The Impact of Glycated Hemoglobin (HbA_{1c}) on Cardiovascular Disease Risk: A Mendelian Randomization Study Using UK Biobank

Shiu Lun Au Yeung,[1] Shan Luo,[1] and C. Mary Schooling[1,2]

Diabetes Care 2018;41:1991–1997 | https://doi.org/10.2337/dc18-0289

OBJECTIVE

Glycated hemoglobin (HbA_{1c}) is positively associated with cardiovascular disease (CVD), although evidence is primarily observational. Mendelian randomization studies have only examined its relation with subtypes of CVD. We examined the relation of HbA_{1c} with CVD and its subtypes in the UK Biobank using Mendelian randomization.

RESEARCH DESIGN AND METHODS

We used 38 genetic variants strongly and independently related to HbA_{1c} (n = 123,665) applied to the UK Biobank (n = 392,038). We used inverse variance weighting (IVW) to obtain the associations of HbA_{1c} with CVD, coronary artery disease (CAD), and stroke (overall and stroke subtypes). Sensitivity analyses included Mendelian randomization (MR)-Egger, a weighted median, and exclusion of potentially invalid single nucleotide polymorphisms (SNPs). We also applied the same genetic instruments to CARDIoGRAMplusC4D (Coronary ARtery DIsease Genome wide Replication and Meta-analysis [CARDIoGRAM] plus The Coronary Artery Disease [C4D] Genetics) 1000 Genomes–based genome-wide association study (n = 184,305) as a validation for CAD.

RESULTS

In the UK Biobank, HbA_{1c} was not associated with CVD using IVW (odds ratio [OR] 1.11 per %, 95% CI 0.83–1.48). However, HbA_{1c} was associated with increased CAD risk (OR 1.50 per %, 95% CI 1.08–2.11) with directionally consistent results from MR-Egger and weighted median. The positive association with CAD was more pronounced when we excluded potentially invalid SNPs (OR 2.24 per %, 95% CI 1.55–3.25). The positive association was replicated in CARDIoGRAM (OR 1.52 per %, 95% CI 1.03–2.26). The association of HbA_{1c} with stroke and its subtypes was less clear given the low number of cases.

CONCLUSIONS

HbA_{1c} likely causes CAD. The underlying mechanisms remain to be elucidated.

Observational studies strongly suggest a link between type 2 diabetes and coronary artery disease (CAD), but these observations could be confounded by lack of physical activity and obesity (1–4). Randomized controlled trials, such as the ACCORD (Action to Control Cardiovascular Risk in Diabetes) trial, unexpectedly showed

[1]School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, People's Republic of China
[2]Graduate School of Public Health and Health Policy, City University of New York, New York, NY
Corresponding author: Shiu Lun Au Yeung, ayslryan@hku.hk.

Received 7 February 2018 and accepted 22 May 2018.
This article contains Supplementary Data online at http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc18-0289/-/DC1.

intensive glycemic control did not substantially reduce the risk of cardiovascular disease (CVD) events and may even have increased overall mortality (5), which remained evident after prolonged follow-up (6,7). However, these results have not been consistently seen in all relevant trials, such as the Steno-2 Study (8). Differences in treatment regimen and sample may have contributed to these discrepancies (9). Other relevant trials, such as the LEADER (Liraglutide Effect and Action in Diabetes: Evaluation of Cardiovascular Outcome Results) trial and the EMPA-REG OUTCOME (BI 10773 [Empagliflozin] Cardiovascular Outcome Event Trial in Type 2 Diabetes Mellitus Patients) trial, also suggested potential beneficial effects of liraglutide and empagliflozin in reducing CVD (10,11). Although randomized controlled trials are less vulnerable to confounding than observational studies, whether results from these trials, which are primarily in people with diabetes, generalize to the general population is uncertain (12). Moreover, interventions in randomized controlled trials may have off-target effects. Notably, some of the agents used to achieve glycemic control, such as sulphonylureas, have been implicated in CVD (13).

Mendelian randomization studies, which are less prone to biases particularly confounding, through use of genetic variants randomly allocated during conception, suggest a positive causal relation of dysglycemia and diabetes with CAD (14–16). However, the number of genetic instruments used for glycated hemoglobin (HbA_{1c}) (n = 9), which represents the overall blood glucose level during the previous 2–3 months, was relatively limited, making assessment of potential violations of the underlying assumptions less reliable. Whether HbA_{1c} has an overall effect on CVD given the heterogeneity of the phenotype and the observational nature of the evidence is also uncertain (17). Trials suggest glycemic traits may have different effects on different CVD subtypes, such as stroke (9), which is a major contributor to the disease burden in many regions, including Asia (18), which also have high rates of diabetes (19). To date, only one Mendelian randomization study has considered the relation of glycemic traits with stroke where they found fasting glucose potentially related to large-artery stroke but not to other ischemic stroke subtypes (20). However, the study did not examine the effect on overall stroke or hemorrhagic stroke.

To address these research gaps, we implemented a two-sample Mendelian randomization study to assess the relation of HbA_{1c} with CVD and its subtypes by using genetic predictors of HbA_{1c} from the most up-to-date genome-wide association study (GWAS) of HbA_{1c} in MAGIC (Meta-Analyses of Glucose and Insulin-Related Traits Consortium) (21) applied to the UK Biobank (22,23), one of the largest population-based cohorts globally with extensive phenotyping and genotyping. We also verified the association of HbA_{1c} with CAD using the largest most extensively genotyped CAD case-control study independent of the UK Biobank; that is, CARDIoGRAMplusC4D (Coronary ARtery DIsease Genome wide Replication and Meta-analysis [CARDIoGRAM] plus The Coronary Artery Disease [C4D] Genetics) 1000 Genomes–based GWAS (24).

RESEARCH DESIGN AND METHODS

This is a two-sample Mendelian randomization study. We obtained genetic associations with HbA_{1c} from MAGIC, and with CVD and its subtypes from the UK Biobank and CARDIoGRAMplusC4D 1000 Genomes–based GWAS.

Assumptions of Mendelian Randomization

Mendelian randomization relies on three stringent assumptions (25). First, the genetic instruments are strongly predictive of HbA_{1c}. Second, the association of genetic instruments with CVD is not confounded. Last, the effect of the genetic instrument on CVD should be fully mediated via HbA_{1c} (i.e., the exclusion restriction assumption).

Participants

MAGIC—Genetic Predictors of HbA_{1c}

MAGIC includes a meta-analysis of GWAS of HbA_{1c} in 159,940 adults without diabetes, including 123,665 participants of European ancestry, with imputation using the phase 2 of the International HapMap Project reference panel (21). The mean age of participants in most of the studies was older than 50 years. HbA_{1c}, NGSP percent, was adjusted for age, sex, study-specific covariates, and genomic control. To reduce confounding by population stratification, we only selected single nucleotide polymorphisms (SNPs) reaching genome-wide significance ($P < 5 \times 10^{-8}$) in participants of European descent, which gave 43 SNPs. After removing 4 SNPs (rs11154792, rs3824065, rs10823343, and rs2408955) in linkage disequilibrium with the other SNPs ($R^2 \geq 0.05$), 39 SNPs were retained.

Genetic Predictors of CVD

UK Biobank

The UK Biobank is one of the largest biobanks globally. It recruited ≥500,000 participants (aged 40–69 years) in the U.K. from 2006 to 2010. Participants completed a questionnaire and physical assessment. Biochemical assays, genotyping, and longitudinal follow-up via record linkage to medical and mortality records are ongoing, as described in detail elsewhere (22,23). Prevalent disease was coded using ICD-9 and ICD-10, and cause of death was coded using ICD-10. Genotyping was performed using two very similar arrays, including Affymetrix UK BiLEVE (UK Biobank Lung Exome Variant Evaluation) Axiom array (~50,000 participants) and Affymetrix UK Biobank Axiom array (~450,000 participants). The SNPs included in this study were directly genotyped or imputed using the Haplotype Reference Consortium panel. We restricted our analysis to people of genetically verified white British descent to reduce confounding by population stratification, as in a previous similar study (26). We also excluded participants who were extensively related (more than 10 putative third-degree relatives in the kinship table), who had poor-quality genotyping (i.e., missing rate ≥1.5%), who had sex chromosome aneuploidy, or whose self-reported and genetic sex did not match. The mean age of the participants was 56.9 years.

CARDIoGRAMplusC4D 1000 Genomes–Based GWAS

CARDIoGRAMplusC4D 1000 Genomes–based GWAS is a meta-analysis of GWAS of CAD case (n = 60,801) and control (n = 123,504) subjects of mainly European descent (77%), with imputation using the 1000 Genomes phase 1 v3 reference panel (24). CAD was defined in various ways, such as diagnosis of myocardial infarction, acute coronary syndrome, chronic stable angina, or coronary stenosis >50%. Diagnoses were based on clinical diagnosis, procedures

(coronary angiography results or bypass surgery), use of medications or symptoms that indicate angina, or self-report of a doctor's diagnosis, as described elsewhere (24). CARDIoGRAMplusC4D 1000 Genomes–based GWAS was adjusted for study-specific covariates (e.g., age and sex) and genomic control.

Exposure

The exposure was genetically predicted HbA_{1c} (%).

Outcomes

The primary outcomes were prevalent CVD (defined as ICD-9 401–459.9 and ICD-10 I10–I99) and its subtypes, including CAD (ICD-9 410–414.9 and ICD-10 I20–I25.9), ischemic stroke (ICD-9 434 and 436 and ICD-10 I63–I64), and hemorrhagic stroke (ICD-9 430–431 and ICD-10 I60–I61), based on self-reports and hospital episodes, and death from CVD (ICD-10 I10–I99) or its subtypes CAD (ICD-10 I20–I25.9l), ischemic stroke (ICD-10 I63–I64), and hemorrhagic stroke (ICD-10 I60–I61) from death records, following the recommended definitions of the UK Biobank Stroke Outcomes Group (27). For completeness, we also considered other CVD (i.e., all CVD excluding stroke and CAD, which was mainly hypertensive diseases) and other stroke (i.e., all stroke excluding ischemic and hemorrhagic stroke). CVD, CAD, and stroke mortality, based on primary cause of death, were also considered separately as secondary outcomes.

Statistical Analysis

We assessed departure from Hardy-Weinberg equilibrium for each SNP using χ^2 tests with Bonferroni correction to correct for multiple comparison (0.05/39 = 0.00128). We used ANOVA (continuous) and χ^2 test (categorical) to examine whether the genetic variants were associated with factors potentially confounding the association of HbA_{1c} with CVD, including Townsend deprivation index, education, age, BMI, smoking, and alcohol drinking, in the UK Biobank, with Bonferroni correction to correct for multiple comparisons (0.0002, based on 0.05/234 derived from 39 SNPs $\times$ 6 traits). We obtained the association of each SNP with CVD and its subtypes using multivariable logistic regression in the UK Biobank, adjusted for age, sex, genotyping array, and 10 principal components.

We conducted our main analysis using inverse variance weighting (IVW) with multiplicative random effects, which is a weighted regression of gene-outcome associations on gene-exposure associations for UK Biobank and CARDIoGRAMplusC4D 1000 Genomes–based GWAS. Given IVW assumes no horizontal pleiotropy, which cannot be empirically assessed, we used the I^2 of the Wald estimates (SNP-outcome association divided by SNP-exposure association) to indicate the presence of invalid instruments. In the presence of invalid SNPs (i.e., SNPs that have effect on the outcome not via HbA_{1c}), IVW will be invalid. As such, we also conducted several sensitivity analyses to assess the robustness of our results to potential violations of the Mendelian randomization assumptions because these analyses have different assumptions for validity, as described below. Although these approaches may have different statistical power (e.g., wider CIs for Mendelian randomization [MR]-Egger), the rationale is that if these approaches give a similar conclusion regarding the association of HbA_{1c} with the outcomes, then we are more confident in inferring that the positive findings are unlikely driven by violation of the Mendelian randomization assumptions (28).

Instrument Strength

To assess instrument strength, we computed the *F* statistic for the association of genetic instruments with HbA_{1c}, assuming the included genetic variants explained at least 4.2% of the HbA_{1c} variance (the lower bound of variance explained based on the previous GWAS) (21). Higher *F* statistics indicate a stronger instrument.

Sensitivity Analysis

1) MR-Egger. We conducted MR-Egger regression, which produces valid estimates even if all the genetic instruments are invalid, as long as the InSIDE (Instrument Strength Independent of Direct Effect) assumption holds (29). We also presented the intercept *P* value from the MR-Egger regression because a significant intercept indicates the IVW estimate may be invalid due to horizontal pleiotropy.

2) Weighted Median. We also used a weighted median, which produces valid estimates as long as more than 50% of the information is derived from valid SNPs (30).

3) Exclusion of Potentially Invalid SNPs. We identified potentially invalid genetic instruments (SNPs) in two ways. First, we excluded SNPs related to potential causes of CAD based on the GWAS Catalog (31)/PhenoScanner (32) and SNPs associated with potential confounders in the UK Biobank (set 1). Second, we additionally excluded SNPs defined as "erythrocytic" in the original GWAS because they did not predict diabetes and hence may be irrelevant to glycemic exposure (i.e., invalid) (21) (set 2). Supplementary Table 1 summarizes the choice of SNPs in different sensitivity analyses.

To rule out the possibility of false positives due to inclusion of related individuals, we also repeated the analyses including only unrelated participants.

All analyses were performed using R version 3.3.2 (R Development Core Team, Vienna, Austria) with the R package (TwosampleMR).

Ethics Approval

UK Biobank received ethics approval from the National Health Service National Research Ethics Service, and participants provided written informed consent. No ethics approval was required for the analysis using publicly available data (CARDIoGRAMplusC4D 1000 Genomes–based GWAS).

RESULTS

Among 502,642 participants in the UK Biobank, 442,698 (88%) were British white. After excluding those who had permanently withdrawn, with poor quality or missing genotype, with a mismatch between self-reported and genetic sex or ancestry, had sex chromosome aneuploidy, or extensive relatedness, 392,038 participants remained for subsequent analyses. Among these 392,038 people, 158,601 had prevalent CVD, 29,293 prevalent CAD, and 9,042 prevalent stroke (3,707 ischemic and 1,655 hemorrhagic), with some participants having more than one condition according to the data available in April 2018. Since the baseline recruitment in March 2006, there were 2,313 CVD deaths, including 1,294 CAD and 356 stroke deaths, where the latest date of death was 16 February 2016.

Of the 39 SNPs for HbA_{1c}, 1 SNP (rs1800562, $P = 0.0006$) violated Hardy-Weinberg equilibrium (Supplementary Table 2) and was discarded. Supplementary

Table 1—Association of HbA_{1c} (%) with CVD, CAD, and stroke (overall and stroke subtypes) using Mendelian randomization in the UK Biobank

		IVW with multiplicative random effects			MR-Egger			Weighted median	
Instrument	Outcome*	OR	95% CI	I^2 of Wald estimates	OR	95% CI	Intercept	*P* value for intercept	OR	95% CI
All SNPs (38)	CVD	1.11	0.83–1.48	0.89	1.12	0.64–1.96	0.000	0.95	1.25	1.07–1.47
	CAD	1.50	1.08–2.11	0.71	1.25	0.66–2.36	0.004	0.50	1.41	1.02–1.94
	Stroke	1.39	0.93–2.07	0.41	0.95	0.45–2.01	0.008	0.25	1.23	0.77–1.95
	Ischemic	1.04	0.62–1.75	0.17	0.49	0.19–1.26	0.017	0.07	0.65	0.33–1.27
	Hemorrhagic	1.27	0.58–2.81	0.20	1.25	0.28–5.68	0.000	0.98	0.82	0.26–2.59
	Others	1.84	0.99–3.39	0.41	1.54	0.47–5.03	0.004	0.73	1.42	0.65–3.11
	Other CVD	0.97	0.75–1.26	0.85	1.07	0.65–1.75	−0.002	0.65	1.08	0.91–1.30
Excluded SNPs if associated with potential causes of CAD or confounders (23)†	CVD	1.20	1.02–1.42	0.57	1.10	0.81–1.50	0.002	0.50	1.26	1.07–1.47
	CAD	1.47	1.10–1.97	0.53	0.97	0.59–1.60	0.011	0.07	1.27	0.89–1.81
	Stroke	1.47	1.04–2.07	0.00	0.81	0.43–1.52	0.016	0.04	1.22	0.76–1.95
	Ischemic	0.90	0.53–1.53	0.00	0.53	0.20–1.40	0.014	0.22	0.64	0.30–1.39
	Hemorrhagic	1.69	0.66–4.35	0.30	0.80	0.14–4.47	0.020	0.32	0.86	0.27–2.75
	Others	2.14	1.20–3.81	0.17	1.25	0.43–3.60	0.014	0.25	1.42	0.66–3.07
	Other CVD	1.06	0.90–1.25	0.53	1.15	0.84–1.56	−0.002	0.56	1.11	0.93–1.32
Excluded SNPs if classified as erythrocytic, associated with potential causes of CAD or confounders (14)‡	CVD	1.27	1.03–1.56	0.43	0.82	0.51–1.31	0.009	0.07	1.20	0.96–1.51
	CAD	2.24	1.55–3.25	0.39	2.10	0.80–5.50	0.001	0.89	1.90	1.25–2.88
	Stroke	1.61	0.98–2.65	0.00	0.88	0.25–3.03	0.013	0.31	1.40	0.71–2.74
	Ischemic	1.14	0.53–2.47	0.00	0.45	0.07–3.02	0.020	0.31	0.97	0.35–2.73
	Hemorrhagic	1.26	0.40–3.93	0.00	1.89	0.10–35.05	−0.009	0.77	0.72	0.16–3.36
	Others	2.32	0.96–5.59	0.25	1.14	0.12–10.76	0.015	0.51	1.20	0.40–3.61
	Other CVD	0.97	0.82–1.14	0.00	0.65	0.43–0.98	0.008	0.06	0.95	0.76–1.18

*Definitions of disease as below: prevalent CVD (defined as ICD-9 401–459.9, ICD-10 I10–I99) and its subtypes, including CAD (defined as ICD-9 410–414.9, ICD-10 I20–I25.9), stroke (defined as ICD-9 430, 431, 434, 436; ICD-10 I60, I61, I63, I64), ischemic stroke (defined as ICD-9 434, 436, ICD-10 I63–I64), and hemorrhagic stroke (defined as ICD-9 430, 431, ICD-10 I60–I61). Cases and mortality were both included. †Set 1: SNP excluded if associated with potential causes of CAD or confounders based on public data sources (PhenoScanner and GWAS Catalog) and UK Biobank. ‡Set 2: SNP excluded if classified as erythrocytic based on the HbA_{1c} GWAS or associated with potential causes of CAD or confounders based on public data sources (PhenoScanner and GWAS Catalog) and UK Biobank.

Table 3 lists the 38 SNPs used as genetic instruments. The *F* statistic for the association of the 38 SNPs on HbA_{1c} was 142, suggesting little weak instrument bias. A few SNPs were associated with potential confounders, two SNPs with education (rs9818758 and rs11964178), five SNPs with BMI (rs8192675, rs7756992, rs17747324, rs10774625, and rs1558902), and two SNPs with smoking (rs10774625 and rs17509001) after Bonferroni correction (Supplementary Table 2). According to GWAS Catalog or PhenoScanner, 13 SNPs were related to potential causes of CAD. Based on this information, we repeated the analyses using the two exclusion criteria for choice of SNPs. First, we excluded 15 SNPs related to potential confounders or causes of CAD, leaving 23 SNPs (set 1). Second, we additionally excluded SNPs defined as "erythrocytic" in the original GWAS (9 SNPs) among the 23 SNPs because they did not predict diabetes and hence may be irrelevant to glycemic exposure (i.e., invalid) (21), leaving 14 SNPs (set 2).

Table 1 shows HbA_{1c} was not clearly associated with CVD using all 38 SNPs (odds ratio [OR] 1.11 per %, 95% CI 0.83–1.48). However, higher HbA_{1c} was associated with higher CAD risk using IVW using all 38 SNPs (OR 1.50 per %, 95% CI 1.08–2.11), with directionally consistent results from MR-Egger and weighted median. After we excluded potentially pleiotropic SNPs or those related to confounders (23 SNPs: set 1), the positive associations remained for CAD in IVW (OR 1.47 per %, 95% CI 1.10–1.97), with directionally consistent results from the weighted median method. The results for CAD were most consistent across IVW, MR-Egger, and weighted median when we further restricted SNPs that were nonerythrocytic (set 2). The association of HbA_{1c} with stroke and its subtypes appeared heterogeneous, although these estimates had wide CIs. MR-Egger intercepts suggested little evidence of directional pleiotropy in all analyses. Heterogeneity in the Wald estimates decreased after potentially invalid SNPs were removed.

Table 2 shows HbA_{1c} was positively associated with CAD in CARDIoGRAMplusC4D 1000 Genomes–based GWAS using IVW (OR 1.52 per %, 95% CI 1.03–2.26), with directionally consistent estimates from

Table 2—Association of HbA_{1c} (%) with CAD using Mendelian randomization in CARDIoGRAMplusC4D 1000 Genomes–based GWAS

		IVW weighting with multiplicative random effects			MR-Egger				Weighted median	
Instrument	Outcome	OR	95% CI	I^2 of Wald estimates	OR	95% CI	Intercept	*P* value for intercept	OR	95% CI
All 38 SNPs	CAD	1.52	1.03–2.26	0.74	1.64	0.73–3.71	−0.002	0.83	1.50	1.09–2.05
23 SNPs (set 1)*	CAD	1.30	0.98–1.73	0.36	2.05	1.21–3.48	−0.011	0.06	1.42	1.02–1.95
14 SNPs (set 2)†	CAD	1.27	0.90–1.78	0.14	1.79	0.75–4.27	−0.007	0.41	1.36	0.90–2.06

*Set 1: SNP excluded if associated with potential causes of CAD or confounders based on public data sources (GWAS Catalog and PhenoScanner) and UK Biobank. †Set 2: SNP excluded if classified as erythrocytic based on the HbA_{1c} GWAS or associated with potential causes of CAD or confounders based on public data sources (GWAS Catalog and PhenoScanner) and UK Biobank.

sensitivity analyses including MR-Egger (OR 1.64 per %, 95% CI 0.73–3.71) and the weighted median (OR 1.50 per %, 95% CI 1.09–2.05). There was little evidence of directional pleiotropy based on the MR-Egger intercept (−0.002, *P* = 0.83). Similar to the results from UK Biobank, the estimates all similarly suggested detrimental effects of HbA_{1c} on CAD regardless of the SNP selection. Heterogeneity in the Wald estimates decreased after potentially invalid SNPs were removed.

The associations of HbA_{1c} with CVD, CAD, and stroke mortality were less clear (Supplementary Table 4), with wide CIs, most likely due to the low mortality rate in the UK Biobank. Similar conclusions were drawn when we restricted our analyses to unrelated participants (Supplementary Table 5).

CONCLUSIONS

To our knowledge, this is the first Mendelian randomization study using the most recently published GWAS of HbA_{1c} applied to both the UK Biobank and CARDIoGRAMplusC4D 1000 Genomes–based GWAS, encompassing more than 700,000 participants. HbA_{1c} was positively associated with CAD in UK Biobank, which was replicated in CARDIoGRAMplusC4D 1000 Genomes–based GWAS, consistent with previous observational studies and an earlier Mendelian randomization study (1–3,15). We cannot exclude HbA_{1c} being associated with CVD. Our study is suggestive of different effects of HbA_{1c} on stroke subtypes, although the number of events in UK Biobank was not enough to allow precise estimates and should be examined further in large stroke consortiums.

Previous observational studies have consistently reported a positive association of HbA_{1c} with CAD, although they are susceptible to confounding (1–3). Randomized controlled trials targeting HbA_{1c} reduction are difficult to interpret given the interventions, primarily on lifestyle, may have multiple effects that do not necessarily only reflect the effect of HbA_{1c} on CAD (33,34). Our study adds by showing that higher HbA_{1c} is positively associated with CAD using a Mendelian randomization study in two different large studies as well as using different analytics and SNP selections. Combining the results obtained from UK Biobank and CARDIoGRAMplusC4D did not change the conclusion (Supplementary Table 6). Considering triangulation of the evidence from different designs with different underlying assumptions, HbA_{1c} may be causal in the development of CAD in the general population (28). The exact mechanistic pathways remain to be elucidated.

Our study does not provide strong evidence for the same magnitude of association of HbA_{1c} with CVD, contrary to previous observational studies (17). This discrepancy could indicate potential confounding or selection bias in observational studies. Potentially different associations of HbA_{1c} with CAD and other CVD subtypes is consistent with the argument that CVD subtypes have different etiologies with different contributions of each factor (35). UK Biobank had a low response rate (~5%) at recruitment, although a low response rate at recruitment does not necessarily invalidate causal inference (36). This is evident from the similar estimates obtained from the UK Biobank and CARDIoGRAMplusC4D, which used different study designs and sampling approaches. Recruitment of generally healthier people into the UK Biobank study would also not explain the different findings for CAD and CVD. Alternatively, given the UK Biobank recruited from age 40 years to 69 years, with an average age of 57 years, a different pattern of death by age from specific types of CVD related to HbA_{1c} genetics would artifactually generate different associations of HbA_{1c} with CVD by subtype (37) because of varying levels of left truncation from the underlying birth cohort.

Although we included more than 700,000 participants in this study and used Mendelian randomization to reduce confounding, some limitations exist. The validity of Mendelian randomization depends on whether the three underlying assumptions, as described in the RESEARCH DESIGN AND METHODS, are satisfied; that is, the instruments predict the exposure, the instruments are not confounded, and the instruments affect the outcome only via the exposure (25). In our study, we used genetic variants predicting HbA_{1c} identified in GWAS of people of European descent to reduce weak instrument bias (indicated by the *F* statistics). Restricting the samples to adults mostly of European descent reduces the likelihood of confounding by population stratification.

We also assessed the associations of the genetic variants with potential confounders and found little association with most confounders (Supplementary Table 2), which would not have been possible using summary statistics from GWAS. Although we could not assess whether the genetic instruments were associated with the outcomes only via their association with HbA_{1c} (exclusion-restriction assumption), we conducted several sensitivity analyses, such as MR-Egger and a weighted median, which have different assumptions for validity, although MR-Egger has reduced statistical power. We also repeated the analyses excluding potentially pleiotropic SNPs, which may violate the exclusion-restriction assumption (38). Given the

consistent results for HbA_{1c} on CAD for these different approaches with different assumptions, the association of HbA_{1c} on CAD is likely to be causal. Repeating the analyses for CAD by sex, as a check, showed similar patterns (Supplementary Table 7). When we repeated the CAD analyses without self-reports, the results were most consistent excluding erythrocytic SNPs (data not shown). We also repeated the analyses restricted to erythrocytic SNPs and found a less clear relation of HbA_{1c} with CAD. This is expected because these SNPs did not predict diabetes, and hence, these SNPs are likely irrelevant to the glycemic exposure (Supplementary Table 8) (21).

Although we used one of the largest possible studies, the relatively low number of stroke cases led to imprecise estimates. The suggestive differences in the relation of HbA_{1c} and stroke subtypes seen in our study should be examined elsewhere using large GWAS consortium or settings where stroke is more prevalent, such as China (39).

Lastly, we were unable to use an allele score approach, which may increase statistical power because HbA_{1c} was not available from the UK Biobank at the time this study was conducted.

Our study provides more evidence of a causal role of HbA_{1c} in CAD. Interventions that target HbA_{1c} reduction may be potential targets for reducing the global burden of CAD. Future studies should also clarify the effect of HbA_{1c} on CVD subtypes, which may provide additional insight into the global distribution of CVD subtypes such as stroke, which is more prevalent in Asians.

Acknowledgments. This research was conducted using the UK Biobank Resource (application number: 14864). Dr. S.L. Lin (The University of Hong Kong) provided valuable advice at the planning stage of this study. The work presented here is the responsibility of the authors and not anyone acknowledged. Data on CAD were contributed by CARDIoGRAMplusC4D investigators and were downloaded from www.cardiogramplusc4d.org.

Funding. This study was funded by the Small Project Funding, The University of Hong Kong (#201409176231).

The funder had no role in the design, analyses, interpretation of results, or writing of the paper.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. S.L.A.Y., S.L., and C.M.S. designed the study, wrote the analysis plan, and interpreted the results. S.L.A.Y. performed analyses with feedback from S.L. and C.M.S. S.L.A.Y. wrote the first draft of the manuscript with critical feedback and revisions from S.L. and C.M.S. S.L.A.Y., S.L., and C.M.S. gave final approval of the version to be published. S.L.A.Y. is the guarantor of this study and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Pai JK, Cahill LE, Hu FB, Rexrode KM, Manson JE, Rimm EB. Hemoglobin A1c is associated with increased risk of incident coronary heart disease among apparently healthy, nondiabetic men and women. J Am Heart Assoc 2013;2:e000077
2. Selvin E, Coresh J, Golden SH, Brancati FL, Folsom AR, Steffes MW. Glycemic control and coronary heart disease risk in persons with and without diabetes: the Atherosclerosis Risk in Communities study. Arch Intern Med 2005;165:1910–1916
3. Selvin E, Steffes MW, Zhu H, et al. Glycated hemoglobin, diabetes, and cardiovascular risk in nondiabetic adults. N Engl J Med 2010;362:800–811
4. Sarwar N, Aspelund T, Eiriksdottir G, et al. Markers of dysglycaemia and risk of coronary heart disease in people without diabetes: Reykjavik prospective study and systematic review. PLoS Med 2010;7:e1000278
5. Gerstein HC, Miller ME, Byington RP, et al.; Action to Control Cardiovascular Risk in Diabetes Study Group. Effects of intensive glucose lowering in type 2 diabetes. N Engl J Med 2008;358:2545–2559
6. ACCORD Study Group. Nine-year effects of 3.7 years of intensive glycemic control on cardiovascular outcomes. Diabetes Care 2016;39:701–708
7. Gerstein HC, Miller ME, Genuth S, et al.; ACCORD Study Group. Long-term effects of intensive glucose lowering on cardiovascular outcomes. N Engl J Med 2011;364:818–828
8. Gaede P, Lund-Andersen H, Parving HH, Pedersen O. Effect of a multifactorial intervention on mortality in type 2 diabetes. N Engl J Med 2008;358:580–591
9. Fang HJ, Zhou YH, Tian YJ, Du HY, Sun YX, Zhong LY. Effects of intensive glucose lowering in treatment of type 2 diabetes mellitus on cardiovascular outcomes: a meta-analysis of data from 58,160 patients in 13 randomized controlled trials. Int J Cardiol 2016;218:50–58
10. Marso SP, Daniels GH, Brown-Frandsen K, et al.; LEADER Steering Committee; LEADER Trial Investigators. Liraglutide and cardiovascular outcomes in type 2 diabetes. N Engl J Med 2016;375:311–322
11. Zinman B, Wanner C, Lachin JM, et al.; EMPA-REG OUTCOME Investigators. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. N Engl J Med 2015;373:2117–2128
12. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?" Lancet 2005;365:82–93
13. Azoulay L, Suissa S. Sulfonylureas and the risks of cardiovascular events and death: a methodological meta-regression analysis of the observational studies. Diabetes Care 2017;40:706–714
14. Ahmad OS, Morris JA, Mujammami M, et al. A Mendelian randomization study of the effect of type-2 diabetes on coronary heart disease. Nat Commun 2015;6:7060
15. Ross S, Gerstein HC, Eikelboom J, Anand SS, Yusuf S, Paré G. Mendelian randomization analysis supports the causal role of dysglycaemia and diabetes in the risk of coronary artery disease. Eur Heart J 2015;36:1454–1462
16. Merino J, Leong A, Posner DC, et al. Genetically driven hyperglycemia increases risk of coronary artery disease separately from type 2 diabetes. Diabetes Care 2017;40:687–693
17. Huang Y, Cai X, Mai W, Li M, Hu Y. Association between prediabetes and risk of cardiovascular disease and all cause mortality: systematic review and meta-analysis. BMJ 2016;355:i5953
18. Feigin VL, Roth GA, Naghavi M, et al.; Global Burden of Diseases, Injuries and Risk Factors Study 2013 and Stroke Experts Writing Group. Global burden of stroke and risk factors in 188 countries, during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet Neurol 2016;15:913–924
19. Ramachandran A, Ma RC, Snehalatha C. Diabetes in Asia. Lancet 2010;375:408–418
20. Larsson SC, Scott RA, Traylor M, et al.; METASTROKE Collaboration and NINDS Stroke Genetics Network (SiGN). Type 2 diabetes, glucose, insulin, BMI, and ischemic stroke subtypes: Mendelian randomization study. Neurology 2017;89:454–460
21. Wheeler E, Leong A, Liu CT, et al.; EPIC-CVD Consortium; EPIC-InterAct Consortium; Lifelines Cohort Study. Impact of common genetic determinants of hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: a transethnic genome-wide meta-analysis. PLoS Med 2017;14:e1002383
22. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015;12:e1001779
23. Collins R. What makes UK Biobank special? Lancet 2012;379:1173–1174
24. Nikpay M, Goel A, Won HH, et al. A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 2015;47:1121–1130
25. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med 2008;27:1133–1163
26. Tyrrell J, Jones SE, Beaumont R, et al. Height, body mass index, and socioeconomic status: Mendelian randomisation study in UK Biobank. BMJ 2016;352:i582
27. Woodfield R, Grant I, Sudlow CL; UK Biobank Stroke Outcomes Group; UK Biobank Follow-Up and Outcomes Working Group. Accuracy of electronic health record data for identifying stroke cases in large-scale epidemiological studies: a systematic review from the UK Biobank Stroke Outcomes Group. PLoS One 2015;10:e0140533
28. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. Int J Epidemiol 2016;45:1866–1886

29. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44:512–525
30. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. Genet Epidemiol 2016;40:304–314
31. Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res 2014; 42:D1001–D1006
32. Staley JR, Blackshaw J, Kamat MA, et al. PhenoScanner: a database of human genotype-phenotype associations. Bioinformatics 2016;32:3207–3209
33. Wing RR, Lang W, Wadden TA, et al.; Look AHEAD Research Group. Benefits of modest weight loss in improving cardiovascular risk factors in overweight and obese individuals with type 2 diabetes. Diabetes Care 2011;34: 1481–1486
34. Lindström J, Ilanne-Parikka P, Peltonen M, et al.; Finnish Diabetes Prevention Study Group. Sustained reduction in the incidence of type 2 diabetes by lifestyle intervention: follow-up of the Finnish Diabetes Prevention Study. Lancet 2006;368:1673–1679
35. Joseph P, Leong D, McKee M, et al. Reducing the global burden of cardiovascular disease, part 1: the epidemiology and risk factors. Circ Res 2017;121:677–694
36. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. Int J Epidemiol 2013;42:1012–1014
37. Schooling CM, Au Yeung SL. "Selection bias by death" and other ways collider bias may cause the obesity paradox. Epidemiology 2017;28:e16–e17
38. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan N, Thompson J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. Stat Med 2017;36:1783–1802
39. Chen Z, Chen J, Collins R, et al.; China Kadoorie Biobank (CKB) Collaborative Group. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. Int J Epidemiol 2011;40:1652–1666