potent effects of LDL-C on CHD than on stroke have also been shown in other observational studies [64].

In contrast, some randomized trials of LDL-C-lowering therapies, such as statin therapy, have shown the consistent effects of LDL-C-lowering treatments on CHD and stroke [62,65]. Tilling et al. also found a positive linear relationship for LDL-C with CHD using G-estimation [35]. Previous studies estimated the 18-years risks of stroke and MI under different scenarios of hypothetical interventions using the parametric g-formula showed that lowering total cholesterol to <6.22 mmol/l did not substantially alter the stroke risk [66], while it was associated with reduced MI risk, with a population risk ratio 0.87 [67]. Another study using the parametric g-formula found that long-term exposure to moderate levels of LDL-cholesterol has the same impact on CHD risk as shorter exposures to levels considered 'high' [68].

In line with Tilling et al. [35] showing a linear inverse relationship between HDL-C and CHD, using G-estimation, high levels of HDL-C in our study tended to have a protective effect on CHD using MSM adjusting for concurrent confounders. However, both MSMs found little evidence for the protective effect of HDL-C on stroke after adjusting for LDL-C and TG. Recent epidemiologic, genetic, and clinical trial studies have suggested that very high levels of HDL-C are associated with cardiovascular events and mortality [29,69–71]. Clinical studies have indicated that not only serum HDL-C concentration/quantity but also its qualities/properties can play a critical role in determining its overall effects on clinical outcomes [72,73]. The results obtained from Mendelian randomization studies have also shown that genetic mutations leading to very high HDL-C may play a role in this association and adversely affect CVD [74–76]. The discrepancy among different studies regarding the cardiovascular effects of high HDL-C may be explained by the very small sample size of individuals with very high HDL-C (i.e., >80 mg/dL) in most of the older studies, which limits the ability to conclude the risk of extremely high levels of HDL-C on CVD [77]. Furthermore, in many of these studies, the focus was on low concentrations of HDL-C, thereby failing to reveal associations at higher concentrations [78]. Our longitudinal study design considered not only variations in HDL-C itself over time but also changes in cardiovascular risk factors over time to estimate the risk of outcomes. In addition, a sufficient sample size of participants with > 80 mg/dl of HDL-C facilitated estimating differential effects of very high HDL-C levels on various CVD manifestations within a single population. Although some other recent studies had these strengths, our study can add evidence to the association.

Moreover, no strong evidence for the detrimental effect of TG on stroke was found in both MSMs after adjustment for LDL-C and HDL-C, even though the results suggest a possible minor effect on CHD. Although the exact role of TG in inducing CVD remains controversial, our results are consistent with those obtained from the largest Emerging Risk Factors Collaboration [79]. Genetic evidence linking TG levels and atherosclerosis risk revealed that almost all genetic variants associated with TG are linked with at least one other lipid trait, commonly reduced HDL-C [80–82]. A genetic study that applied inverse-variance-weighted Mendelian randomization to estimate lipid indices' causal effect on stroke demonstrated genetically elevated triglycerides are not associated with ischemic stroke or its subtypes [83]. On the other hand, randomized trials of agents lowering TG did not find any reduction in cardiovascular events [84,85]. In contrast, a recent systematic review

and meta-regression analysis of randomized trials of TG-lowering therapies showed an association between TG-lowering and CVD risk reduction, even after adjusting for LDL-C lowering. It should be noted that the effect attenuated after excluding the Reduction of Cardiovascular Events with Icosapent Ethyl-Intervention Trial (REDUCE-IT) from the analysis [86].

Furthermore, both MSMs showed that a TC/HDL-C ratio higher than four and an LDL-C/HDL-C higher than 2.5 is associated with the greater risks of CHD and stroke. A TG/HDL-C higher than two was associated with CHD; however, there was little evidence for its association with stroke. Our findings are consistent with some previous observational studies [87–90].

In cohort studies, since the exposure may change before current time-varying confounders, lagged time-varying confounders were adjusted for sensitivity analysis [52,54]. However, effect estimates in the MSM model with weights adjusted for concurrent confounders were weaker than those adjusting for lagged confounders. Generally, these two MSM models can produce different effect estimates, given that concurrent confounders are expected to have less measurement error than lagged confounders because of long measurement intervals. Also, the MSM model with weights adjusting for concurrent confounders may give biased estimates if they are affected by the exposure [52,54].

Our results should be interpreted in light of several limitations. First, measurement error in self-reported confounders such as smoking and drinking status may induce residual confounding [91,92]. Second, HR is not an ideal effect measure for causal inference as it is non-collapsible and has a built-in selection bias [93–95]. Third, using quartiles of lipid ratios instead of actual values in the analyses may reduce power and change non-differential to differential error [96,97]. Fourth, as discussed in more detail elsewhere [53], the validity of causal estimates using G-methods, including MSM, relies on four identifiability assumptions: well-defined exposure, exchangeability, positivity, and correct model specification [98–101]. Our exposures were not well-defined as many interventions can be used to change them with different effects on CVD. Although lipid-lowering medications are potentially a more well-defined exposure to study than the lipid levels themselves, novel evidence supports a continuous and graded benefit of LDL-C reduction on cardiovascular risk across different types and doses of statin-based therapies [102,103]. Furthermore, many individuals cannot achieve sufficient LDL-C reduction, do not tolerate effective doses, or are unwilling to take LDL-C-lowering medications [104]. Therefore, considering the different types, duration, and treatment doses administered are essential when evaluating lipid-lowering effects; unfortunately, these data were not available.

In addition, the issues of unmeasured confounding and model misspecification cannot be excluded. Although many time-varying and time-fixed confounders were used to justify the exchangeability assumption, there might still be some unmeasured confounders, such as diet quality, sodium intake, and household and ambient pollution. Furthermore, due to data availability, diet and physical activity, both potentially modifiable and affected by prior exposure, could not be treated as time-varying covariates. Another assumption for causal methods is positivity. We used V-stabilized weights with all baseline confounders in our analyses, which decreased the chance of random violations of the positivity assumption.

To determine the correct model, as described by Cole and Hernán [53], we informally explored the sensitivity of bias and precision of the effect estimate to different weight models specifications (e.g., considering non-linear vs. linear terms for continuous confounders using multivariate fractional polynomial models, adding lagged confounders and so on). The mean weights were also very close to one, suggested as a necessary condition for correct model specification [53].

Lastly, treating the competing risk of death as a right-censoring event can lead to a "cause-specific" interpretation including what would be observed in a counterfactual world where the competing event did not occur. Although it is unrealistic to imagine a scenario where we could

entirely remove all other causes of death and produce such a population, this will not be the case for all research questions. On the other hand, the subdistribution method may help generate predicted probabilities that individuals will be in particular states. However, evidence stated that the subdistribution approach is more “predictive” while the cause-specific hazards are more fundamentally “causal.” In other words, the subdistribution risk is helpful for causal prediction, i.e., predicting counterfactual outcomes under a certain policy relative to a second policy [105,106].

Despite these limitations, this study had several strengths. First, we used a large harmonized data set from two US representative cohort studies with a large ethnically diverse population enrolled, a considerable follow-up time with multiple measurements on various covariates, and active follow-up for incident CVD and death. Second, to the best of our knowledge, this is the first study to evaluate the effects of lipid indices and lipid ratios on the CHD and stroke separately while appropriately adjusting for time-varying confounders affected by previous exposure.

In conclusion, using MSM appropriately adjusting for time-varying confounders, this study highlights the effect of TC and LDL-C on cardiovascular events, with a stronger effect on CHD than on stroke. Our results suggest that high levels of HDL-C may be protective for CHD, but not for stroke. There was also no evidence of an adverse effect of high triglyceride levels on stroke.

### Ethics approval

The current study was performed following the ethical standards in the Declaration of Helsinki. The institutional review board approved the ARIC and MESA study protocols of each participating center. All methods were carried out following relevant guidelines and regulations. Approval for undertaking the current project was also obtained from the School of Public Health & Neuroscience Research center, Shahid Beheshti University of Medical Sciences, Tehran, Iran (IR.SBMU.PHNS.REC.1398.090).

### Consent to participate

All ARIC and MESA participants provided written informed consent.

### Consent for publication

Not applicable.

### Availability of data and material

The datasets used during the current study are available from the corresponding author on reasonable request.

### Funding

No funds, grants, or other support was received.

### Authors' contributions

All authors contributed to the study's conception and design. FK, MA.M, DK, and HS raised the presented idea and designed the study. FK performed statistical analyses under the supervision of senior authors MA.M, DK, and HS. FK, DK, MA.M, MN, and HS interpreted the results. FK developed the first draft of the manuscript. MA.M, DK, MN, and HS critically reviewed the manuscript. All authors read and approved the final manuscript.

### Declaration of Competing Interest

The authors have no relevant financial or non-financial interests to

disclose.

### Acknowledgments

This article is part of the Ph.D. thesis of the first author supported by the School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran. This study was carried out using the ARIC and MESA Research Materials obtained from the NHLBI; BioLINCC Research Materials Distribution Agreement (RMDA) V02 1d20120806. We would like to extend our thanks to the staff and coordinators of BioLINC and the other investigators, the staff, and the participants of the ARIC study and MESA for their valuable contributions.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.gloepi.2022.100075>.

### References

- [1] Roth GA, Johnson C, Abajobir A, Abd-Allah F, Abera SF, Abyu G, et al. Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015. *J Am Coll Cardiol* 2017;70(1):1–25.
- [2] Ryan A, Heath S, Cook P. Managing dyslipidaemia for the primary prevention of cardiovascular disease. *Bmj*. 2018;360.
- [3] Kobayashi D, Noto H, Shimbo T, Ino T, Osugi Y, Takahashi O, et al. Repeated measures of extremely high levels of high-density lipoprotein cholesterol and subsequent all-cause mortality and cardiovascular events: A longitudinal study. *Atherosclerosis*. 2019;288:17–25.
- [4] Hernán MÁ, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*. 2000;561–70.
- [5] Shakiba M, Mansournia MA, Salari A, Soori H, Mansournia N, Kaufman JS. Accounting for time-varying confounding in the relationship between obesity and coronary heart disease: analysis with G-estimation: the ARIC Study. *Am J Epidemiol* 2018;187(6):1319–26.
- [6] Shakiba M, Mansournia MA, Kaufman JS. Estimating effect of obesity on stroke using G-Estimation: the ARIC study. *Obesity*. 2019;27(2):304–8.
- [7] Doosti-Irani A, Mansournia MA, Collins G. Use of G-methods for handling time-varying confounding in observational research. *Lancet Glob Health* 2019;7(1):e35.
- [8] Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology*. 2003;300–6.
- [9] Mansournia MA, Nazempour M, Etminan M. Interaction contrasts and collider bias. *Am J Epidemiol* 2022 (in press).
- [10] Fekri N, Hadaegh F, Ramezankhani A, Mansournia MA. The protective effect of obesity on mortality among those with (or without) CVD cannot be fully explained by collider-stratification bias. *Int J Obes (Lond)* 2021;45(4):918–9.
- [11] Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. *bmj*. 2017;359.
- [12] Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;615–25.
- [13] Shakiba M, Mansournia M, Soori H. The bias of standard methods in estimating causal effect. 2017.
- [14] Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *LWW* 2000;11(5):550–60.
- [15] Robins JM. Association, causation, and marginal structural models. *Synthese*. 1999;151–79.
- [16] Mansournia MA, Altman DG. Inverse probability weighting. *Bmj* 2016;352:i189.
- [17] Almasi-Hashiani A, Mansournia MA, Rezaeifard A, Mohammad K. Causal effect of donor source on survival of renal transplantation using marginal structural models. *Iran J Public Health* 2018;47(5):706–12.
- [18] Almasi-Hashiani A, Nedjat S, Ghiasvand R, Safiri S, Nazempour M, Mansournia N, et al. The causal effect and impact of reproductive factors on breast cancer using super learner and targeted maximum likelihood estimation: a case-control study in Fars Province, Iran. *BMC Public Health* 2021;21(1):1219.
- [19] Khodamoradi F, Nazempour M, Mansournia N, Yazdani K, Khalili D, Mansournia MA. The effects of smoking on metabolic syndrome and its components using causal methods in the Iranian population. *Int J Prev Med* 2021; 12:118.
- [20] Smith MJ, Mansournia MA, Maringe C, Zivich PN, Cole SR, Leyrat C, et al. Introduction to computational causal inference using reproducible Stata, R, and Python code: a tutorial. *Stat Med* 2022;41(2):407–32.
- [21] ARIC I. The atherosclerosis risk in communities (ARIC) study: design and objectives. *Am J Epidemiol* 1989;129(4):687–702.
- [22] Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, et al. Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol* 2002;156(9):871–81.

- [23] Fortier I, Raina P, Van den Heuvel ER, Griffith LE, Craig C, Saliba M, et al. Maelstrom research guidelines for rigorous retrospective data harmonization. *Int J Epidemiol* 2017;46(1):103–5.
- [24] Atherosclerosis. Risk in Communities Study Description. Available from. <http://sites.csc.unc.edu/aric/desc.pub>.
- [25] MESA. Website. available from: <https://www.mesa-nhlbi.org/>.
- [26] National Cholesterol Education Program Expert Panel on Detection E, Treatment of High Blood Cholesterol in A. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. *Circulation*. 2002;106(25):3143–421.
- [27] Expert Panel on Detection E. 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(19):2486–97.
- [28] Wilkins JT, Ning H, Stone NJ, Criqui MH, Zhao L, Greenland P, et al. Coronary heart disease risks associated with high levels of HDL cholesterol. *J Am Heart Assoc* 2014;3(2):e000519.
- [29] Hirata A, Sugiyama D, Watanabe M, Tamakoshi A, Iso H, Kotani K, et al. Association of extremely high levels of high-density lipoprotein cholesterol with cardiovascular mortality in a pooled analysis of 9 cohort studies including 43,407 individuals: the EPOCH-JAPAN study. *J Clin Lipidol* 2018;12(3):674–84. e5.
- [30] Whelton PK, Carey RM, Aronow WS, Casey Jr DE, Collins KJ, Dennison Himmelfarb C, et al. ACC/AHA/AAPA/ABC/ACPM/AGS/APha/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Hypertension (Dallas, Tex: 1979) 2017;71(6):e13–115.
- [31] Association AD. 2. Classification and diagnosis of diabetes: standards of medical care in diabetes—2021. *Diabetes Care* 2020;44(Supplement 1):S15–33.
- [32] Organization WH. Physical status: The use of and interpretation of anthropometry, Report of a WHO Expert Committee. World Health Organization; 1995.
- [33] Organization WH. Waist circumference and waist-hip ratio: report of a WHO expert consultation, Geneva, 8–11 December 2008. 2011.
- [34] Duncan MS, Vasan RS, Xanthakis V. Trajectories of blood lipid concentrations over the adult life course and risk of cardiovascular disease and all-cause mortality: observations from the Framingham Study over 35 years. *J Am Heart Assoc* 2019;8(11):e011433.
- [35] Tilling K, Sterne JA, Szklo M. Estimating the effect of cardiovascular risk factors on all-cause mortality and incidence of coronary heart disease using G-estimation: the atherosclerosis risk in communities study. *Am J Epidemiol* 2002;155(8):710–8.
- [36] Hajar R. Risk factors for coronary artery disease: historical perspectives. *Heart Views* 2017;18(3):109–14.
- [37] Wang SY, Tan ASL, Claggett B, Chandra A, Khatana SAM, Lutsey PL, et al. Longitudinal associations between income changes and incident cardiovascular disease: the atherosclerosis risk in communities study. *JAMA Cardiol* 2019;4(12):1203–12.
- [38] Schultz WM, Kelli HM, Lisko JC, Varghese T, Shen J, Sandesara P, et al. Socioeconomic status and cardiovascular outcomes: challenges and interventions. *Circulation*. 2018;137(20):2166–78.
- [39] Brooks EL, Preis SR, Hwang S-J, Murabito JM, Benjamin EJ, Kelly-Hayes M, et al. Health insurance and cardiovascular disease risk factors. *Am J Med* 2010;123(8):741–7.
- [40] de Koning L, Merchant AT, Pogue J, Anand SS. Waist circumference and waist-to-hip ratio as predictors of cardiovascular events: meta-regression analysis of prospective studies. *Eur Heart J* 2007;28(7):850–6.
- [41] Larsson SC, Burgess S, Mason AM, Michaëlsson K. Alcohol consumption and cardiovascular disease: a mendelian randomization study. *Circul Genomic Precision Med* 2020;13(3):e002814.
- [42] Etmninan M, Collins GS, Mansournia MA. Using causal diagrams to improve the design and interpretation of medical research. *Chest*. 2020;158(1s):S21–s8.
- [43] Mansournia MA, Hernán MA, Greenland S. Matched designs and causal diagrams. *Int J Epidemiol* 2013;42(3):860–9.
- [44] Mansournia MA, Higgins JP, Sterne JA, Hernán MA. Biases in randomized trials: a conversation between trialists and epidemiologists. *Epidemiology*. 2017;28(1):54–9.
- [45] Etmninan M, Nazempour M, Candidate MS, Mansournia MA. Potential biases in studies of acid-suppressing drugs and COVID-19 infection. *Gastroenterology*. 2021;160(5):1443–6.
- [46] Mansournia MA, Collins GS, Nielsen RO, Nazempour M, Jewell NP, Altman DG, et al. A Checklist for statistical Assessment of Medical Papers (the CHAMP statement): explanation and elaboration. *Br J Sports Med* 2021;55(18):1009–17.
- [47] Mansournia MA, Collins GS, Nielsen RO, Nazempour M, Jewell NP, Altman DG, et al. Checklist for statistical Assessment of Medical Papers: the CHAMP statement. *Br J Sports Med* 2021;55(18):1002–3.
- [48] Mansournia MA, Nazempour M, Etmninan M. Causal diagrams for immortal time bias. *Int J Epidemiol* 2021;50(5):1405–9.
- [49] Etmninan M, Brophy JM, Collins G, Nazempour M, Mansournia MA. To adjust or not to adjust: the role of different covariates in cardiovascular observational studies. *Am Heart J* 2021;237:62–7.
- [50] Mansournia MA, Nazempour M, Etmninan M. Time-fixed vs time-varying causal diagrams for immortal time bias. *Int J Epidemiol* 2022.
- [51] D'Agostino RB, Lee ML, Belanger AJ, Cupples LA, Anderson K, Kannel WB. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. *Stat Med* 1990;9(12):1501–15.
- [52] Mansournia MA, Naimi AI, Greenland S. The implications of using lagged and baseline exposure terms in longitudinal causal and regression models. *Am J Epidemiol* 2019;188(4):753–9.
- [53] Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol* 2008;168(6):656–64.
- [54] Mansournia MA, Danaei G, Forouzanfar MH, Mahmoodi M, Jamali M, Mansournia N, et al. Effect of physical activity on functional performance and knee pain in patients with osteoarthritis: analysis with marginal structural models. *Epidemiology*. 2012;631–40.
- [55] Fewell Z, Hernán MA, Wolfe F, Tilling K, Choi H, Sterne JA. Controlling for time-dependent confounding using marginal structural models. *Stata J* 2004;4(4):402–20.
- [56] Mansournia MA, Nazempour M, Naimi AI, Collins GS, Campbell MJ. Reflection on modern methods: demystifying robust standard errors for epidemiologists. *Int J Epidemiol* 2021;50(1):346–51.
- [57] StataCorp L. Stata treatment-effects reference manual. College Station, TX: A Stata Press Publication; 2015.
- [58] Navar-Boggan AM, Peterson ED, D'Agostino Sr RB, Neely B, Sniderman AD, Pencina MJ. Hyperlipidemia in early adulthood increases long-term risk of coronary heart disease. *Circulation*. 2015;131(5):451–8.
- [59] Wadhera RK, Steen DL, Khan I, Giugliano RP, Foody JM. A review of low-density lipoprotein cholesterol, treatment strategies, and its impact on cardiovascular disease morbidity and mortality. *J Clin Lipidol* 2016;10(3):472–89.
- [60] Abdullah SM, Defina LF, Leonard D, Barlow CE, Radford NB, Willis BL, et al. Long-term association of low-density lipoprotein cholesterol with cardiovascular mortality in individuals at low 10-year risk of atherosclerotic cardiovascular disease: results from the Cooper Center Longitudinal Study. *Circulation*. 2018;138(21):2315–25.
- [61] Valdes-Marquez E, Parish S, Clarke R, Stari T, Worrall BB, Hopewell JC. Relative effects of LDL-C on ischemic stroke and coronary disease: a Mendelian randomization study. *Neurology*. 2019;92(11):e1176–87.
- [62] Baigent C, Blackwell L, Emberson J, Holland LE, Reith C, Bhalra N, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet (London, England)* 2010;376(9753):1670–81.
- [63] Allara E, Morani G, Carter P, Gkatzionis A, Zuber V, Foley CN, et al. Genetic determinants of lipids and cardiovascular disease outcomes: a wide-angled Mendelian randomization investigation. *Circul Genomic Precision Med* 2019;12(12):e002711.
- [64] Di Angelantonio E, Sarwar N, Perry P, Kaptoge S, Ray KK, Thompson A, et al. Major lipids, apolipoproteins, and risk of vascular disease. *Jama*. 2009;302(18):1993–2000.
- [65] Cannon CP, Blazing MA, Giugliano RP, McCagg A, White JA, Theroux P, et al. Ezetimibe added to statin therapy after acute coronary syndromes. *N Engl J Med* 2015;372(25):2387–97.
- [66] Vangen-Lønne AM, Ueda P, Gulayin P, Wilsgaard T, Mathiesen EB, Danaei G. Hypothetical interventions to prevent stroke: an application of the parametric g-formula to a healthy middle-aged population. *Eur J Epidemiol* 2018;33(6):557–66.
- [67] Wilsgaard T, Vangen-Lønne AM, Mathiesen E, Løchen M-L, Njølstad I, Heiss G, et al. Hypothetical interventions and risk of myocardial infarction in a general population: application of the parametric g-formula in a longitudinal cohort study—the Tromsø Study. *BMJ Open* 2020;10(5):e035584.
- [68] Ueda P, Gulayin P, Danaei G. Long-term moderately elevated LDL-cholesterol and blood pressure and risk of coronary heart disease. *PLoS One* 2018;13(7):e0200017.
- [69] Koohi F, Khalili D, Mansournia MA, Hadaegh F, Soori H. Multi-trajectories of lipid indices with incident cardiovascular disease, heart failure, and all-cause mortality: 23 years follow-up of two US cohort studies. *J Transl Med* 2021;19(1):1–13.
- [70] Haase CL, Tybjærg-Hansen A, Ali Qayyum A, Schou J, Nordestgaard BG, Frikke-Schmidt R. LCAT, HDL cholesterol and ischemic cardiovascular disease: a Mendelian randomization study of HDL cholesterol in 54,500 individuals. *J Clin Endocrinol Metabol* 2012;97(2):E248–56.
- [71] Investigators A-H. Niacin in patients with low HDL cholesterol levels receiving intensive statin therapy. *N Engl J Med* 2011;365(24):2255–67.
- [72] Kosmas CE, Silverio D, Sourlas A, Montan PD, Guzman E. Dysfunctional high-density lipoprotein and atherogenesis. *Vessel Plus* 2019;3:2.
- [73] Chang TI, Streja E, Moradi H. Could high-density lipoprotein cholesterol predict increased cardiovascular risk? *Curr Opin Endocrinol Diabetes Obes* 2017;24(2):140–7.
- [74] Holmes MV, Asselbergs FW, Palmer TM, Drenos F, Lanktree MB, Nelson CP, et al. Mendelian randomization of blood lipids for coronary heart disease. *Eur Heart J* 2015;36(9):539–50.
- [75] Zanon P, Khetarpal SA, Larach DB, Hancock-Cerutti WF, Millar JS, Cuchel M, et al. Rare variant in scavenger receptor BI raises HDL cholesterol and increases risk of coronary heart disease. *Science (New York, NY)* 2016;351(6278):1166–71.
- [76] Frikke-Schmidt R, Nordestgaard BG, Jensen GB, Steffensen R, Tybjærg-Hansen A. Genetic variation in ABCA1 predicts ischemic heart disease in the general population. *Arterioscler Thromb Vasc Biol* 2008;28(1):180–6.
- [77] Wilkins JT, Ning H, Stone NJ, Criqui MH, Zhao L, Greenland P, et al. Coronary heart disease risks associated with high levels of HDL cholesterol. *J Am Heart Assoc* 2014;3(2):e000519.

- [78] Madsen CM, Varbo A, Nordestgaard BG. Extreme high high-density lipoprotein cholesterol is paradoxically associated with high mortality in men and women: two prospective cohort studies. *Eur Heart J* 2017;38(32):2478–86.
- [79] Di Angelantonio E, Sarwar N, Perry P, Kaptoge S, Ray K, Thompson A, et al. Emerging risk factors collaboration. Major lipids, apolipoproteins, and risk of vascular disease. *JAMA* 2009;302:1993–2000.
- [80] Sarwar N, Sandhu MS, Ricketts SL, Butterworth AS, Di Angelantonio E, Boekholdt SM, et al. Triglyceride-mediated pathways and coronary disease: collaborative analysis of 101 studies. *Lancet (London, England)* 2010;375(9726):1634–9.
- [81] Dewey FE, Gusarova V, O'Dushlaine C, Gottesman O, Trejos J, Hunt C, et al. Inactivating variants in ANGPTL4 and risk of coronary artery disease. *N Engl J Med* 2016;374(12):1123–33.
- [82] Helgadottir A, Gretarsdottir S, Thorleifsson G, Hjartarson E, Sigurdsson A, Magnusdottir A, et al. Variants with large effects on blood lipids and the role of cholesterol and triglycerides in coronary disease. *Nat Genet* 2016;48(6):634–9.
- [83] Hindy G, Engström G, Larsson SC, Traylor M, Markus HS, Melander O, et al. Role of blood lipids in the development of ischemic stroke and its subtypes: a Mendelian randomization study. *Stroke*. 2018;49(4):820–7.
- [84] Ganda OP, Bhatt DL, Mason RP, Miller M, Boden WE. Unmet need for adjunctive dyslipidemia therapy in hypertriglyceridemia management. *J Am Coll Cardiol* 2018;72(3):330–43.
- [85] Nichols GA, Philip S, Reynolds K, Granowitz CB, Fazio S. Increased residual cardiovascular risk in patients with diabetes and high versus normal triglycerides despite statin-controlled LDL cholesterol. *Diabetes Obes Metab* 2019;21(2):366–71.
- [86] Marston NA, Giugliano RP, Im K, Silverman MG, O'Donoghue ML, Wiviott SD, et al. Association between triglyceride lowering and reduction of cardiovascular risk across multiple lipid-lowering therapeutic classes: a systematic review and meta-regression analysis of randomized controlled trials. *Circulation*. 2019;140(16):1308–17.
- [87] Millán J, Pintó X, Muñoz A, Zúñiga M, Rubiés-Prat J, Pallardo LF, et al. Lipoprotein ratios: physiological significance and clinical usefulness in cardiovascular prevention. *Vasc Health Risk Manag* 2009;5:757.
- [88] Lemieux I, Lamarche B, Couillard C, Pascot A, Cantin B, Bergeron J, et al. Total cholesterol/HDL cholesterol ratio vs LDL cholesterol/hdl cholesterol ratio as indices of ischemic heart disease risk in men: the Quebec CARDIOVASCULAR Study. *Arch Intern Med* 2001;161(22):2685–92.
- [89] Tamada M, Makita S, Abiko A, Naganuma Y, Nagai M, Nakamura M. Low-density lipoprotein cholesterol to high-density lipoprotein cholesterol ratio as a useful marker for early-stage carotid atherosclerosis. *Metab Clin Exp* 2010;59(5):653–7.
- [90] Hadaegh F, Khalili D, Ghasemi A, Tohidi M, Sheikholeslami F, Azizi F. Triglyceride/HDL-cholesterol ratio is an independent predictor for coronary heart disease in a population of Iranian men. *Nutr Metab Cardiovasc Dis* 2009;19(6):401–8.
- [91] Greenland S. The effect of misclassification in the presence of covariates. *Am J Epidemiol* 1980;112(4):564–9.
- [92] Pakzad R, Nedjat S, Yaseri M, Salehiniya H, Mansournia N, Nazempour M, et al. Effect of smoking on breast cancer by adjusting for smoking misclassification bias and confounders using a probabilistic bias analysis method. *Clin Epidemiol* 2020;12:557–68.
- [93] Hernán MA. The hazards of hazard ratios. *Epidemiology* 2010;21(1):13–5.
- [94] Greenland S. Absence of confounding does not correspond to collapsibility of the rate ratio or rate difference. *Epidemiology*. 1996;7(5):498–501.
- [95] Mansournia MA, Greenland S. The relation of collapsibility and confounding to faithfulness and stability. *Epidemiology*. 2015;26(4):466–72.
- [96] Greenland S. Avoiding power loss associated with categorization and ordinal scores in dose-response and trend analysis. *Epidemiology*. 1995;450–4.
- [97] Flegal KM, Keyl PM, Nieto FJ. Differential misclassification arising from nondifferential errors in exposure measurement. *Am J Epidemiol* 1991;134(10):1233–46.
- [98] Abdollahpour I, Nedjat S, Mansournia MA, Schuster T. Estimation of the marginal effect of regular drug use on multiple sclerosis in the Iranian population. *PLoS One* 2018;13(4):e0196244.
- [99] Almasi-Hashiani A, Nedjat S, Mansournia MA. Causal methods for observational research: a primer. *Arch Iran Med* 2018;21(4):164–9.
- [100] Mokhayeri Y, Hashemi-Nazari SS, Khodakarim S, Safiri S, Mansournia N, Mansournia MA, et al. Effects of hypothetical interventions on ischemic stroke using parametric G-formula. *Stroke*. 2019;50(11):3286–8.
- [101] Abdollahpour I, Nedjat S, Almasi-Hashiani A, Nazempour M, Mansournia MA, Luque-Fernandez MA. Estimating the marginal causal effect and potential impact of waterpipe smoking on risk of multiple sclerosis using the targeted maximum likelihood estimation method: a large, population-based incident case-control study. *Am J Epidemiol* 2021;190(7):1332–40.
- [102] Trialists CT. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90 056 participants in 14 randomised trials of statins. *The Lancet* 2005;366(9493):1267–78.
- [103] Trialists CT. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170 000 participants in 26 randomised trials. *The Lancet* 2010;376(9753):1670–81.
- [104] Koskinas KC, Siontis GC, Piccolo R, Mavridis D, Räber L, Mach F, et al. Effect of statins and non-statin LDL-lowering medications on cardiovascular outcomes in secondary prevention: a meta-analysis of randomized trials. *Eur Heart J* 2018;39(14):1172–80.
- [105] Rudolph JE, Lesko CR, Naimi AI. Causal inference in the face of competing events. *Curr Epidemiol Rep* 2020;7(3):125–31.
- [106] Young JG, Stensrud MJ, Tchetgen Tchetgen EJ, Hernán MA. A causal framework for classical statistical estimands in failure-time settings with competing events. *Stat Med* 2020;39(8):1199–236.