# Assessment of Causal Direction Between Gut Microbiota–Dependent Metabolites and Cardiometabolic Health: A Bidirectional Mendelian Randomization Analysis

Jinzhu Jia,[1,2] Pan Dou,[3] Meng Gao,[1] Xuejun Kong,[4] Changwei Li,[5] Zhonghua Liu,[6] and Tao Huang[1,7,8]

*Diabetes 2019;68:1747–1755 | https://doi.org/10.2337/db19-0153*

**We examined the causal direction between gut microbiota–dependent metabolite trimethylamine *N*-oxide (TMAO) or its predecessors and cardiometabolic diseases, such as risk of type 2 diabetes mellitus (T2DM), coronary artery disease (CAD), myocardial infarction (MI), stroke, atrial fibrillation (AF), and chronic kidney disease (CKD). We used genetic variants as instruments to test the causal associations. Genetically predicted higher TMAO and carnitine were not associated with higher odds of T2DM, AF, CAD, MI, stroke, and CKD after Bonferroni correction ($P \leq 0.0005$). However, we observed that genetically increased choline showed a suggestive association with higher risk of T2DM (odds ratio 1.84 [95% CI 1.00–3.42] per 10 units, $P = 0.05$). In contrast, genetically predicted higher betaine (0.68 [0.48–0.95] per 10 units, $P = 0.023$) was suggestively associated with a lower risk of T2DM. We observed a suggestive association of genetically increased choline with a lower level of body fat percentage ($\beta \pm$ SE $-0.28 \pm 0.11$, $P = 0.013$) but a higher estimated glomerular filtration rate ($0.10 \pm 0.05$, $P = 0.034$). We further found that T2DM ($0.130 \pm 0.036$, $P < 0.0001$) and CKD ($0.483 \pm 0.168$, $P = 0.004$) were causally associated with higher TMAO levels. Our Mendelian randomization findings support that T2DM and kidney disease increase TMAO levels and that observational evidence for cardiovascular diseases may be due to confounding or reverse causality.**

Gut microbiota has been recently implicated as a novel endocrine organ that plays an important role in the development of cardiovascular disease (CVD) and type 2 diabetes mellitus (T2DM) through modulating blood levels of bioactive metabolites (1,2). There is a growing appreciation that a gut microbiota–dependent metabolite, trimethylamine *N*-oxide (TMAO), formed from ingested choline, betaine, and carnitine in humans (3,4), is predictive of cardiovascular events (1,5). Furthermore, a greater increase in TMAO, choline, and L-carnitine has been implicated in lesser improvements in insulin sensitivity (6), adiposity, and energy metabolism (7). Animal studies have shown that the TMAO meta-organismal pathway enhances the accumulation of cholesterol in macrophages and platelet hyperreactivity (1) and exacerbates impaired glucose tolerance by blocking the hepatic insulin signaling pathway (8), which promotes CVD and diabetes in animal models.

However, the results from human studies are not always consistent (9,10). Although a recent meta-analysis of 19 observational cohorts found that elevated levels of TMAO and its precursors are associated with an increased risk of major adverse cardiovascular events (11), available evidence from observational studies mainly relies on self-reported information and is susceptible to confounding or reverse causation bias. Therefore, the causality of these observations remains unclear. Documentation of such

[1]Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China
[2]Department of Biostatistics, School of Public Health, Peking University, Beijing, China
[3]Department of Clinical Nutrition, Peking University First Hospital, Beijing, China
[4]Martinos Center, Massachusetts General Hospital, Charleston, MA
[5]Department of Epidemiology and Biostatistics, University of Georgia, Athens, GA
[6]Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong, China
[7]Department of Global Health, School of Public Health, Peking University, Beijing, China
[8]Key Laboratory of Molecular Cardiovascular Sciences (Peking University), Ministry of Education, Beijing, China

Corresponding author: Tao Huang, huangtaotao@pku.edu.cn, or Zhonghua Liu, zhhliu@hku.hk

Received 12 February 2019 and accepted 29 May 2019

This article contains Supplementary Data online at http://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/db19-0153/-/DC1.

causality can inform potential lifestyle or drug treatment targets for prevention of cardiometabolic disease.

The Mendelian randomization (MR) approach, which uses genetic variants as an instrumental variable in epidemiological study, has been widely accepted to explore the potential causal effect of exposure on disease (12–16). Such an MR method is analogous to a randomized controlled trial, where genetic alleles are randomly assorted during conception, and is less likely to be affected by confounding or reverse causation (13,17). Therefore, we performed a bidirectional MR analysis to examine the causal direction between the gut microbiota–dependent metabolite TMAO or its predecessors (e.g., choline, betaine, and carnitine) and cardiometabolic diseases (e.g., T2DM, coronary artery disease [CAD], myocardial infarction [MI], stroke, atrial fibrillation [AF], and chronic kidney disease [CKD]) and related traits using summary data from genome-wide association studies (GWAS).

## RESEARCH DESIGN AND METHODS

### Study Design

The MR approach must satisfy three assumptions (18) (Fig. 1): *1*) the genetic variant selected as the instrumental variable is associated with the gut microbiota–dependent metabolite TMAO and its predecessors; *2*) the genetic variant is not associated with any unmeasured confounders of the gut microbiota–dependent metabolites and the cardiometabolic relationship; and *3*) the genetic variant is associated with cardiometabolic events only through gut microbiota–dependent metabolites, not through other pathways.

### Gut Microbiota–Dependent Metabolites and Instrumental Variable Selection

The gut microbiota–dependent metabolites include TMAO, choline, betaine, and carnitine. We searched PubMed for GWAS of the gut microbiota–dependent metabolites and cardiometabolic diseases and identified genetic variants for each metabolite (19) and disease (20–24). For each metabolite, we selected genetic variants (single nucleotide polymorphisms [SNPs]) at thresholds for suggestive genome-wide significance ($P < 5 \times 10^{-5}$) from the published GWAS (19) (Supplementary Table 1). Details on GWAS where we extracted summary-level data are presented in Table 1 and Supplementary Table 1. Contributing studies received ethical approval from their respective institutional review boards. Informed consent was obtained from all participants of contributing studies.

### Cardiometabolic Diseases and Data Sources

For disease outcomes, summary-level data were extracted from the Diabetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium for T2DM ($n$ = 149,821) (20); the Atrial Fibrillation Consortium (AFGen) for AF (21); the Coronary Artery Disease Genomewide Replication and Meta-analysis (CARDIoGRAM) plus the Coronary Artery Disease (C4D) Genetics (CARDIoGRAMplusC4D) Consortium for CAD (60,801 patients and 123,504 control subjects) and MI (43,676 patients and 128,197 control subjects), respectively (22); the National Institute of Neurological Disorders and Stroke Genetics Network (SiGN) and International Stroke Genetics Consortium (ISGC) for stroke (37,792 patients and 397,209 control subjects) (23); and the Chronic Kidney Disease Consortium (CKDGen) ($n$ = 133,814) for CKD (24).

Patients with T2DM were diagnosed according to 1999 World Health Organization criteria of fasting plasma glucose concentration ≥7.0 mmol/L or 2-h plasma glucose concentration ≥11.1 mmol/L, by report of diabetes medication use, or on the basis of medical record review (25). Ascertainment of AF in UK Biobank included samples with one or more of the following codes: *1*) noncancer illness code, self-reported (1471, 1483); *2*) operation code (1524); *3*) diagnoses main/secondary (ICD-10 codes I48, I48.0–4,

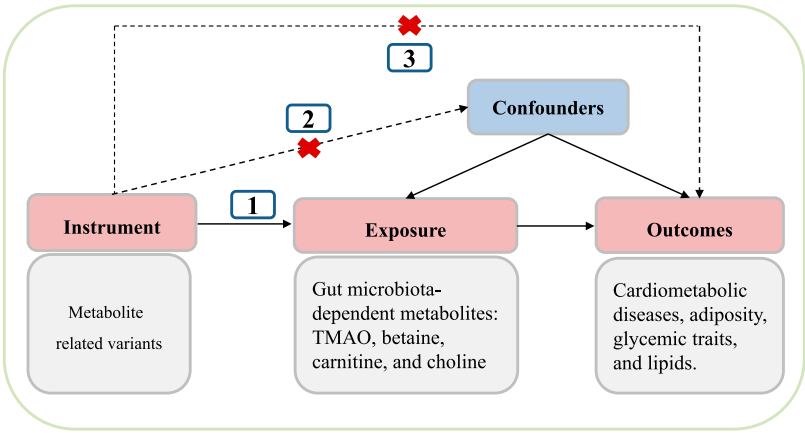


**Figure 1**—Schematic representation of an MR analysis. MR can be used to test the hypothesis that exposure (gut microbiota–related metabolites) causes T2DM, CAD, MI, AF, stroke, and CKD. The three assumptions of MR are as follows: *1*) Genetic variants must be associated with gut microbiota–related metabolites; *2*) genetic variants must not be associated with confounders; and *3*) genetic variants must influence disease outcomes only through gut microbiota–related metabolites, not through other pathways.

**Table 1—Description of modifiable factors or cardiometabolic disease**

| | Consortium or study | Sample size | Population | Year |
|---|---|---|---|---|
| Adiposity | | | | |
| BMI | GIANT | 339,224 | European | 2015 |
| Body fat | GIANT, SAT-VAT, LEPgen, GLGC, MAGIC, DIAGRAM, CARDIoGRAMplusC4D | 100,716 | Transethnic | 2016 |
| Waist-to-hip ratio | GIANT | 224,459 | European | 2015 |
| Glycemic traits | | | | |
| Fasting glucose | MAGIC | 133,010 | European | 2013 |
| Fasting insulin | MAGIC | 133,010 | European | 2013 |
| HOMA-IR | MAGIC | 133,010 | European | 2013 |
| HOMA-B | MAGIC | 133,010 | European | 2013 |
| Fasting proinsulin | DIAGRAM | 16,378 | European | 2011 |
| 2-h glucose | MAGIC | 133,010 | European | 2013 |
| $HbA_{1c}$ | DIAGRAM | 46,368 | European | 2010 |
| Lipids | | | | |
| HDL | GLGC | 188,578 | Transethnic | 2013 |
| LDL | GLGC | 188,578 | Transethnic | 2013 |
| Total cholesterol | GLGC | 188,578 | Transethnic | 2013 |
| Triglycerides | GLGC | 188,578 | Transethnic | 2013 |
| Kidney function | | | | |
| eGFRcreat in T2DM | CKDGen | 133,814 | European | 2016 |
| eGFRcreat | CKDGen | 133,814 | European | 2016 |
| eGFRcys | CKDGen | 133,814 | European | 2016 |
| Other | | | | |
| Leptin | ADIPOGen | 52,140 | Transethnic | 2015 |
| Heart rate | UK Biobank | 265,046 | European | 2016 |
| Diseases* | | | | |
| T2DM | DIAGRAM, AGEN-T2D, SAT2D, MAT2D | 47,979/139,611 | European | 2014 |
| AF | AFGen | 65,446/522,744 | Transethnic | 2018 |
| CAD | CARDIoGRAM and CARDIoGRAMplusC4D | 60,801/123,504 | Transethnic | 2018 |
| MI | CARDIoGRAM and CARDIoGRAMplusC4D | 43,676/128,197 | Transethnic | 2018 |
| Stroke | NINDS SiGN and ISGC | 37,792/397,209 | Transethnic | 2015 |
| CKD | CKDGen | 133,814 | Transethnic | 2016 |

ADIPOGen, Adiponectin Genetics Consortium; AGEN-T2D, Asian Genetic Epidemiology Network Type 2 Diabetes; LEPgen, Leptin Genetic Consortium; MAT2D, Mexican American Type 2 Diabetes Consortium; NINDS, National Institute of Neurological Disorders and Stroke; SAT2D, South Asian Type 2 Diabetes Consortium; SAT-VAT, subcutaneous adipose tissue-abdominal visceral adipose tissue.
*Except for CKD, sample size data are presented as case subjects/control subjects.

I48.9); *4*) underlying (primary/secondary) cause of death (ICD-10 codes I48, I48.0–4, I48.9); *5*) diagnoses main/secondary (ICD-9 code 4273); and *6*) operative procedures main/secondary (OPCS codes K57.1, K62.1–4) (21). CKD was classified in patients with an estimated glomerular filtration rate based on creatinine (eGFRcreat) $<60$ mL/min/1.73 $m^2$ (24). Cases of CAD were determined with a broad definition, including MI, acute coronary syndrome, chronic stable angina, and coronary artery stenosis $>50\%$ (22).

### Cardiometabolic Traits and Data Sources

We searched PubMed for GWAS of cardiometabolic traits. Summary-level data were extracted from the Genetic Investigation of Anthropometric Traits (GIANT) Consortium for adiposity indicators such as BMI (26), body fat (27), and waist-to-hip ratio–adjusted BMI (28); the Meta-Analyses of Glucose and Insulin-Related Traits Consortium (MAGIC) (25,29) for glycemic traits such as fasting glucose, log-transformed fasting insulin, 2-h glucose, log-transformed HOMA of β-cell function (HOMA-B), log-transformed HOMA of insulin resistance (HOMA-IR), log-transformed proinsulin, and hemoglobin $A_{1c}$ ($HbA_{1c}$); the Global Lipids Genetics Consortium (GLGC) ($n = 188,577$) for lipids such as HDL (30), LDL (30), total cholesterol (30), and triglycerides (30); the CKDGen ($n = 133,814$) for kidney function indicators such as eGFRcreat, which was estimated using the four-variable MDRD equation, and eGFR based on cystatin C (eGFRcys), which was estimated as $76.7 \times$ (serum cystatin C) $- 1.19$ (24); and GWAS summary data for leptin (31) and heart rate (32).

### Statistical Analysis

#### *Linkage Disequilibrium Assessment and Pleiotropy Assessment*

To verify that the SNPs selected in this study met assumptions 1 and 2, we examined the genetic association with each metabolite, measured linkage disequilibrium (LD) among all the SNPs for the same metabolite (33), and selected independent genetic variants for each metabolite (13). We chose the variant with the lowest *P* value for association with each metabolite if genetic variants were in

LD; therefore, the SNPs selected did not violate assumption 2. All SNPs for the same metabolite showed strong association ($F$ statistic >21.10, the strength of the instrument), thus meeting assumption 1. We used MR-Egger regression to assess the presence of pleiotropic effects on cardiometabolic outcomes (18). Using the MR-Egger method, the SNP's effect upon each metabolite is plotted against its effect upon outcomes, and an intercept distinct from the origin provides evidence for pleiotropic effects.

### *MR Analysis*

The estimates of the causal effect of each metabolite on outcomes were analyzed using inverse-variance weighting (IVW), which provides a combined estimate of the causal estimate from each SNP. IVW is equivalent to a two-stage least squares or allele score analysis using individual-level data and is hence considered here as conventional MR (34). Results are presented as odds ratios (ORs) (95% CIs) per 10 units of genetically predicted increase in each metabolite. In addition, results for causal effects of diseases on metabolites are presented as $\beta \pm SE$ per each unit higher in log odds of disease. Furthermore, complementary approaches such as simple median method, weighted median method, the mode-based estimate (MBE), and MR-Egger for multiple genetic variants were used to examine causal effect (34–37). Detailed information on these MR methods has been described previously (34–37). Finally, the MR-Egger regression test was used to evaluate the pleiotropic effects (18). Using the MR-Egger regression method, the effect of an instrumental variable on the exposure is plotted against its effect on the outcome, and an intercept distinct from the origin provides evidence for pleiotropic effects. The slope of the MR-Egger regression can provide pleiotropy-corrected causal estimates (18). In the present MR analyses, median weighted, MBE, and MR-Egger methods were considered as sensitivity analyses for MR investigations with multiple genetic variants (34–36). Power calculations for MR were conducted on the basis of mRnd (http://cnsgenomics.com/shiny/mRnd).

Analyses were performed using R 3.2.3 software (R Project for Statistical Computing). The threshold of statistical significance was $P \leq 0.0005$ (0.05/100) after Bonferroni correction. $P \leq 0.05$ but above the Bonferroni-corrected significance threshold was considered as suggestive of evidence for a potential association.

## RESULTS

### Characteristics for Selected SNPs

The characteristics of the selected SNPs for each gut microbiota–dependent metabolite are presented in Table 1 and Supplementary Table 1. To examine assumptions 2 and 3, we tested whether any of the selected SNPs were influenced by LD. We chose the variant with the lowest $P$ value for association with each metabolite if genetic variants were in LD. Our genetic analysis showed that 15.4% of betaine, 17.1% of carnitine, 8.0% of choline, and 9.6% of TMAO were explained by their SNPs (Supplementary Table 2). MR power calculation showed that we had 87%, 84%, 78%, and 81% power to test significant ($P < 0.05$) causal effects (OR 1.3) of betaine, carnitine, choline, and TMAO, respectively, on cardiometabolic events.

### Gut Microbiota–Dependent Metabolites and Cardiometabolic Disease

Genetically predicted higher TMAO was not associated with a higher odds of T2DM (OR 0.96 [95% CI 0.59–1.57], $P$ = 0.863), AF (0.99 [0.79–1.26], $P$ = 0.961), CAD (1.00 [0.70–1.43], $P$ = 0.986), MI (1.08 [0.73–1.60], $P$ = 0.708), stroke (0.94 [0.53–1.65], $P$ = 0.83), and CKD (1.08 [0.59–1.99], $P$ = 0.794) per 10 units after Bonferroni correction (Fig. 2). Likewise, genetically predicted higher carnitine was also not associated with each cardiometabolic disease. However, we observed that genetically increased choline was suggestively associated with higher risk of T2DM (1.84 [1.00–3.42] per 10 units, $P$ = 0.05) but not with cardiovascular events and CKD. In contrast, genetically predicted higher betaine (0.68 [0.48–0.95] per 10 units, $P$ = 0.023) was suggestively associated with a lower risk of T2DM (Fig. 2).

In addition, the associations were consistent in sensitivity analyses that used the simple median and weighted median methods (Supplementary Table 3). The MR-Egger regression test was used to examine the presence of a pleiotropic effect. The intercepts (SE) from MR-Egger regression testing of each outcome were centered at the

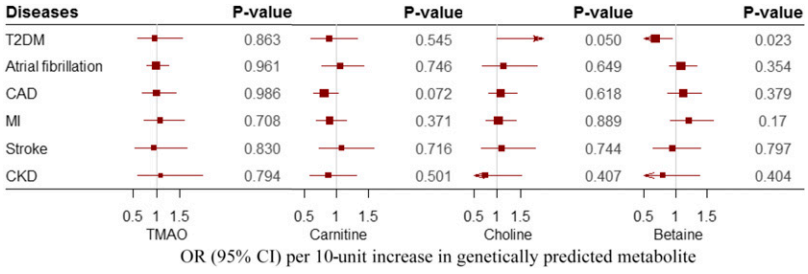


**Figure 2**—MR of gut microbiota–related metabolites and risk of T2DM, CAD, MI, AF, stroke, and CKD.

origin with a CI including the null, suggesting that the results were not influenced by pleiotropy (Supplementary Table 3).

### Gut Microbiota–Dependent Metabolites and Cardiometabolic Traits

Genetically predicted higher TMAO and betaine were not associated with any cardiometabolic factors after Bonferroni correction (Table 2). However, we observed a suggestive association ($P < 0.05$) of genetically increased choline with a lower level of body fat percentage ($\beta \pm SE$ $-0.28 \pm 0.11$, $P = 0.013$) but a higher eGFRcys ($0.10 \pm 0.05$, $P = 0.034$). Additionally, genetically predicted higher carnitine was suggestively associated with a lower level of triglycerides ($-0.17 \pm 0.07$, $P = 0.012$) (Table 2). Results were similar in sensitivity analyses that used the simple median and weighted median methods (Supplementary Table 4). The results were not influenced by directional pleiotropy ($P > 0.05$). We further examined the genetic association of TMAO SNPs with potential confounders, such as alcohol intake and dietary factors (Supplementary Table 5). We did not find that TMAO SNPs were associated with any potential confounders, thus validating assumption 2.

### Causal Effects of Cardiometabolic Diseases on Gut-Dependent Metabolites

We further examined the causal effects of cardiometabolic diseases on gut-dependent metabolites. We found that T2DM was causally associated with lower betaine ($\beta \pm SE$ $-0.111 \pm 0.035$, $P = 0.002$) and higher TMAO levels ($0.13 \pm 0.036$, $P < 0.0001$) per 1-unit higher log odds. CKD was also causally associated with higher TMAO levels ($0.483 \pm 0.168$, $P = 0.004$) per 1-unit higher log odds (Table 3 and Supplementary Table 6).

## DISCUSSION

In the present MR study, we did not observe a significant association of genetically predicted higher gut microbiota–dependent metabolite TMAO and its predecessor with cardiometabolic disease and related traits. However, we observed a suggestive association of genetically increased choline with a higher risk of T2DM. Furthermore, genetically increased choline showed a suggestive association with a lower body fat percentage but a higher level of eGFR, whereas genetically predicted higher carnitine was suggestively associated with a lower level of triglycerides. We further found that T2DM and CKD were causally associated with higher TMAO levels. Our findings support that T2DM and CKD increase TMAO levels and that the previously observed association of TMAO and CVD may be due to confounding or reverse causality.

Gut microbiota alteration has been implicated in the development of diabetes (38). Previous observational findings suggested that higher plasma TMAO is associated with an increased odds of newly diagnosed T2DM (39). Animal studies showed that TMAO exacerbates impaired glucose tolerance and hyperglycemia through causing inflammation or blocking the hepatic insulin signaling pathway in adipose tissue (8). Likewise, a weight loss diet intervention study demonstrated that a greater increase in TMAO is related to lesser improvements in the glycemic traits (6). However, previous studies have yielded inconsistent associations between plasma TMAO and T2DM (39–41). Our MR analysis also did not confirm the causal effect of genetically increased TMAO on risk of T2DM. Interestingly, we found that T2DM increases TMAO levels. In addition, a recent genome-wide scan suggested that human gut microbiota is influenced by a host gene that has been associated with diabetes (42), suggesting that a TMAO-T2DM association may share a genetic basis rather than a causal relationship. We cannot exclude the possibility that the suggestive genome-wide SNP ($P < 5.0 \times 10^{-5}$) used in the MR analysis might introduce a weak instrument bias (19), thus diluting the association. Future investigation for identifying genome-wide significant SNPs for TMAO and its causal role in the development of T2DM is required.

Furthermore, a recent prospective study indicated that a higher intake of phosphatidylcholine, the precursor to the generation of TMAO, is independently associated with an increased risk of T2DM in three U.S. populations (43). Interestingly, our MR results showed a suggestive association of genetically increased choline with a higher risk of T2DM. Our findings were supported by a previous weight loss diet intervention study that demonstrated that greater decreases in choline and L-carnitine are significantly associated with greater improvements in fasting insulin concentrations and insulin resistance (6), underscoring the importance of changes in choline and L-carnitine in improving insulin sensitivity (6).

Our findings did not corroborate the results from a previous meta-analysis of data from 19 prospective studies, which provided quantitative pooled estimates of the associations of circulating TMAO level with the incidence of cardiovascular events (11). It was estimated that participants with high TMAO levels had a 62% increased risk for the development of cardiovascular events compared with those with low TMAO levels (11). However, our MR results showed that genetically increased TMAO is not associated with a higher risk of cardiovascular events, such as CAD, MI, stroke, and AF, suggesting that observational evidence might result from confounding and reverse causality. Although we measured LD among all selected SNPs using northern Europeans from Utah samples from the 1000 Genomes Project (33), we cannot exclude the possibility that our results might be affected by unmeasured confounders. Furthermore, the MR-Egger regression showed that our results are not influenced by pleiotropy (18). Nevertheless, it is possible that the results reflect a shared genetic basis between gut microbiota–related metabolites and CVD rather than a causal relationship.

Despite strong observational evidence for an association of predecessors of TMAO, such as carnitine, choline, or betaine, with risk of CAD, MI, stroke, and AF, our results

**Table 2—MR analysis of gut microbiota–dependent metabolite and cardiometabolic traits**

| | Betaine | | Carnitine | | Choline | | TMAO | |
|---|---|---|---|---|---|---|---|---|
| | β ± SE | *P* value | β ± SE | *P* value | β ± SE | *P* value | β ± SE | *P* value |
| Adiposity | | | | | | | | |
| BMI (kg/m$^2$) | 0.02 ± 0.06 | 0.686 | −0.06 ± 0.05 | 0.246 | 0.04 ± 0.08 | 0.605 | 0.06 ± 0.07 | 0.411 |
| Waist-to-hip ratio | −0.05 ± 0.08 | 0.509 | −0.07 ± 0.06 | 0.294 | −0.04 ± 0.09 | 0.683 | 0.01 ± 0.09 | 0.988 |
| Body fat (%) | 0.07 ± 0.07 | 0.312 | 0.01 ± 0.10 | 0.997 | −0.28 ± 0.11 | 0.013 | 0.13 ± 0.08 | 0.112 |
| Glycemic traits | | | | | | | | |
| Fasting glucose (mmol/L) | 0.06 ± 0.06 | 0.325 | −0.04 ± 0.08 | 0.608 | 0.16 ± 0.09 | 0.065 | −0.02 ± 0.08 | 0.775 |
| Fasting insulin (pmol/L) | 0.10 ± 0.07 | 0.159 | −0.02 ± 0.06 | 0.716 | −0.06 ± 0.08 | 0.439 | −0.02 ± 0.08 | 0.768 |
| $HbA_{1c}$ | | 0.43 | | 0.942 | | 0.818 | | 0.866 |
| % | 0.05 ± 0.06 | | 0.01 ± 0.05 | | −0.02 ± 0.07 | | −0.01 ± 0.05 | |
| mmol/mol | 0.31 ± 0.37 | | 0.06 ± 0.31 | | −0.12 ± 0.43 | | 0.06 ± 0.31 | |
| HOMA-B | 0.07 ± 0.06 | 0.227 | 0.03 ± 0.06 | 0.527 | −0.05 ± 0.06 | 0.485 | 0.02 ± 0.09 | 0.858 |
| HOMA-IR | 0.09 ± 0.07 | 0.178 | 0.01 ± 0.06 | 0.879 | 0.01 ± 0.08 | 0.974 | −0.02 ± 0.08 | 0.772 |
| Fasting proinsulin (pmol/L) | −0.03 ± 0.11 | 0.791 | 0.09 ± 0.11 | 0.436 | −0.19 ± 0.18 | 0.291 | −0.01 ± 0.12 | 0.936 |
| 2-h glucose (mmol/L) | 0.12 ± 0.32 | 0.708 | 0.35 ± 0.29 | 0.216 | 0.32 ± 0.39 | 0.415 | 0.31 ± 0.39 | 0.438 |
| Lipids (mg/dL) | | | | | | | | |
| HDL | 0.01 ± 0.10 | 0.953 | −0.01 ± 0.07 | 0.919 | 0.10 ± 0.10 | 0.356 | 0.03 ± 0.10 | 0.744 |
| LDL | 0.01 ± 0.12 | 0.94 | −0.07 ± 0.08 | 0.347 | 0.22 ± 0.13 | 0.094 | 0.05 ± 0.11 | 0.608 |
| Total cholesterol | −0.01 ± 0.12 | 0.906 | −0.09 ± 0.07 | 0.242 | 0.10 ± 0.13 | 0.433 | 0.05 ± 0.10 | 0.652 |
| Triglycerides | 0.01 ± 0.08 | 0.934 | −0.17 ± 0.07 | 0.012 | 0.08 ± 0.11 | 0.476 | 0.05 ± 0.10 | 0.636 |
| Kidney function (mL/min/1.73 m$^2$) | | | | | | | | |
| eGFRcreat in T2DM | 0.05 ± 0.06 | 0.367 | −0.03 ± 0.06 | 0.636 | 0.07 ± 0.09 | 0.475 | 0.05 ± 0.09 | 0.552 |
| eGFRcreat | 0.01 ± 0.02 | 0.792 | 0.01 ± 0.01 | 0.536 | −0.02 ± 0.02 | 0.38 | −0.02 ± 0.01 | 0.186 |
| eGFRcys | −0.04 ± 0.03 | 0.307 | 0.01 ± 0.03 | 0.642 | 0.10 ± 0.05 | 0.034 | −0.01 ± 0.04 | 0.844 |
| Other | | | | | | | | |
| Leptin (μg/mL) | −0.08 ± 0.08 | 0.272 | 0.03 ± 0.07 | 0.644 | −0.01 ± 0.13 | 0.912 | −0.07 ± 0.12 | 0.535 |
| Heart rate (beats/min) | −0.11 ± 0.31 | 0.729 | −0.15 ± 0.48 | 0.75 | −1.03 ± 0.87 | 0.237 | −0.14 ± 0.45 | 0.757 |

Causal effects were estimated using instrumental variables. The β-coefficient means each SD change in outcomes per 10-unit increase in genetically predicted gut microbiota–dependent metabolite.

disagree with an earlier meta-analysis of observational studies showing a positive association of L-carnitine, choline, or betaine with risk of cardiovascular events (11). These null findings suggest that our findings do not support a causal role of predecessors of TMAO in the development of CVD, but we have to acknowledge that more variants, especially suggestive genome-wide SNPs, may lead to greater pleiotropy and dilute the association in our analysis. However, we found that genetically predicted higher carnitine was suggestively associated with a lower level of triglycerides. The potential mechanism may be that dietary carnitine is metabolized in the liver by intestinal bacteria to produce TMAO by the enzyme flavin monooxygenase 3 (3,4), which is reported to be a key integrator of hepatic cholesterol and lipid metabolism (1,5).

Interestingly, an observational study has reported that TMAO, cleared by the kidney, is elevated in patients with impaired renal function (44) and portends poorer long-term survival (45), suggesting that circulating TMAO might play a role in kidney dysfunction (9,44–46). In the present MR study, our findings did not support a causal role of TMAO in kidney dysfunction. Surprisingly, we found that genetically increased choline showed a suggestive association with a higher level of eGFR. Importantly, we demonstrate that kidney disease increases TMAO levels. Given that the gut microbiota–mediated metabolism of choline and carnitine each ultimately produces TMAO (3,4), we speculate that there is a threshold in the observed association between choline and kidney function and that the optimal level of choline may be beneficial for improving kidney function. However, chronic high dietary exposure to choline or carnitine that increases TMAO appears to be toxic and directly contributes to progressive renal fibrosis and dysfunction (45); further investigation on the mechanism is thus required.

This study has several strengths. First, we have systematically assessed for the first time in our knowledge the causal role of gut microbiota–dependent metabolites in the development of T2DM, CAD, MI, stroke, AF, and CKD. Second, the present MR analysis used summary-level data from large GWAS, which might avoid bias from reverse causation and reduce confounding. Furthermore, the large sample size from GWAS summary statistics has sufficient power for reliable and lifelong causal estimation. Third, the consistent causal estimation across five methods, such as the conventional IVW, simple median, weighted median, MBE, and MR-Egger methods, suggests robustness of our findings.

However, several limitations merit consideration. First, we assumed that the associations between continuous metabolites and outcomes are linear, but our findings might suggest that there is a threshold in observed association. Therefore, further investigation is warranted on the causal role of gut microbiota–dependent metabolites in the development of cardiometabolic disease. Second, we used the suggestive genome-wide SNPs for the

**Table 3—MR analysis of cardiometabolic diseases on gut microbiota–dependent metabolites**

| | Betaine | | Carnitine | | Choline | | TMAO | |
|---|---|---|---|---|---|---|---|---|
| Disease | β ± SE | *P* value | β ± SE | *P* value | β ± SE | *P* value | β ± SE | *P* value |
| T2DM | −0.111 ± 0.035 | 0.002 | 0.054 ± 0.036 | 0.133 | 0.001 ± 0.036 | 0.969 | 0.13 ± 0.036 | <0.0001 |
| AF | −0.069 ± 0.054 | 0.203 | −0.076 ± 0.051 | 0.14 | −0.036 ± 0.057 | 0.522 | −0.004 ± 0.051 | 0.942 |
| CAD | 0.039 ± 0.06 | 0.513 | 0.024 ± 0.06 | 0.686 | 0.062 ± 0.077 | 0.42 | 0.005 ± 0.073 | 0.944 |
| MI | −0.004 ± 0.059 | 0.947 | 0.007 ± 0.059 | 0.908 | 0.083 ± 0.076 | 0.277 | 0.018 ± 0.068 | 0.793 |
| Stroke | 0.025 ± 0.067 | 0.707 | 0.128 ± 0.068 | 0.059 | 0.056 ± 0.086 | 0.515 | 0.044 ± 0.067 | 0.51 |
| CKD | 0.085 ± 0.123 | 0.49 | 0.257 ± 0.149 | 0.085 | 0.039 ± 0.144 | 0.787 | 0.483 ± 0.168 | 0.004 |

The β-coefficient means each SD change in outcomes per 10-unit increase in genetically predicted gut microbiota–dependent metabolite.

metabolites; thus, we cannot exclude the possibility that our findings might have been affected by weak instrument bias, which depends on the strength of the genetic instrument through the *F* statistic. Third, we do not have diet or lifestyle information; therefore, we cannot test whether the instrumental variables are associated with these factors. However, we used diet/lifestyle information from UK Biobank and demonstrated that the genetic variants are not associated with confounders. Fourth, we cannot exclude the possible diet-gene or gene-environment interactions on outcomes, which may influence the observed results. Finally, completely ruling out an alternative direct causal pathway is a challenge for all MR analyses, particularly for metabolites determined by both gut microbiota and multiple genetic variants.

In summary, our findings from an MR approach support that TMAO and kidney disease increase TMAO levels and that observational evidence for gut microbiota–dependent metabolites and CVDs may be due to confounding or reverse causality. The human genome influences both gut microbiota and chronic disease; therefore, we cannot exclude the possibility that such an association may have a genetic basis rather than a causal relationship. A large-scale genome-wide scan for genetic variants of gut microbiota–related metabolites and further investigation in understanding the potential role of gut microbiota–dependent metabolites in the development of cardiometabolic disease are required.

**Acknowledgments.** The authors thank Dr. Robert Gerszten (Division of Cardiovascular Medicine, Beth Israel Deaconess Medical Center, Boston, MA), Dr. Ming-Huei Chen (Population Sciences Branch, National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA), and Dr. Qiong Yang (Department of Biostatistics, Boston University School of Public Health, Boston, MA) for sharing the GWAS summary statistics of gut-dependent metabolites.

**Funding.** The study was supported by Peking University Start-up grant BMU2018YJ002, High-performance Computing Platform of Peking University, and Beijing Technology and Business University grant 88442Y0033.

The funding organizations had no role in the preparation of the manuscript.

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

**Author Contributions.** J.J., P.D., M.G., X.K., C.L., Z.L., and T.H. contributed to the statistical analysis, performed the data analysis, wrote the manuscript, critically reviewed the manuscript during the writing process, and approved the final version to be published. J.J., P.D., and T.H. designed the research. J.J., Z.L., and T.H. are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

**Data and Resource Availability.** All data used in the current study were obtained from GWAS summary statistics, which were publicly released by genetic consortia.

## References

1. Wang Z, Klipfell E, Bennett BJ, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. Nature 2011;472:57–63

2. Forslund K, Hildebrand F, Nielsen T, et al.; MetaHIT Consortium. Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. Nature 2015;528:262–266

3. Tang WH, Wang Z, Levison BS, et al. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. N Engl J Med 2013;368:1575–1584

4. Bennett BJ, de Aguiar Vallim TQ, Wang Z, et al. Trimethylamine-N-oxide, a metabolite associated with atherosclerosis, exhibits complex genetic and dietary regulation. Cell Metab 2013;17:49–60

5. LeLeiko RM, Vaccari CS, Sola S, et al. Usefulness of elevations in serum choline and free F2)-isoprostane to predict 30-day cardiovascular outcomes in patients with acute coronary syndrome. Am J Cardiol 2009;104:638–643

6. Heianza Y, Sun D, Li X, et al. Gut microbiota metabolites, amino acid metabolites and improvements in insulin sensitivity and glucose metabolism: the POUNDS Lost trial. Gut 2019;68:263–270

7. Heianza Y, Sun D, Smith SR, Bray GA, Sacks FM, Qi L. Changes in gut microbiota-related metabolites and long-term successful weight loss in response to weight-loss diets: the POUNDS Lost trial. Diabetes Care 2018;41:413–419

8. Gao X, Liu X, Xu J, Xue C, Xue Y, Wang Y. Dietary trimethylamine N-oxide exacerbates impaired glucose tolerance in mice fed a high fat diet. J Biosci Bioeng 2014;118:476–481

9. Kaysen GA, Johansen KL, Chertow GM, et al. Associations of trimethylamine N-oxide with nutritional and inflammatory biomarkers and cardiovascular outcomes in patients new to dialysis. J Ren Nutr 2015;25:351–356

10. Shafi T, Powe NR, Meyer TW, et al. Trimethylamine N-oxide and cardiovascular events in hemodialysis patients. J Am Soc Nephrol 2017;28:321–331

11. Heianza Y, Ma W, Manson JE, Rexrode KM, Qi L. Gut microbiota metabolites and risk of major adverse cardiovascular disease events and death: a systematic review and meta-analysis of prospective studies. J Am Heart Assoc 2017;6:e004947

12. Geng T, Smith CE, Li C, Huang T. Childhood BMI and adult type 2 diabetes, coronary artery diseases, chronic kidney disease, and cardiometabolic traits: a Mendelian randomization analysis. Diabetes Care 2018;41:1089–1096

13. 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

14. Ding M, Huang T, Bergholdt HK, Nordestgaard BG, Ellervik C, Qi L; CHARGE Consortium. Dairy consumption, systolic blood pressure, and risk of hypertension: Mendelian randomization study. BMJ 2017;356:j1000

15. Huang T, Ren J, Huang J, Li D. Association of homocysteine with type 2 diabetes: a meta-analysis implementing Mendelian randomization approach. BMC Genomics 2013;14:867

16. Benn M, Nordestgaard BG. From genome-wide association studies to Mendelian randomization: novel opportunities for understanding cardiovascular disease causality, pathogenesis, prevention, and treatment. Cardiovasc Res 2018;114:1192–1208

17. Didelez V, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. Stat Methods Med Res 2007;16:309–330

18. 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

19. Rhee EP, Ho JE, Chen MH, et al. A genome-wide association study of the human metabolome in a community-based cohort. Cell Metab 2013;18:130–143

20. Mahajan A, Go MJ, Zhang W, et al.; DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium; Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium; Mexican American Type 2 Diabetes (MAT2D) Consortium; Type 2 Diabetes Genetic Exploration by Next-generation sequencing in multi-Ethnic Samples (T2D-GENES) Consortium. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. Nat Genet 2014;46:234–244

21. Roselli C, Chaffin MD, Weng LC, et al. Multi-ethnic genome-wide association study for atrial fibrillation. Nat Genet 2018;50:1225–1233

22. Nikpay M, Goel A, Won HH, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 2015;47:1121–1130

23. NINDS Stroke Genetics Network (SiGN); International Stroke Genetics Consortium (ISGC). Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study [published correction appears in Lancet Neurol 2016;15:241]. Lancet Neurol 2016;15:174–184
24. Pattaro C, Teumer A, Gorski M, et al.; ICBP Consortium; AGEN Consortium; CARDIOGRAM; CHARGe-Heart Failure Group; ECHOGen Consortium. Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. Nat Commun 2016;7:10023
25. Scott RA, Lagou V, Welch RP, et al.; DIAbetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. Nat Genet 2012;44:991–1005
26. Locke AE, Kahali B, Berndt SI, et al.; LifeLines Cohort Study; ADIPOGen Consortium; AGEN-BMI Working Group; CARDIOGRAMplusC4D Consortium; CKDGen Consortium; GLGC; ICBP; MAGIC Investigators; MuTHER Consortium; MIGen Consortium; PAGE Consortium; ReproGen Consortium; GENIE Consortium; International Endogene Consortium. Genetic studies of body mass index yield new insights for obesity biology. Nature 2015;518:197–206
27. Lu Y, Day FR, Gustafsson S, et al. New loci for body fat percentage reveal link between adiposity and cardiometabolic disease risk. Nat Commun 2016;7:10495
28. Shungin D, Winkler TW, Croteau-Chonka DC, et al.; ADIPOGen Consortium; CARDIOGRAMplusC4D Consortium; CKDGen Consortium; GEFOS Consortium; GENIE Consortium; GLGC; ICBP; International Endogene Consortium; LifeLines Cohort Study; MAGIC Investigators; MuTHER Consortium; PAGE Consortium; ReproGen Consortium. New genetic loci link adipose and insulin biology to body fat distribution. Nature 2015;518:187–196
29. Dupuis J, Langenberg C, Prokopenko I, et al.; DIAGRAM Consortium; GIANT Consortium; Global BPgen Consortium; Anders Hamsten on behalf of Procardis Consortium; MAGIC Investigators. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet 2010;42:105–116
30. Willer CJ, Schmidt EM, Sengupta S, et al.; Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. Nat Genet 2013;45:1274–1283
31. Kilpeläinen TO, Carli JFM, Skowronski AA, et al. Genome-wide meta-analysis uncovers novel loci influencing circulating leptin levels. Nat Commun 2016;7:10494
32. Eppinga RN, Hagemeijer Y, Burgess S, et al. Identification of genomic loci associated with resting heart rate and shared genetic predictors with all-cause mortality. Nat Genet 2016;48:1557–1563
33. Abecasis GR, Auton A, Brooks LD, et al.; 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. Nature 2012;491:56–65
34. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013;37:658–665
35. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol 2017;46:1734–1739
36. 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
37. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. Int J Epidemiol 2017;46:1985–1998
38. Moon JY, Zolnik CP, Wang Z, et al. Gut microbiota and plasma metabolites associated with diabetes in women with, or at high risk for, HIV infection. EBioMedicine 2018;37:392–400
39. Shan Z, Sun T, Huang H, et al. Association between microbiota-dependent metabolite trimethylamine-*N*-oxide and type 2 diabetes. Am J Clin Nutr 2017;106:888–894
40. Qiu G, Zheng Y, Wang H, et al. Plasma metabolomics identified novel metabolites associated with risk of type 2 diabetes in two prospective cohorts of Chinese adults. Int J Epidemiol 2016;45:1507–1516
41. Svingen GF, Schartum-Hansen H, Pedersen ER, et al. Prospective associations of systemic and urinary choline metabolites with incident type 2 diabetes. Clin Chem 2016;62:755–765
42. Wang J, Thingholm LB, Skiecevičienė J, et al. Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota. Nat Genet 2016;48:1396–1406
43. Li Y, Wang DD, Chiuve SE, et al. Dietary phosphatidylcholine intake and type 2 diabetes in men and women. Diabetes Care 2015;38:e13–e14
44. Bain MA, Faull R, Fornasini G, Milne RW, Evans AM. Accumulation of trimethylamine and trimethylamine-N-oxide in end-stage renal disease patients undergoing haemodialysis. Nephrol Dial Transplant 2006;21:1300–1304
45. Tang WHW, Wang Z, Kennedy DJ, et al. Gut microbiota-dependent trimethylamine N-oxide (TMAO) pathway contributes to both development of renal insufficiency and mortality risk in chronic kidney disease. Circ Res 2015;116:448–455
46. Stubbs JR, House JA, Ocque AJ, et al. Serum trimethylamine-N-oxide is elevated in CKD and correlates with coronary atherosclerosis burden. J Am Soc Nephrol 2016;27:305–313