iMeta WILEY

Unraveling interindividual variation of trimethylamine *N*-oxide and its precursors at the population level

Sergio Andreu-Sánchez[1,2] | Shahzad Ahmad[3,4] | Alexander Kurilshikov[1] |
Marian Beekman[5] | Mohsen Ghanbari[3] | Martijn van Faassen[6] |
Inge C. L. van den Munckhof[7] | Marinka Steur[3] | Amy Harms[4] |
Thomas Hankemeier[4] | M. Arfan Ikram[3] | Maryam Kavousi[3] |
Trudy Voortman[3] | Robert Kraaij[8] | Mihai G. Netea[7] |
Joost H. W. Rutten[7] | Niels P. Riksen[7] | Alexandra Zhernakova[1] |
Folkert Kuipers[2,6,9] | P. Eline Slagboom[5] | Cornelia M. van Duijn[10] |
Jingyuan Fu[1,2] | Dina Vojinovic[3,5]

[1]Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

[2]Department of Pediatrics, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

[3]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands

[4]Metabolomics & Analytics Centre, Leiden Academic Center for Drug Research, Leiden University, Leiden, The Netherlands

[5]Molecular Epidemiology, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands

[6]Department of Laboratory Medicine, University Medical Center Groningen, University of Groningen, Groningen, The Netherland

[7]Department of Internal Medicine and Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, Nijmegen, The Netherlands

Abstract

Trimethylamine *N*-oxide (TMAO) is a circulating microbiome-derived metabolite implicated in the development of atherosclerosis and cardiovascular disease (CVD). We investigated whether plasma levels of TMAO, its precursors (betaine, carnitine, deoxycarnitine, choline), and TMAO-to-precursor ratios are associated with clinical outcomes, including CVD and mortality. This was followed by an in-depth analysis of their genetic, gut microbial, and dietary determinants. The analyses were conducted in five Dutch prospective cohort studies including 7834 individuals. To further investigate association results, Mendelian Randomization (MR) was also explored. We found only plasma choline levels (hazard ratio [HR] 1.17, [95% CI 1.07; 1.28]) and not TMAO to be associated with CVD risk. Our association analyses uncovered 10 genome-wide significant loci, including novel genomic regions for betaine (6p21.1, 6q25.3), choline (2q34, 5q31.1), and deoxycarnitine (10q21.2, 11p14.2) comprising several metabolic gene associations, for example, *CPS1* or *PEMT*. Furthermore, our analyses uncovered 68 gut microbiota associations, mainly related to TMAO-to-precursors ratios and the Ruminococcaceae family, and 16 associations of food groups and metabolites including fish-TMAO, meat-carnitine, and plant-based food-betaine associations. No significant association was identified by the MR approach. Our analyses provide novel insights into the TMAO pathway, its determinants, and pathophysiological impact on the general population.

Jingyuan Fu and Dina Vojinovic contributed equally to this article and share senior authorship.

----

[8]Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands

[9]European Institute for the Biology of Ageing, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

[10]Nuffield Department of Population Health, University of Oxford, Oxford, UK

Correspondence
Dina Vojinovic
Email: dina.vojinovic@iqvia.com

Jingyuan Fu
Email: j.fu@umcg.nl

Highlights
- Exploration of microbiome-related metabolites (trimethylamine *N*-oxide [TMAO], choline, betaine, L-carnitine, and deoxycarnitine) in 7834 participants from five population cohorts.
- Cardiovascular risk was associated with elevated choline concentrations, but not with TMAO concentrations.
- Characterization of the genetic architecture behind metabolite concentration variability.
- Identification of gut microbial taxonomic abundance associated with metabolite's plasma concentration levels.
- Fish intake is the major dietary driver of TMAO concentrations, and betaine is related to grains and vegetable intake.

INTRODUCTION

There is a growing interest in the role of gut microbiome-related metabolites in cardiovascular disease (CVD) [1, 2]. Trimethylamine *N*-oxide (TMAO), in particular, has received a lot of attention as a potential promoter of CVD and atherosclerosis [1]. Elevated fasting plasma levels of TMAO have been associated with increased risk for CVD and mortality independently of traditional risk factors in clinical studies [3–8]. However, investigations on TMAO have mainly been conducted in individuals with a high risk of CVD, existing disease, or multimorbidity while studies on the general population are comparatively scarcer [9–12]. Proposed mechanisms through which TMAO may promote the development of atherosclerosis and CVD include vascular inflammation, activation of platelets, disturbance of bile acid metabolism, and inhibition of reverse cholesterol transport [13]. Yet, the actual mode of action in disease development may be context-dependent.

Variation in circulating levels of TMAO is driven by a complex interplay of multiple determinants, such as host genetics, gut microbiome, diet, and kidney function [14]. TMAO can be acquired directly from the diet from fish but it is mainly produced by gut microbiota from the dietary precursors such as choline, L-carnitine, the carnitine-derived metabolite, deoxycarnitine (also known as γ-butyrobetaine), and betaine [15]. Gut microbiota converts these ubiquitous dietary components into trimethylamine (TMA) which is subsequently absorbed from the intestine, transported to the liver, and oxidized into TMAO by hepatic flavin monooxygenases (FMOs), followed by its distribution to different tissues or kidney clearance [16]. However, the extent to which consumption of different animal-based foods may affect plasma TMAO levels is debated [17–19], and TMA producers in the human gut microbiome have not been well-characterized. Association studies point to several genera; however, published results are, in most cases, heterogeneous among studies [15, 20–22]. At the same time, human genetic variation also contributes to TMAO variability. Rare genetic mutations in the *FMO* type 3 gene (*FMO3*) have been shown to affect the oxidation of TMA to TMAO [23]. However, the role of common genetic variation in TMAO homeostasis remains to be elucidated. As identifying potential drivers for alterations in circulating TMAO levels could have preventive and therapeutic implications for CVD [24], a number of studies have explored determinants of TMAO [15, 21]. However, these cross-sectional multiomics studies lack sample sizes and are often limited to exploring TMAO variability while overlooking its precursors. The physiological impact of choline and betaine associations with CVD risk remains controversial [25–28]. Importantly, the major sources of intraindividual variability of these metabolites, together with carnitine and deoxycarnitine, remain unknown.

In addition, CVD risk has often been regarded as a sex-related disease, with clear prevalence differences between males and females [29]. However, sex differences in the effect of TMAO metabolites on CVD risk remain understudied.

The objectives of this study are twofold: First, we aim to explore the relationships between plasma levels of TMAO and its precursors (betaine, carnitine, deoxycarnitine, choline) and various clinical outcomes, encompassing

CVD and mortality. In contrast to much of the existing research conducted within clinical cohorts, our approach involves the utilization of population data derived from 7834 participants enrolled in five prospective Dutch cohorts. By doing so, we aim to shed light on whether elevated concentrations of these metabolites pose a risk to the general population. Second, we attempt to unravel the sources of interindividual variability in these metabolites within the general population. To achieve this, we harness host genetic information, gut microbiome composition, and dietary patterns obtained from distinct subsets of participants within the same five prospective studies.

## RESULTS

Characteristics of the study population

An overview of the study design is depicted in Figure 1. Our study population included 7834 participants from five Dutch prospective cohort studies including the Rotterdam Study I-4, Rotterdam Study III-2, Leiden Longevity Study (LLS), LifeLines-DEEP (LLD), and 300-Obese cohort (300-OB). A description of the contributing cohorts is provided in Supplementary Material and descriptive characteristics of the study participants are shown in Table 1. The mean age of study participants ranged from 43.4 years (standard deviation [sd] = 14.2) in the LLD to 75.1 years (sd = 6.1) in the Rotterdam Study I-4. The sex ratio of participants was roughly balanced, with slightly more females in most of the participating studies (up to 58%), with the exception of 300-OB in which the majority of participants were males (55%). Detailed information on clinical outcomes such as CVD and mortality, host genetics, gut microbiome composition, and diet was available (Figure 1).

Incident clinical outcomes are associated with TMAO precursors but not with TMA

Plasma levels of TMAO and its precursors betaine, carnitine, deoxycarnitine, and choline were measured in all participating cohorts (Supporting Information S2: Table S1). We observed weak correlations between TMAO and its precursors ($|r| \leq 0.20$) and moderate correlations between precursors ($0.30 \leq |r| \leq 0.44$) (Supporting Information S1: Figure S1).

To assess the relationship between metabolites and incident clinical outcomes including CVD and mortality, Cox proportional hazards regression models with age as a time scale were used. In total, 571 incident CVD events
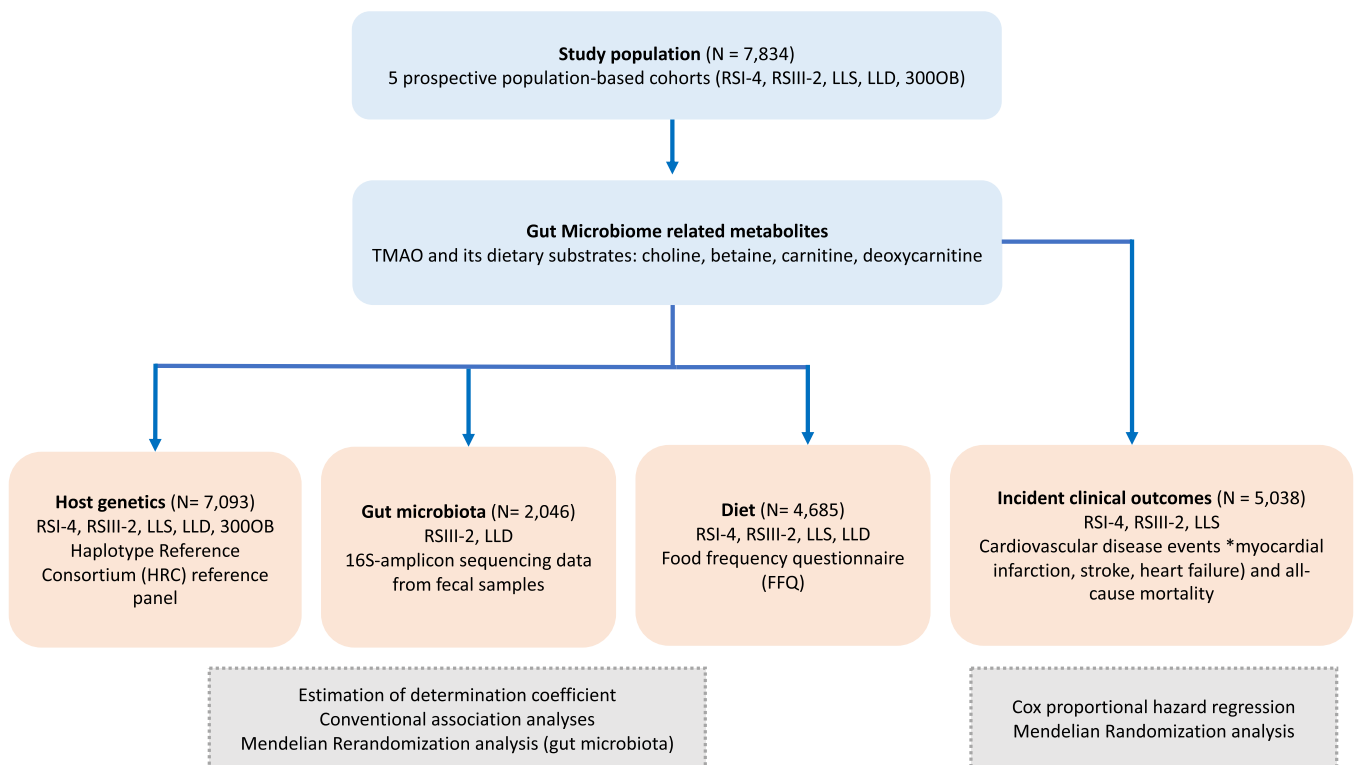
FIGURE 1 Overview of study design and performed analyses. The study population included participants from Rotterdam Study I-4 (RSI-4), Rotterdam Study III-2 (RSIII-2), Leiden Longevity Study (LLS), LifeLines-DEEP (LLD), and 300-Obese cohort (300-OB).

TABLE 1 Descriptive characteristics of the study population.

	Rotterdam Study I-4	Rotterdam Study III-2	LLS	LLD	300-OB
N	2556	1377	2158	1650	302
Age (years), mean (sd)	75.14 (6.08)	62.68 (5.82)	59.12 (6.71)	43.94 (14.15)	67.05 (5.39)
Women, N (%)	1486 (58.14)	748 (54.32)	1208 (55.98)	833 (57.52)	167 (55.29)
Smoking, N (%)	301 (12.05)	238 (17.32)	241 (11.17)	289 (19.63)	26 (8.63)
Diabetes, N (%)	346 (13.61)	123 (8.94)	90 (4.17)	28 (1.7)	37 (12.25)
Hypertension, N (%)	2182 (85.60)	752 (54.77)	159 (7.37)	321 (19.47%)	175 (57.94)
LDLC (mmol/L), mean (sd)	1.63 (0.44)	1.76 (0.51)	1.60 (0.47)	3.14 (0.91)	4.13 (0.96)
HDLC (mmol/L), mean (sd)	1.40 (0.28)	1.37 (0.35)	1.49 (0.33)	1.52 (0.40)	1.33 (0.31)
Serum triglycerides (mmol/L), mean (sd)	1.38 (0.57)	1.33 (0.62)	1.55 (0.82)	1.16 (0.86)	1.83 (1.02)
Lipid-lowering medication, N (%)	583 (22.90)	362 (26.51)	207 (9.59)	29 (2.55)	83 (27.48)
BMI (kg/m^2), mean (sd)	27.43 (4.13)	27.44 (4.50)	25.43 (3.57)	25.30 (4.22)	30.73 (3.48)
eGFRa[a]	79.41 (13.10)	95.31 (9.29)	94.88 (11.60)	73.13 (12.6)	80.4 (15.68)
Incident cardiovascular disease	544	9	85	23	—
Follow-up time cardiovascular disease, mean (years)	8.62 (3.49)	2.48	10.7 (1.77)	—	—
Incident mortality events	1295	37	209	—	—
Follow-up time mortality, mean (years)	10.15 (3.88)	2.48 (1.17)	12.74 (2.22)	—	—

Abbreviations: 300-OB, 300-Obese cohort; BMI, body mass index; eGFR, estimated glomerular filtration rate; HDLC, high-density lipoproteins; LDLC, cholesterol in low-density lipoproteins; LLD, LifeLines-DEEP; LLS, Leiden Longevity Study; sd, standard deviation.

[a]eGFR—calculated using the Chronic Kidney Disease Epidemiology Collaboration equation (CKD-EPI).

and 1440 mortality events were observed among up to 5011 participants with available covariates from Rotterdam Study I-4, Rotterdam Study III-2, and LLS (Supporting Information S2: Tables S2 and S3). A statistically significant association was observed between higher levels of choline and risk of CVD (hazard ratio [HR] 1.17, [95% CI 1.07; 1.28]) after adjusting for sex, body mass index (BMI), hypertension, diabetes, cholesterol in low-density lipoproteins (LDLC) and high-density lipoproteins (HDLC), serum triglycerides (TGs), use of lipid-lowering medication, current smoking, fasting status (if appropriate) (model 1), and multiple testing. The association did not change after further adjustment for the estimated glomerular filtration rate (eGFR) (model 2). The direction of effect size was concordant among the cohorts and no differences were observed between men and women (Supporting Information S2: Table S2). Additionally, a nominal significant association was observed between betaine and risk of CVD (HR 1.10; 95% CI 1.00; 1.21). This association was driven by a strong association in men (HR 1.23; 95% CI 1.07; 1.43). TMAO was associated neither with CVD risk nor mortality. Additional analyses were repeated using concentration quartiles, identifying no significant TMAO association.

To assess the potential causal effect, we performed a Mendelian Randomization (MR) analysis. The results did not provide evidence for any causal effect of TMAO or its precursor metabolites on CVD (Supporting Information S2: Table S4).

Drivers of variation in gut microbiome-related metabolite levels

We next investigated the sources of variability of each of the measured metabolites. The combined effect of host genetics, gut microbiome, and dietary variation in explaining the variability of plasma metabolite levels and ratios was evaluated by fitting linear regularized additive models (elastic net) on a train set and by estimating the determination coefficient (R^2) on a test set (see details in 5 section). We trained two models, using LLD or Rotterdam Study III-2 as the train set and the left-out cohort as the test set, respectively (Figure 2). Genetic contributions to metabolite variability were large for TMAO precursors but small for TMAO and TMAO-to-precursor ratios, although it showed low replicability in the testing cohort. We observed a similar pattern
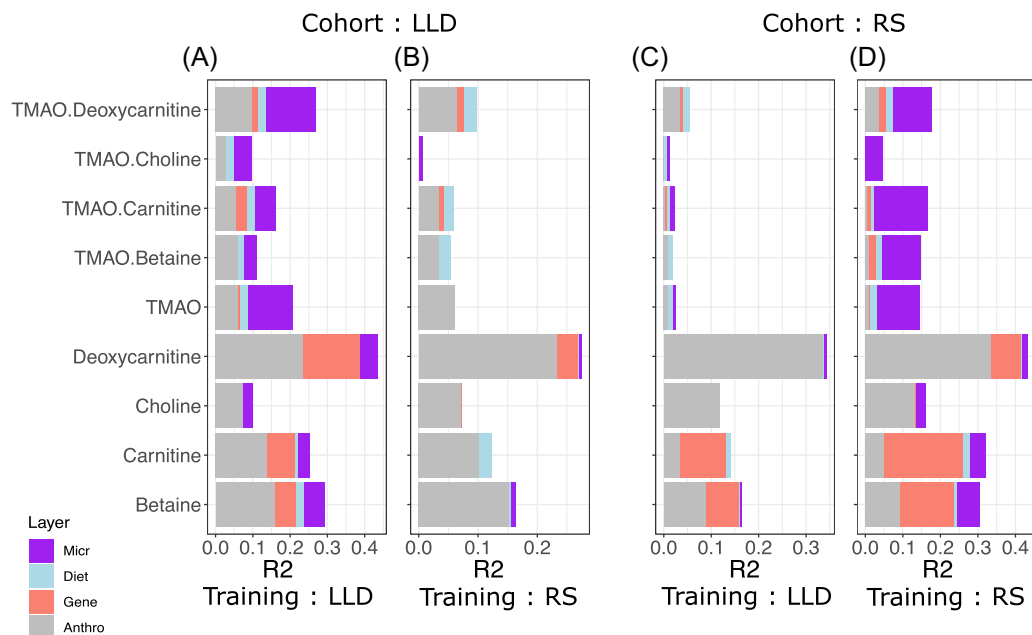
FIGURE 2 Variance explained in gut microbiome-related metabolite levels and ratios by different data layers. The X-axis shows the coefficient of determinations R^2 gained with each additional data layer (gray - anthropometrics; red - genetics; blue - diet; purple - microbiome). Two models were trained, using LLD (A and C) or Rotterdam Study III-2 (named RS) (B and D) as the train set, and the left-out cohort as the test set respectively.

regarding microbial features. For both LLD and Rotterdam Study III-2-trained models, a large proportion of TMAO and TMAO ratio variability could be explained by the microbiome in the training set. However, this effect was lost in the test sets. On the other hand, diet showed small but consistent effects between cohorts, while anthropometric effects seemed to be larger in LLD, both in the Rotterdam Study III-2-trained and LLD-trained models.

Subsequently, due to the lack of consistency found in the cohort-trained models, we sought to identify consistent associations between individual metabolites and host genetics, gut microbiome, or diet through meta-analysis.

Genetic variants are underlying levels of TMAO precursors, but not of TMAO itself

To evaluate host genetic determinants of TMAO-related metabolites and ratios, we performed a genome-wide association study (GWAS) ($N = 7093$) (Figure 3A). The quantile–quantile plots indicated that any cryptic relatedness and/or population stratification were well-controlled after genomic correction (λ ranged between 1.00 and 1.02) (Supporting Information S1: Figure S2, Supporting Information S2: Table S5). Meta-analyses identified 55 independent genetic variants mapped to five genomic regions for betaine, 89 independent

genetic variants mapped to three genomic regions for carnitine, 10 mapped to three genomic regions for choline, and 37 mapped to three genomic regions for deoxycarnitine (Bonferroni corrected genome-wide significance level, p value $< 8.33 \times 10^{-9}$) (Supporting Information S2: Tables S6–S8). Of these genomic regions, two genomic regions for betaine (6p21.1, 6q25.3), two for choline (2q34, 5q31.1), and two for deoxycarnitine concentration (10q21.2, 11p14.2) have not been reported in previous association studies (Supporting Information S2: Table S7). GWAS of TMAO revealed no genetic variant associated with TMAO concentration at a genome-wide significant level (Figure 3A), in line with previous studies [30–33]. As this might suggest that genetic variants have a weak effect on variation in TMAO levels, we combined our results with the results from Hartiala et al. [32] to increase our sample size and improve power. However, no difference in signal was observed. Similarly, an exploration of genetic common and rare variants in the *FMO* gene yielded no significant associations (minimum p value > 0.1), in contrast to previous findings [34]. On the other hand, GWAS of TMAO-to-precursor ratio revealed genetic loci associated with TMAO to betaine ratio ($n = 2$), TMAO to carnitine ratio ($n = 2$), and TMAO to deoxycarnitine ratio ($n = 1$) (p value $< 8.33 \times 10^{-9}$) (Figure 3B, Supporting Information S2: Table S9), all of which are overlapping with the precursor's findings.
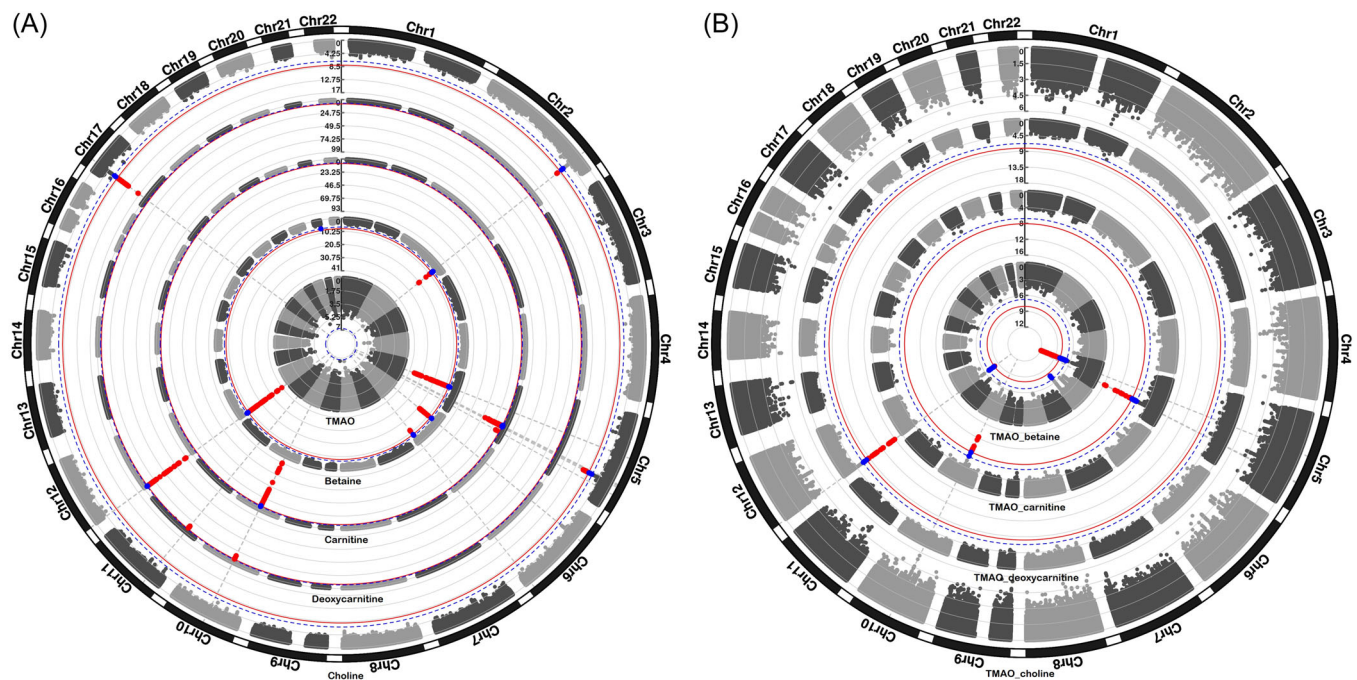
FIGURE 3 Results of genome-wide association analyses for (A) individual metabolites and (B) TMAO-to-precursor ratios. Each dot represents a genetic variant. Genetic variants surpassing the Bonferroni corrected significance threshold (p value $< 8.33 \times 10^{-9}$) are highlighted in red. Genetic variants showing suggestive evidence of association (p value $< 1.7 \times 10^{-7}$) are highlighted in blue.

The genetic variants found have previously been associated with various metabolic (e.g., plasma cholesterol and triglycerides levels, metabolite levels, liver enzyme levels), anthropometric (e.g., height, waist circumference, BMI), and medical traits (e.g., chronic kidney disease, diabetes mellitus, blood pressure) (Supporting Information S2: Table S10).

Furthermore, we performed a sex-stratified analysis ($N_{females} = 4026$, $N_{males} = 3067$). All significant associations observed in males and almost all significant associations found in females were also significant in the overall analysis (Supporting Information S1: Figure S3, Supporting Information S2: Table S11). An exception was the intronic variant mapped to the *RP1* gene which showed a significant association with TMAO in females (beta = −0.29, p value = 2.63×10^{-9}) but not in males (beta = 0.02, p value = 0.78) or the overall analysis. This variant showed heterogeneity between males and females (p value = 3.35×10^{-5}). The *RP1* gene has been reported to function in photoreceptor differentiation (GeneCards Version 3: the human gene integrator).

Gene-based association analysis revealed 11 genes associated with betaine, 16 with carnitine, six with deoxycarnitine, and two with choline at the gene-wide significance level (p value $< 4.42 \times 10^{-7}$) (Supporting Information S2: Table S12). No significant gene sets were identified in the gene-set analysis (Supporting Information S2: Table S13).

Heritability estimates and genetic correlation

Single nucleotide polymorphism (SNP)-based heritability of metabolites was estimated in a range from 0.16 (standard error [SE] = 0.07) for choline to 0.28 (SE = 0.08) for betaine using linkage disequilibrium (LD) score regression (Supporting Information S1: Figure S4a). An overlap of lead genetic loci was observed between betaine and choline (2q34), betaine and deoxycarnitine (12p13.33), carnitine and deoxycarnitine (10q21.2), and carnitine and choline (5q31.1) (Supporting Information S1: Figure S4b). Additionally, we examined genetic overlap on a genome-wide level by computing genetic correlations. Evidence of suggestive genetic overlap was observed for TMAO and choline ($\rho_{genetic} = 0.63$, p value = 2.77×10^{-2}) and betaine and choline ($\rho_{genetic} = 0.54$, p value = 5.7×10^{-3}), while evidence of significant genome-wide genetic overlap was observed between TMAO and TMAO-to-precursor ratios (Supporting Information S2: Table S14).

Microbial taxa are associated with plasma levels of TMAO, but not with those of its precursors

Next, we conducted a meta-analysis between 241 relative abundances from gut taxa and plasma metabolite

concentrations (Figure 4A) identifying 68 associations (Bonferroni multiple test threshold p value $< 6.21 \times 10^{-5}$). Association effect sizes across cohorts were generally concordant, with no evidence of heterogeneity in most associations (Supporting Information S2: Table S15). Significant associations were predominantly seen for TMAO (11/68) and TMAO-to-precursor ratios (56/68) (Supporting Information S2: Table S15). From those, the TMAO/carnitine ratio was associated with an abundance of 23 microbial taxa and the TMAO/choline ratio with 15 microbial taxa. The top associated taxa corresponded to several *Ruminococcaceae* genera (NK4A21, UCG003, UCG005), which were positively associated with TMAO abundance ratio to its precursors, carnitine and choline. Other top associations included the class *Actinobacteria*, which was consistently negatively associated with TMAO-precursor ratios. The only significant association not directly related to TMAO was a positive association between the genus *Haemophilus* and betaine. Some members of this taxonomic group are able to oxidize choline to generate betaine, although we did not observe a negative association between these taxa and plasma choline [35]. Interestingly, most of the significant or close to significant TMAO associations showed a different direction of effect than associations to at least one precursor (Supporting Information S1: Figure S5). This might highlight taxa with the potential to metabolize TMAO precursors into TMA.

Sex-stratified analyses identified four significant results in women (Supporting Information S2: Table S16) and four in men (Supporting Information S2: Table S17) (Supporting Information S1: Figure S6). If we focus on suggestive associations in one of the stratified analyses based on sex (p value $< 1.2 \times 10^{-3}$) but not found to be associated in the overall analysis (p value $> 1.2 \times 10^{-3}$), we could identify 14 associations, 10 of which showed a significant heterogeneity (p value < 0.05) between women and men (Supporting Information S1: Figure S7). The most noteworthy heterogeneous association is between *Coprococcus* and carnitine, for which a clear negative association is seen in males (beta $= -0.0178$, p value $= 0.039$), which does not appear in females (beta $= 0.006$, p value $= 0.583$). This taxon has previously been linked to be sex-linked in mice and pigs [36, 37].

In addition, we leveraged metagenomic shotgun sequencing data available for LLD to quantify specific
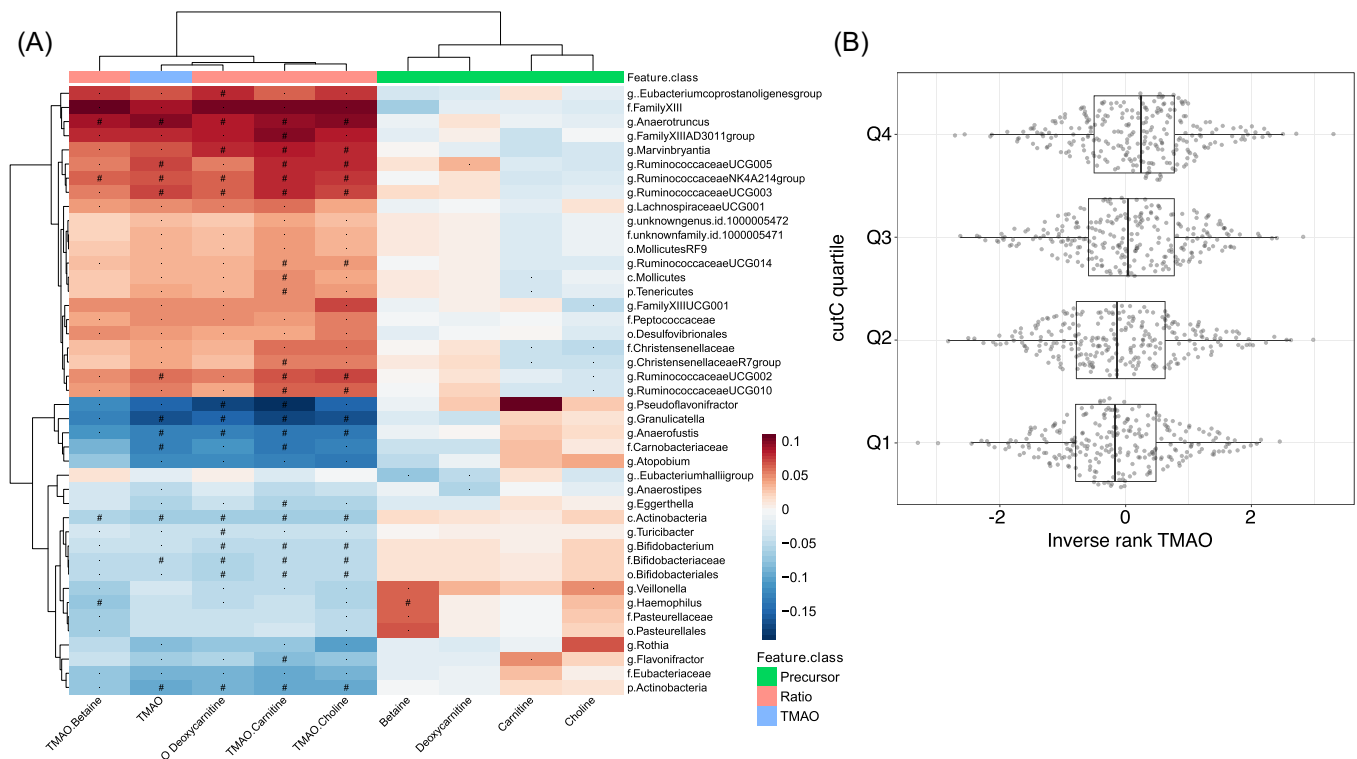
FIGURE 4 Microbial associations with plasma metabolic concentrations (A) Heatmap showing results of the association analysis between metabolites and gut microbial taxa. Displayed results are after adjustment for age, sex, BMI, and study-specific covariates. Metabolites are displayed on the *x*-axis and gut microbial taxa are shown on the *y*-axis. Red color denotes positive associations and blue color stands for negative associations. Hash symbol (#) represents the Bonferroni significant associations (p value $< 6.2 \times 10^{-5}$), while star denotes suggestive associations (p value $< 1.2 \times 10^{-3}$). (B) Boxplots displaying the distribution of normalized TMAO concentration per clr-transformed abundance quartiles from cutC, measured from metagenomic shotgun sequencing in LLD.

gene pathways previously related to TMA metabolism from dietary precursors. We quantified a total of 624 proteins belonging to glycine betaine/carnitine transport cluster (gbu A, B, and C), in addition to other known TMA metabolic genes, as choline trimethylamine-lyase (cutC), choline trimethylamine-lyase activating enzyme (cutD), the carnitine monooxygenase oxygenase subunit (YeaW), and L-carnitine/deoxycarnitine antiporter (caiT) gene families. In total, we observed a positive significant association between the abundance of cutC and TMAO (Figure 4B) concentrations and TMAO-to-precursor ratios (Supporting Information S2: Table S18). This indicates that this gene might be the major contributor to TMAO variability, or that the major contributors using the other genes were not present in the UniRef90 database we used to extract the gene sequences.

To examine causality, we performed two-sample MR between taxa and metabolites (excluding ratios). However, we failed to find any significant association (Supporting Information S2: Table S19).

Diet is associated with plasma levels of TMAO and its precursors

Using 4685 samples with food frequency questionnaires (FFQs) belonging to all but 300-OB studies, we ran correlations between metabolites and dietary data categorized into 13 major food groups while adjusting for age, sex, and BMI (Supporting Information S1: Figure S8, Supporting Information S2: Table S20). There were 16 correlations that surpassed the Bonferroni corrected threshold for multiple testing (p value $< 7.58 \times 10^{-4}$). Top positive correlations were observed between TMAO or TMAO-to-precursor ratios and fish intake. A significant correlation between fish and vegetable intake led to a correlation between TMAO and vegetable intake, which was not significant if fish intake was accounted for. Betaine levels were positively associated with grains, vegetables, and nuts, while carnitine levels showed a positive association with meat, and negative correlations were observed between cheese intake and choline, and egg intake and deoxycarnitine. Overall, our findings are in line with previous work [14, 26, 38].

DISCUSSION

We have performed an in-depth study to identify the potential relationship between plasma levels of TMAO and its precursors (betaine, carnitine, deoxycarnitine, and choline) and various factors including clinical outcomes, host genetics, gut microbiota composition, and diet in up to 7834 participants from five prospective cohort studies mainly based on the general Dutch population.

We observed a significant association between plasma choline concentrations and higher CVD risk. Higher circulating levels of choline correlated with a 17% increase in risk for CVD events. Choline is an essential nutrient that plays a role in various metabolic processes such as C1 metabolism and the synthesis of phospholipids. As such, choline metabolism interacts with the pathways of insulin sensitivity, fat deposition, and energy metabolism [39]. The relationship between dietary choline intake and circulating levels of free choline and its metabolites is concealed by homeostatic regulations and rapid tissue uptake, resulting in a narrow concentration range [39]. Despite being a common dietary compound, we did not identify a strong predictive potential of diet in plasma choline levels. This might be related to the fact that blood samples in our study were mainly collected after overnight fasting. Previous studies reported associations of circulating levels of choline with a higher risk of CVD, mortality and with some traditional cardiovascular risk factors including lower HDL, higher systolic pressure, and triglycerides [4, 8, 20, 40, 41]. However, there are also studies focusing on dietary choline intake with conflicting results. Higher dietary intake of choline was not predictive of incident coronary heart disease or CVD mortality [25–28]. However, a systematic review and meta-analysis of six prospective studies reported an association between dietary choline intake and incident CVD [41]. Inconsistency in findings may be related to differences in dietary patterns, sample sizes, follow-up periods, and geographical location. We also investigated the causality of our association by means of an MR analysis, which failed to support a causal relationship, as seen in previous work [42]. Future studies should focus on improving the strength of the instrumental variables for plasma levels of choline.

In contrast to circulating choline, we did not find an association between plasma TMAO levels and CVD risk or mortality in our population-based cohorts. Even though a number of studies have shown a relationship between TMAO levels and CVD risk, these results are mostly derived from individuals with a high risk of CVD, existing disease, or multimorbidity [6–8, 43–46]. In line with our results, a few population-based studies did not identify an association between TMAO and cardiometabolic markers, carotid intima media thickness, CVD events including heart attack and stroke and mortality, either [9, 21, 47, 48], while some other studies did report significant associations [10–12]. To the best of our knowledge, no apparent factor can explain this between-study heterogeneity, such as sample size, concentration levels, ethnic groups,

follow-up times, and employed methods. These factors were comparable between negative and positive studies. Taken together, we believe that the predictive value of TMAO in CVD onset is still not robust, and further studies need to be done to address its complex interaction with diet and gut microbiome, to understand the existent inter-study heterogeneity.

Next, we aimed to understand what factors are the most important drivers of metabolite variability in the general population. Exploring the role of host genetics in underlying plasma levels of TMAO and its precursors revealed genome-wide significant variant associations with plasma levels of TMAO precursors but not for TMAO itself. We confirmed some of the previously reported associations implicated in determining plasma levels of all TMAO precursors. The *DMGDH, BHMT*, and *BHMT2* genes, mapped to the 5q14.1 region, have previously been linked to betaine concentration [49, 50]. These genes are involved in betaine metabolism which is related to a series of interlinking metabolic pathways that include the methionine and folate cycles [51]. Choline is also related to this pathway and we confirmed association with the *PEMT* gene (17p11.2) encoding an enzyme critical in phosphatidylcholine synthesis [52]. Further-more, the *SLC6A13* gene, mapped to the 12p13.33 locus, has previously been linked to deoxycarnitine levels while the *SLC16A9* gene, mapped to the 10q21.2 locus, has been associated with carnitine levels [53]. Interestingly, we have identified the 10q21.2 region as a novel genetic region underlying deoxycarnitine levels. The lead variant of this region mapped to the *SLC16A9* gene which is involved in urate metabolism. *SLC16A9* encodes a membrane transporter and is expressed in the intestine (GTEx Analysis Release V8), which might indicate a role in deoxycarnitine absorption. Previous studies showed that deoxycarnitine is an intermediary metabolite pro-duced from carnitine by gut microbiota, but it may also act as a precursor to carnitine synthesis [54]. Interestingly, we identified a novel genetic locus in the GWAS of deoxycarnitine that depicts this process. More specifically, the top lead intronic variant of the 11p14.2 locus was mapped to the *BBOX1* gene which is known to catalyze the formation of carnitine from deoxycarnitine and is therefore involved in the carnitine synthesis pathway.

Additionally, we identified the 6p21.1 region as a novel region associated with betaine levels. This region has previously been associated with stroke and type 2 diabetes [52, 53]. Genetic variants in linkage disequilibrium ($r^2 > 0.8$) with our lead variant were associated with differential expression of *GNMT* and *PEX6* genes. Interestingly, the *GNMT* gene is involved in the metabo-lism of methionine. Among novel regions, we have also identified the 2q34 locus underlying the plasma

concentrations of choline. The lead genetic variant of this region is mapped to the *CPS1* gene which is involved in the urea cycle. Variants in this gene have previously been linked to creatinine, glycine, betaine and homocysteine levels, BMI, systolic blood pressure, and cholesterol levels [49, 55–59].

Although we were not able to detect genetic variants underlying plasma TMAO concentrations, we estimated that 20% of genetic variability in TMAO concentration could be explained by common genetic variants. To discover genetic determinants of TMAO, future studies should further increase sample size and focus on complex genetic effects. Additionally, diet intervention studies might be of help, as these could decrease the variability attributed to gut microbiome and diet.

Microbial abundance was mainly associated with TMAO and TMAO-to-precursor ratios, which may be interpreted as a proxy for microbial conversion rates. Dietary precursors, on the other hand, did not show strong microbial associations. Several of the taxa we identified have previously been linked to TMAO in European, North American, and Asian studies. For instance, among the taxa belonging to the positively associated cluster, *Ruminococcus* or uncultured *Rumino-coccaceae* have frequently been described to correlate with TMA and/or TMAO concentrations in mice and humans [20, 60–62]. A member of *Family XIII* was associated with TMAO concentrations in mice [63]. *Anaerotruncus* was seen to be decreased upon resveratrol treatment in a mouse model and has been linked to TMAO metabolism [61]. In the same study, several bacterial taxa were increased after resveratrol treatment, including *Bifidobacterium, Bifidobacteriaceae*, and *Bifido-bacteriales*, which are negatively associated with TMAO in the present study, also in agreement with other observations in humans [15]. Interestingly, the strongest negative associations were found between TMAO-to-precursor ratios and *Pseudoflavonifractor* or *Granulica-tella* genera, which have not been reported before. Conversely, other taxa that are often linked to TMAO [64] did not show any significant association in our study, including *Clostridia* or *Escherichia* genera. *Desul-fovibrionales*, although did not pass the Bonferroni corrected p-value threshold, showed consistent positive associations with TMAO.

In addition to gut microbiota, diet composition was observed to be an important determinant of metabolite concentrations. For instance, TMAO concentrations and TMAO-to-precursor ratios showed a positive association with fish intake. Previous studies linked TMAO concen-trations to fish intake and this association has been demonstrated to vary throughout populations and/or regions [15]. Fish consumption was associated with

TMAO concentrations in Asian countries and some European countries, while intake of eggs and red meat showed a stronger correlation with TMAO in the population of the United States [17–19, 65]. In our study population, no association was found between TMAO and eggs or red meat. Fish and seafood are rich in TMAO and TMA and they can be directly absorbed without being transformed by gut microbiota [19]. As previous studies linked TMAO to atherosclerosis and CVD, these findings might be counterintuitive as fish is generally accepted to be cardioprotective [66].

Overall, the large sample size, population-based design, and comprehensive molecular and epidemiological data of our study helped us to investigate the sources of variation of TMAO and its precursor metabolites in the general population and their health-related consequences. We were able to study not only TMAO but also the compounds implicated in the TMAO biosynthesis pathway. Furthermore, we were also able to improve statistical power and internally cross-check the findings by combining data from five population-based studies. However, our study also has limitations. Microbial species genetic variation is known to modulate bacterial-related metabolites, thus the taxonomic resolution of 16S might not properly reflect the metabolic potential of the present strains [67]. Metagenomic-shotgun sequencing experiments will be needed to address that level of variation. The cross-sectional nature of our metabolomics measurements and gut microbiota assessment in our study only allowed us to investigate the relationship between the two at one time point. To complement our findings and advance our understanding, future studies should focus on assessing longitudinal changes. For instance, a 1-year follow-up study in the general population reported large variations in plasma TMAO concentrations, which might underlie the heterogeneous associations related to this metabolite [68].

CONCLUSION

In conclusion, our data add up to the mounting evidence of research showing that TMAO is not associated with an increased risk of CVD in the general population, despite earlier evidence suggesting this to be the case among patient groups. However, we did show a significant relation between plasma choline levels and higher CVD risk. Our MR revealed no evidence of a causal link between TMAO or its precursors with incident CVD. Furthermore, we also identified several determinants explaining the variability of TMAO and its precursors' blood levels in humans. Gut microbiome was mainly associated with TMAO-to-precursor ratios, although the total variability explained of TMAO concentration remains mild and cohort-specific. Diet was associated with both

TMAO and its precursors but could not explain a great proportion of their variation. Genetic contributions to precursor concentrations were greater than to TMAO itself, where no strong genetic effects were seen. The biological mechanisms underlying these associations should be the subject of further studies. Overall, our investigation of the factors determining the interindividual variability of the metabolite concentrations might be used in the future to target interventions aimed at module circulating metabolites. Diet and microbes offer an ethical and effective intervention target that shall be further investigated in follow-up interventional studies.

METHODS

Study population

Our study population included 7834 participants from five cohort studies. Detailed description of participating studies can be found in Supporting Information and descriptive characteristics of study participants are shown in Table 1. Each study was approved by ethical committees (please see Supporting Information for details). Written informed consent was obtained from all participants.

Metabolite profiling

TMAO and its precursors betaine, carnitine, deoxycarnitine, and choline were quantified in plasma samples of participants from five cohorts by using the liquid chromatography tandem mass spectrometry (LC–MS/MS) method. A detailed description of the method can be found elsewhere [33]. Briefly, before introducing the sample to the mass spectrometer, an analytic column with a C_{18} stationary phase was used to realize an online cleanup of it. The analytes were not retained by this stationary phase but important matrix interferences were retained, such as (phosphor–)lipids [69]. The descriptive statistics of metabolites were coherent across the cohorts (Supporting Information S2: Table S1). In addition to individual metabolites, ratios of TMAO to its dietary precursors were also calculated.

Incident clinical outcomes

The Rotterdam Study I-4, Rotterdam Study III-2, and LLS cohorts had data on incident clinical outcomes including CVD and mortality. Incident CVD events were defined as incident stroke, myocardial infarction, angina pectoris, and heart failure according to the codes of International

Classification of Disease, 10th edition. Incident CVD events were assessed continuously through an automated digital linkage of the study database to medical records maintained by general practitioners in the Rotterdam Study and from general practitioner records in the LLS [70, 71]. Information on vital status is additionally obtained from the central registry of the municipality of the city of Rotterdam, and in January 2021, the vital status for participants of the Leiden Longevity Study was updated through the Personal Records Database which is managed by the Dutch governmental service for identity information (https://www.government. nl/topics/personal-data/personal-records-database-brp) [72].

Baseline clinical characteristics

The baseline clinical characteristics were obtained by means of interviews, physical examination, blood sampling, or medical records from the general practitioner. Assessment of current smoking status, weight, height, blood pressure, glucose concentrations, total LDLC and HDLC, serum TG, creatinine, and medication use including lipid-lowering medication and use of medication indicated for the treatment of diabetes. Diabetes was defined as fasting glucose concentrations above 7 mmol/L, nonfasting glucose concentrations above 11.1 mmol/L (only if nonfasting concentrations were unavailable), or use of medication indicated for the treatment of diabetes [73], or through medical records from general practitioners. Hypertension was defined as systolic blood pressure ≥140 mmHg, diastolic blood pressure ≥90 mmHg, or the use of medication for the treatment of hypertension [73], or through medical records from the general practitioner. BMI was calculated as weight in kilograms divided by the square of height in meters. Quantified creatinine was used to calculate the eGFR using the Chronic Kidney Disease Epidemiology Collaboration equation (CKD-EPI) [74].

Genotyping and imputation

Details on genotyping platforms, calling methods, and quality control (QC) procedures in participating studies are shown in Supporting Information S2: Table S20. Commercially available genotyping arrays were used for genotyping. Similar QC procedures were applied in each study before genotype imputation (Supporting Information S2: Table S20). Genotypes in each cohort were imputed by the Haplotype Reference Consortium reference panel on a Michigan Imputation Server [75].

All genomic coordinates were lifted to Human Build GRCh37/hg19.

Microbiome processing

The Rotterdam Study III-2 and LLD cohorts had 16S-amplicon sequence data available from fecal samples matching plasma collection. Fecal sample collection and 16S sequencing was described elsewhere [76, 77]. In brief, DNA was isolated from the fecal samples belonging to the Rotterdam Study III-2 cohort, and the V3 and V4 variable regions of the 16S rRNA gene were amplified and sequenced on the Illumina MiSeq sequencer. Similarly, DNA was isolated from fecal samples of LLD participants and the 16S V4 region was sequenced at the Broad Institute using Illumina MiSeq. 16S rRNA data were processed as previously described in a large 16S meta-analysis including both cohorts [78]. In brief, samples were rarified to 10,000 reads. Reads were classified to a given taxonomic level (genus to phyla) using the RDP classifier (v2.12) [79]. Reads below 0.8 posterior probability to belong to a given taxonomic level were discarded. For each sample, each taxonomic level was centered log-ratio (clr) transformed. Only taxa observed in above 10% of participants per cohort were used for association resulting in 241 microbial taxa.

Dietary assessment

The Rotterdam Study cohorts, LLS and LLD had data on dietary intake collected via a validated FFQ. Data on dietary intake in the LLS, LLD, and Rotterdam Study cohorts were collected via validated FFQs [80–82]. The FFQs assess the frequency of consumption of food items and the number of servings per day. Additionally, information on portion size, type of food item, and preparation methods was collected. The average daily energy and nutrient intake were calculated using the Dutch Food Composition Database. Specific food items were aggregated into food groups in grams per day. The major food groups overlapping between the cohorts were used for subsequent analysis including vegetables, fruit, grains, nuts, eggs, fish, meat, poultry, processed meat, cheese, milk, yogurt, and total dairy products.

Incident clinical outcomes analysis

Metabolites were transformed using rank-based inverse normal transformation. The relationship between metabolites and incident clinical outcomes was assessed using Cox proportional hazards regression models with age as a time scale. The analyses were adjusted for sex, BMI, hypertension, diabetes, LDLC, HDLC, TG, lipid-lowering medication use, current smoking, and fasting status (if appropriate) (model 1). Subsequently, the associations were adjusted for eGFR (model 2). The proportional

hazard assumption was checked using statistical tests incorporated in the *survival* package. Violation of this assumption was observed for some of the covariates and was resolved by stratification. In the LLS, Cox-type random effect (frailty) regression models were used to adjust for family relations. All analyses were performed using *R*.

The summary statistics results of participating studies were combined using inverse variance-weighted fixed-effect meta-analysis in METAL [83]. The heterogeneity of effects was assessed by I^2 which indicates the percentage of variance in the meta-analysis attributable to study heterogeneity [74]. To model the correlation between metabolites, we first used the method of Li and Ji to calculate the number of independent tests [84]. The Bonferroni corrected significance threshold was calculated based on the number of independent tests and set at 0.05/6 independent metabolites $= 8.33 \times 10^{-3}$. Additionally, all analyses were stratified by sex. The same steps were followed for the overall analysis.

The potential causal effect of metabolites on clinical CVD outcomes was assessed by MR analysis using the TwoSampleMR package. GWAS summary statistics results were obtained from a large meta-analysis comprising coronary artery disease cases and controls of UK Biobank resource and CARDIoGRAMplusC4D [85]. Genetic variants with *p* value $< 1 \times 10^{-5}$ were used as instruments. Independent genetic variants were selected based on r^2 in European reference data. The results were kept if these were based on at least three shared genetic variants. Causality was estimated using various MR methods including inverse variance weighted (IVW), MR-Egger, Wald ratio, Weighted median, Simple Mode, and Weighted Mode.

Estimation of determination coefficient in different data layers

The Rotterdam Study cohorts and LLD were used to estimate the total determination coefficient (R^2) from each of the analyzed data layers in each of the metabolites or metabolite ratios. We used features present in both cohorts including anthropometric covariates (age, sex, BMI), eight overlapping dietary items, 242 bacterial taxonomic abundances, and the number of suggestive genetic variants (*p* value $< 1 \times 10^{-5}$) from a meta-analysis of the GWAS results from LLS and 300-OB. These two cohorts were used for preselecting variants and were not used to train the model or estimate R^2. Taxa-abundance was clr-transformed, while metabolites and diet were inverse-rank normal transformed.

We trained a regularized additive linear model, elastic net (glmnet v4.0), and selected the best combinations of

hyperparameters alpha (regularization mix) and lambda (regularization strength) through a five-repeated 10-fold cross-validation procedure using the root mean square error as a performance metric (caret v6.0, tunelength = 10). We trained a model for each metabolite (or metabolite ratio) using two different training sets, a training set consisting of the LLD cohort (784 samples) and a training set consisting of the Rotterdam Study III-2 cohort (772 samples). For the test set (LLD in the Rotterdam Study III-2-trained or the Rotterdam Study III-2 in the LLD-trained model), the determination coefficient (R^2) was estimated in nested models. To estimate anthropological R^2, all other coefficients were made 0. To estimate genetics R^2, all nongenetics, nonanthropological covariate coefficients were made 0. This was followed by the addition of non-0 diet parameters and finally the complete model including microbial features. Individual layer R^2 was quantified by subtracting the R^2 from the nested models, for example, microbial R^2 was estimated by subtracting the complete model R^2 and the diet model.

Genome-wide association analysis

Each participating study performed genome-wide association analysis under an additive model using metabolites as a dependent variable and variant allele dosage as a predictor. Before the analysis, metabolites were transformed using rank-based inverse normal transformation. The association analysis was adjusted for age, sex, fasting status if applicable, familial relatedness if appropriate, and principal components if needed. Study-specific details on covariates and software used to run the analysis are provided in Supporting Information S2: Table S20. The QC was performed using a standardized protocol provided by Winkler et al. [86]. Genetic variants with minor allele count below 10 and low imputation quality ($r^2 < 0.3$) were excluded. The summary statistic results were combined using fixed-effect meta-analysis in METAL. To account for a small amount of population stratification or unaccounted relatedness, genomic control was applied. After meta-analysis, genetic variants that were present in less than three participating studies were filtered out. The Bonferroni corrected genome-wide significance threshold was set at $5 \times 10^{-8}/6$ independent metabolites $= 8.33 \times 10^{-9}$. Additionally, the analyses were stratified by sex. The same QC steps were followed for the overall analysis. The sex-stratified summary statistic results were combined using fixed-effect meta-analysis in METAL while applying genomic control. Test statistics of each variant were tested for heterogeneity between males and females.

Functional mapping, annotation, and enrichment analysis

Functional Mapping and Annotation of genetic associations (FUMA) was used to characterize genomic loci [87]. Genetic variants that passed the Bonferroni corrected genome-wide significance threshold and were independent of each other ($r^2 < 0.6$) were defined as independent genetic variants. Independent significant genetic variants with $r^2 < 0.1$ were defined as lead genetic variants. Independent significant genetic variants with $r^2 \geq 0.1$ or that were 250 bp or closer were assigned to the same genomic risk locus. Each locus was represented by the top lead genetic variant with a minimal p value in the locus. Functional annotation was performed using Combined Annotation Dependent Depletion [88], HaploReg [89], and RegulomeDB [90] tools as implemented in FUMA.

Genome-wide summary statistics were used to perform gene-based analysis using multi-marker analysis of genomic annotation as implemented in FUMA. Genetic variants were assigned to the genes from Ensembl build 85 based on genomic location. All genetic variants mapped to the protein-coding genes were tested for association with metabolites using the SNP-wide mean model. 1000G phase 3 was used as a reference panel to calculate LD across SNPs and genes. To account for multiple testing, Bonferroni correction was calculated, and the gene-wide significance threshold was set at $0.05/(18\,861$ tested genes $\times 6$ independent metabolites) = 4.42×10^{-7}. Subsequently, gene sets enrichment analysis was also performed using FUMA. Hypergeometric tests were performed to test if genes of interest are overrepresented in any of the 15,496 predefined gene sets obtained from MsigDB. Multiple test correction was calculated based on the total number of gene sets (p value = $0.05/15,496 = 3.23 \times 10^{-6}$).

Heritability estimates and genetic correlation

The heritability of metabolites and the genetic correlation between them were estimated from GWAS results using the LD Score Regression approach [91]. We used precomputed LD scores for Europeans. Only genetic variants available in HapMap3 were used.

Gut microbial taxa and metabolites association analysis

Each of the metabolites and TMAO-to-precursor ratios were rank-based inverse normal transformed. Standard linear regression models were carried out to associate the transformed metabolite and taxonomy abundances while adjusting for age, sex, BMI, and study-specific covariates (sample batch, time in mail, and storage time in the Rotterdam Study III-2). This analysis was also reproduced in sex-stratified samples. Common taxonomy associations in the Rotterdam Study III-2 and LLD were meta-analyzed using a fixed-effects, inverse-variance analysis (R package meta v4.12). The association's heterogeneity was measured by Cochran's Q statistic.

To correct for multiple testing, we determined the number of independent tests using the method of Li and Ji [84]. There were 134 independent tests among microbial taxa and six independent tests among metabolomic measures. The Bonferroni significance threshold was set at $0.05/(134$ independent microbial taxa $\times 6$ independent metabolites) = 6.2×10^{-5}, while a suggestive threshold was set at $1/(134 \times 6) = 1.2 \times 10^{-3}$.

The potential causal relation between gut microbial taxa and metabolite levels was tested by two-sample MR using the TwoSampleMR package. Only taxa-to-metabolite associations that surpassed our significant thresholds (p value $< 6.2 \times 10^{-5}$) were considered for the analysis. We used the GWAS summary statistics for metabolites produced in this work, while for microbial taxa we obtained summary statistics from a large 16S meta-analysis comprising 18,340 individuals from 24 cohorts including both Rotterdam Study III-2 and LLD [88]. Genetic variants with p value $< 1 \times 10^{-5}$ were selected as instruments. Bacterial taxa with no available instruments were removed from the analysis. Independent genetic variants were selected as instrumental variables based on an R^2 threshold of 0.001 (1000 Genomes in the European reference population). The number of instruments varied between 11 and 20. The causality was estimated using various MR methods including IVW, MR-Egger, Wald ratio, Weighted median, Simple Mode, and Weighted Mode. In addition, we also assessed genetic variant heterogeneity and evidence of horizontal pleiotropy (using Egger). Individual summary statistics for genetic variants were estimated using Wald ratio tests.

To investigate the correlation of distinct microbial pathways implicated in the metabolism of TMA from dietary precursors [92, 93], including the glycine betaine/carnitine transport cluster (gbuA, gbuB, and gbuC), cutC, cutD, the YeaW, and caiT, we conducted an analysis based on 845 matching annotations extracted from UniRef90 (June 2019). For the analysis, we utilized the remaining UniRef90 data set as a background universe. A database of unique oligopeptides was constructed using ShortBRED (v 0.9.4) [94] to enable unambiguous mapping of reads for subsequent quantification.

Metagenomic shotgun sequencing data from the LifeLines-DEEP cohort [95] was obtained for further analysis (EGA accesion: EGAC00001000457). Preprocessing of shotgun data was carried out using publicly available pipelines (https://github.com/SegataLab/preprocessing). Specifically, we employed Trim Galore (v0.6.6) to remove low-quality reads (<Q20), reads with lengths <75 bp, and those with more than two ambiguous nucleotides. Host DNA contamination was mitigated by aligning reads to the hg19 human genome using Bowtie2 (v2.3.4.3) [96]. Subsequently, forward and reverse reads were merged, and gene quantification was performed using ShortBRED. This approach yielded the quantification of 624 proteins (caiT = 96, cutC = 81, cutD = 109, GbuA = 86, GbuB = 39, GbuC = 184, YeaW = 29). Normalization of read numbers per marker length was carried out, and the total abundance per gene family was computed. To address data compositionality, we applied the clr-transformation. Correlation analysis between gene family clr-transformed abundances and (inverse-ranked) metabolite concentrations was performed using a unique multivariable model while accounting for age and sex. To control for multiple testing, we implemented the Benjamini-Hochberg False Discovery Rate procedure.

Diet and metabolites correlation analysis

Each food group, metabolite and TMAO-to-precursor ratio were rank-based inverse normal transformed before analysis. Partial correlation coefficients were calculated between each transformed food group item and metabolite or metabolite ratio while adjusting for age, sex, and BMI. Summary statistics results of participating studies were combined by performing a fixed-effect meta-analysis in METAL. To model the correlation between food groups, the method of Li and Ji was used to calculate the number of independent tests. Associations were considered significant if they surpassed the Bonferroni corrected significance threshold of $0.05/(11 \text{ independent food groups} \times 6 \text{ independent metabolite}) = 7.58 \times 10^{-4}$.

AUTHOR CONTRIBUTIONS
Conception and design of the study: Sergio Andreu-Sánchez, P. Eline Slagboom, Cornelia M. van Duijn, Jingyuan Fu, Dina Vojinovic. Collection of the data: Alexander Kurilshikov, Marian Beekman, Mohsen Ghanbari, Martijn van Faassen, Inge C. L. van den Munckhof, Marinka Steur, Amy Harms, Thomas Hankemeier, M. Arfan Ikram, Maryam Kavousi, Trudy Voortman, Robert Kraaij, Mihai G. Netea, Joost H. W. Rutten, Niels P. Riksen, Alexandra Zhernakova, Folkert Kuipers, P. Eline Slagboom, Cornelia M. van Duijn, Jingyuan Fu. Analysis: Sergio Andreu-Sánchez, Shahzad Ahmad, Dina Vojinovic. Interpretation of the data: Sergio Andreu-Sánchez, Shahzad Ahmad, Alexander Kurilshikov, Marian Beekman, Trudy Voortman, P. Eline Slagboom, Cornelia M. van Duijn, Jingyuan Fu, Dina Vojinovic. Drafting of the manuscript: Sergio Andreu-Sánchez, Dina Vojinovic. All authors have read the final manuscript and approved it for publication.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data are available from the author on request.

All relevant data supporting the key findings of this study are available within the article and its supplementary information files. Other data are available from the corresponding author upon reasonable request. Due to ethical and legal restrictions, individual-level data from the Rotterdam Study (RS) cannot be made publicly available. Data are available upon request to the data manager of the Rotterdam Study Frank van Rooij (f.vanrooij@erasmusmc.nl) and subject to local rules and regulations. This includes submitting a proposal to the management team of RS, where upon approval, analysis needs to be done on a local server with protected access, complying with GDPR regulations.

The use of Lifelines data and materials must comply with the informed consent signed by Lifelines participants specifying that their collected data will not be used for commercial purposes. The raw 16S amplicon sequencing, metabolomics, and basic phenotype data (age, sex, and BMI) are deposited in the EGA database with the study ID EGAS00001001704 (https://ega-archive.org/studies/EGAS00001001704), which includes Data set ID EGAD00001003453 for raw sequencing data and data set for metabolomics data. There is a minimal access procedure for the access of the EGA data set that includes a contact address and an online data access form https://goo.gl/forms/TWHlrmbXaXNqWnnl2, which is very simple and is only intended to ensure that the data is being requested for research/scientific purposes only. Submitted data access forms will be evaluated by the data manager and Lifelines. For requests from verified academic parties, access will be granted within 2 weeks. For requests from commercial parties, Lifelines will perform a pre-Data Privacy Impact Assessment (pre-DPIA) to assess the risks of the proposed processing of personal data (e.g., purpose, storage, access, archiving, etc.) with respect to the GDPR subject rights. Based on the outcome of the pre-DPIA, Lifelines will decide whether sharing data with the commercial entity is allowed and/or whether additional measures have to be taken. Lifelines diet and genotype information should be directly requested from Lifelines. (1) Data are requested by filling out the application form to request "Available Lifelines-data" at https://www.lifelines.nl/researcher/how-to-apply/apply-here; (2) Lifelines will evaluate project proposals to ensure compliance with the Lifelines data access policy, informed consent of Lifelines participants and the GDPR, and that the data is being requested for

noncommercial research; (3) Upon approval, Lifelines will send Data and Material Transfer Agreement (DMTA) contracts to the applicants; and (4) After the required contracts are signed, Lifelines will provide access to data via the Workspace or HPC and link the raw and processed DMP sequencing data to the Lifelines phenotypes. Lifelines strives to accomplish steps 2–4 at 2 weeks per step, assuming that no extra actions by the applicant or Lifelines are required.

The source data of the Leiden Longevity Study are not automatically publicly available due to legal and informed consent restrictions. Reasonable requests to access the data sets should be directed to the PI of the Leiden Longevity Study P. Eline Slagboom (p.slagboom@lumc.nl) and Marian Beekman (m.beekman@lumc.nl). If data access is approved, the data users are bound by a data access agreement. This includes responsibilities with respect to third party data sharing and maintaining participant privacy. Further responsibilities include a responsibility to acknowledge data sharing.

The source genetic data of the 300OB cannot be shared. Interested researchers can send a proposal to Niels.Riksen@Radboudumc.nl with regard to the genetics data of the 300OB.

The scripts used for analysis are saved in GitHub https://github.com/GRONINGEN-MICROBIOME-CENTRE/Groningen-Microbiome/tree/master/Projects/TMAO_metagenomics/16S_association. Supplementary materials (figures, tables, graphical abstract, slides, videos, Chinese translated version and update materials) may be found in the online DOI or iMeta Science http://www.imeta.science/.

ETHICS STATEMENT

Leiden Longevity Study (LLS)

In accordance with the Declaration of Helsinki, we obtained informed consent from all participants before they entered the study. Good clinical practice guidelines were maintained. The study protocol was approved by the ethical committee of the Leiden University Medical Center before the start of the study (P01.113).

LifeLines-DEEP (LLD)

The Lifelines study was approved by the ethics committee of the University Medical Center Groningen (METc2007/152). All participants signed an informed consent form before enrollment. The LLD study was approved by the Institutional Ethics Review Board of the University Medical Center Groningen (ref. M12.113965), the Netherlands.

Rotterdam Study (RS)

The Rotterdam Study has been approved by the Medical Ethics Committee of the Erasmus MC (registration number MEC 02.1015) and by the Dutch Ministry of Health, Welfare and Sport (Population Screening Act WBO, license number 1071272-159521-PG). The Rotterdam Study Personal Registration Data collection is filed with the Erasmus MC Data Protection Officer under registration number EMC1712001. The Rotterdam Study has been entered into the Netherlands National Trial Register (NTR; www.trialregister.nl) and into the WHO International Clinical Trials Registry Platform (ICTRP; www.who.int/ictrp/network/primary/en/) under shared catalog number NTR6831. All participants provided written informed consent to participate in the study and to have their information obtained from treating physicians.

300 Obesity (300-OB) cohort

The 300OB study was approved by the IRB CMO Regio Arnhem-Nijmegen (number 46846.091.13).

ORCID

Sergio Andreu-Sánchez http://orcid.org/0000-0002-3503-9971
Shahzad Ahmad https://orcid.org/0000-0002-8658-3790
Alexander Kurilshikov https://orcid.org/0000-0003-2541-5627
Marian Beekman https://orcid.org/0000-0003-0585-6206
Mohsen Ghanbari https://orcid.org/0000-0002-9476-7143
M. Arfan Ikram https://orcid.org/0000-0003-0372-8585
Maryam Kavousi https://orcid.org/0000-0001-5976-6519
Trudy Voortman https://orcid.org/0000-0003-2830-6813
Robert Kraaij https://orcid.org/0000-0003-0416-999X
Mihai G. Netea https://orcid.org/0000-0003-2421-6052
Niels P. Riksen https://orcid.org/0000-0001-9197-8124
Alexandra Zhernakova https://orcid.org/0000-0002-4574-0841
Folkert Kuipers https://orcid.org/0000-0003-2518-737X
P. Eline Slagboom https://orcid.org/0000-0002-2875-4723
Jingyuan Fu https://orcid.org/0000-0001-5578-1236
Dina Vojinovic https://orcid.org/0000-0001-5744-0232

REFERENCES

1. Tang, W. H. Wilson, Takeshi Kitai, and Stanley L. Hazen. 2017. "Gut Microbiota in Cardiovascular Health and Disease." *Circulation Research* 120(7): 1183–96. https://doi.org/10.1161/CIRCRESAHA.117.309715
2. Wang, Zeneng, Elizabeth Klipfell, Brian J. Bennett, Robert Koeth, Bruce S. Levison, Brandon Dugar, Ariel E. Feldstein, et al. 2011. "Gut Flora Metabolism of Phosphatidylcholine Promotes Cardiovascular Disease." *Nature* 472(7341): 57–63. https://doi.org/10.1038/nature09922

3. Tang, W. H. Wilson, Zeneng Wang, Bruce S. Levison, Robert A. Koeth, Earl B. Britt, Xiaoming Fu, Yuping Wu, and Stanley L. Hazen. 2013. "Intestinal Microbial Metabolism of Phosphatidylcholine and Cardiovascular Risk." *The New England Journal of Medicine* 368(17): 1575–84. https://doi.org/10.1056/NEJMOA1109400

4. Wang, Zeneng, Wilson H. W. Tang, Jennifer A. Buffa, Xiaoming Fu, Earl B. Britt, Robert A. Koeth, Bruce S. Levison, Yiying Fan, Yuping Wu, and Stanley L. Hazen. 2014. "Prognostic Value of Choline and Betaine Depends on Intestinal Microbiota-Generated Metabolite Trimethylamine-*N*-Oxide." *European Heart Journal* 35(14): 904–10. https://doi.org/10.1093/EURHEARTJ/EHU002

5. Tang, W. H. Wilson, Zeneng Wang, Yiying Fan, Bruce Levison, Jennie E. Hazen, Lillian M. Donahue, Yuping Wu, and Stanley L. Hazen. 2014. "Prognostic Value of Elevated Levels of Intestinal Microbe-Generated Metabolite Trimethylamine-*N*-Oxide in Patients With Heart Failure: Refining the Gut Hypothesis." *Journal of the American College of Cardiology* 64(18): 1908–14. https://doi.org/10.1016/J.JACC.2014.02.617

6. Schiattarella, Gabriele Giacomo, Anna Sannino, Evelina Toscano, Giuseppe Giugliano, Giuseppe Gargiulo, Anna Franzone, Bruno Trimarco, Giovanni Esposito, and Cinzia Perrino. 2017. "Gut Microbe-Generated Metabolite Trimethylamine-*N*-Oxide as Cardiovascular Risk Biomarker: A Systematic Review and Dose-Response Meta-Analysis." *European Heart Journal* 38(39): 2948–56. https://doi.org/10.1093/EURHEARTJ/EHX342

7. Qi, Jiaqian, Tao You, Jing Li, Tingting Pan, Li Xiang, Yue Han, and Li Zhu. 2018. "Circulating Trimethylamine *N*-Oxide and the Risk of Cardiovascular Diseases: A Systematic Review and Meta-Analysis of 11 Prospective Cohort Studies." *Journal of Cellular and Molecular Medicine* 22(1): 185–94. https://doi.org/10.1111/JCMM.13307

8. Heianza, Yoriko, Wenjie Ma, Jo Ann E. Manson, Kathryn M. Rexrode, and Lu Qi. 2017. "Gut Microbiota Metabolites and Risk of Major Adverse Cardiovascular Disease Events and Death: A Systematic Review and Meta-Analysis of Prospective Studies." *Journal of the American Heart Association* 6(7), e004947. https://doi.org/10.1161/JAHA.116.004947

9. Meyer, Katie A., Thomas Z. Benton, Brian J. Bennett, David R. Jacobs, Donald M. Lloyd-Jones, Myron D. Gross, J. Jeffrey Carr, Penny Gordon-Larsen, and Steven H. Zeisel. 2016. "Microbiota-Dependent Metabolite Trimethylamine *N*-Oxide and Coronary Artery Calcium in the Coronary Artery Risk Development in Young Adults Study (CARDIA)." *Journal of the American Heart Association* 5(10), e003970. https://doi.org/10.1161/JAHA.116.003970

10. Tang, W. H. Wilson, Xinmin S. Li, Yuping Wu, Zeneng Wang, Kay Tee Khaw, Nicholas J. Wareham, Max Nieuwdorp, S. Matthijs Boekholdt, and Stanley L. Hazen. 2021. "Plasma Trimethylamine *N*-Oxide (TMAO) Levels Predict Future Risk of Coronary Artery Disease in Apparently Healthy Individuals in the EPIC-Norfolk Prospective Population Study." *American Heart Journal* 236: 80–6. https://doi.org/10.1016/J.AHJ.2021.01.020

11. Fretts, Amanda M., Stanley L. Hazen, Paul Jensen, Matthew Budoff, Colleen M. Sitlani, Meng Wang, Marcia C. de Oliveira Otto, et al. 2022. "Association of Trimethylamine *N*-Oxide and Metabolites With Mortality in Older Adults." *JAMA Network Open* 5(5): e2213242. https://doi.org/10.1001/JAMANETWORKOPEN.2022.13242

12. Wang, Meng, Xinmin S. Li, Zeneng Wang, Marcia C. de Oliveira Otto, Rozenn N. Lemaitre, Amanda Fretts, Nona Sotoodehnia, et al. 2023. "Trimethylamine *N*-Oxide is Associated With Long-Term Mortality Risk: The Multi-Ethnic Study of Atherosclerosis." *European Heart Journal* 44(18): 1608–18. https://doi.org/10.1093/EURHEARTJ/EHAD089

13. Liu, Yarong, and Min Dai. 2020. "Trimethylamine *N*-Oxide Generated by the Gut Microbiota is Associated With Vascular Inflammation: New Insights into Atherosclerosis." *Mediators of Inflammation* 2020: 1–15. https://doi.org/10.1155/2020/4634172

14. Yang, Jae Jeong, Xiao Ou Shu, David M. Herrington, Steven C. Moore, Katie A. Meyer, Jennifer Ose, Cristina Menni, et al. 2021. "Circulating Trimethylamine *N*-Oxide in Association with Diet and Cardiometabolic Biomarkers: An International Pooled Analysis." *The American Journal of Clinical Nutrition* 113(5): 1145–56. https://doi.org/10.1093/AJCN/NQAA430

15. Manor, Ohad, Niha Zubair, Matthew P. Conomos, Xiaojing Xu, Jesse E. Rohwer, Cynthia E. Krafft, Jennifer C. Lovejoy, and Andrew T. Magis. 2018. "A Multi-Omic Association Study of Trimethylamine *N*-Oxide." *Cell Reports* 24(4): 935–46. https://doi.org/10.1016/j.celrep.2018.06.096

16. Bain, Marcus, Gianfranco Fornasini, and Allan Evans. 2005. "Trimethylamine: Metabolic, Pharmacokinetic and Safety Aspects." *Current Drug Metabolism* 6(3): 227–40. https://doi.org/10.2174/1389200054021807

17. Cheung, William, Pekka Keski-Rahkonen, Nada Assi, Pietro Ferrari, Heinz Freisling, Sabina Rinaldi, Nadia Slimani, et al. 2017. "A Metabolomic Study of Biomarkers of Meat and Fish Intake." *The American Journal of Clinical Nutrition* 105(3): 600–8. https://doi.org/10.3945/AJCN.116.146639

18. Gibson, Rachel, Chung Ho E. Lau, Ruey Leng Loo, Timothy M. D. Ebbels, Elena Chekmeneva, Alan R. Dyer, Katsuyuki Miura, et al. 2020. "The Association of Fish Consumption and its Urinary Metabolites With Cardiovascular Risk Factors: The International Study of Macro-/Micronutrients and Blood Pressure (INTERMAP)." *The American Journal of Clinical Nutrition* 111(2): 280–90. https://doi.org/10.1093/AJCN/NQZ293

19. Rohrmann, Sabine, Jakob Linseisen, Martina Allenspach, Arnold Von Eckardstein, and Daniel Müller. 2016. "Plasma Concentrations of Trimethylamine-*N*-oxide are Directly Associated With Dairy Food Consumption and Low-Grade Inflammation in a German Adult Population." *The Journal of Nutrition* 146(2): 283–9. https://doi.org/10.3945/JN.115.220103

20. Fu, Benjamin C., Meredith A. J. Hullar, Timothy W. Randolph, Adrian A. Franke, Kristine R. Monroe, Iona Cheng, Lynne R. Wilkens, et al. 2020. "Associations of Plasma Trimethylamine *N*-Oxide, Choline, Carnitine, and Betaine With Inflammatory and Cardiometabolic Risk Biomarkers and the Fecal Microbiome in the Multiethnic Cohort Adiposity Phenotype Study." *The American Journal of Clinical Nutrition* 111(6): 1226–34. https://doi.org/10.1093/ajcn/nqaa015

21. James, Kristen L., Erik R. Gertz, Eduardo Cervantes, Ellen L. Bonnel, Charles B. Stephensen, Mary E. Kable, and Brian J. Bennett. 2022. "Diet, Fecal Microbiome, and

Trimethylamine *N*-Oxide in a Cohort of Metabolically Healthy United States Adults." *Nutrients* 14(7): 1376. https://doi.org/10.3390/NU14071376

22. Li, Jun, Yanping Li, Kerry L. Ivey, Dong D. Wang, Jeremy E. Wilkinson, Adrian Franke, Kyu Ha Lee, et al. 2022. "Interplay Between Diet and Gut Microbiome, and Circulating Concentrations of Trimethylamine *N*-Oxide: Findings From a Longitudinal Cohort of US Men." *Gut* 71(4): 724–33. https://doi.org/10.1136/GUTJNL-2020-322473

23. Lambert, Deborah M., Orval A. Mamer, Beverley R. Akerman, Luc Choinière, Daniel Gaudet, Pavel Hamet, and Eileen P. Treacy. 2001. "In Vivo Variability of TMA Oxidation is Partially Mediated by Polymorphisms of the FMO3 Gene." *Molecular Genetics and Metabolism* 73(3): 224–9. https://doi.org/10.1006/MGME.2001.3189

24. Doran, Stephen, Muhammad Arif, Simon Lam, Abdulahad Bayraktar, Hasan Turkez, Mathias Uhlen, Jan Boren, and Adil Mardinoglu. 2021. "Multi-Omics Approaches for Revealing the Complexity of Cardiovascular Disease." *Briefings in Bioinformatics* 22(5): bbab061. https://doi.org/10.1093/BIB/BBAB061

25. Dalmeijer, G. W., M. R. Olthof, P. Verhoef, M. L. Bots, and Y. T. van der Schouw. 2008. "Prospective Study on Dietary Intakes of Folate, Betaine, and Choline and Cardiovascular Disease Risk in Women." *European Journal of Clinical Nutrition* 62(3): 386–94. https://doi.org/10.1038/SJ.EJCN.1602725

26. Golzarand, Mahdieh, Parvin Mirmiran, and Fereidoun Azizi. 2022. "Association Between Dietary Choline and Betaine Intake and 10.6-year Cardiovascular Disease in Adults." *Nutrition Journal* 21(1): 1–11. https://doi.org/10.1186/s12937-021-00755-9

27. Nagata, Chisato, Keiko Wada, Takashi Tamura, Kie Konishi, Toshiaki Kawachi, Michiko Tsuji, and Kozue Nakamura. 2015. "Choline and Betaine Intakes are not Associated With Cardiovascular Disease Mortality Risk in Japanese Men and Women." *The Journal of Nutrition* 145(8): 1787–92. https://doi.org/10.3945/JN.114.209296

28. Bidulescu, Aurelian, Lloyd E. Chambless, Anna Maria Siega-Riz, Steven H. Zeisel, and Gerardo Heiss. 2007. "Usual Choline and Betaine Dietary Intake and Incident Coronary Heart Disease: The Atherosclerosis Risk in Communities (ARIC) Study." *BMC Cardiovascular Disorders* 7: 20. https://doi.org/10.1186/1471-2261-7-20

29. Peters, Sanne A. E., Paul Muntner, and Mark Woodward. 2019. "Sex Differences in the Prevalence of, and Trends in, Cardiovascular Risk Factors, Treatment, and Control in the United States, 2001 to 2016." *Circulation* 139(8): 1025–35. https://doi.org/10.1161/CIRCULATIONAHA.118.035550

30. Yousri, Noha A., Khalid A. Fakhro, Amal Robay, Juan L. Rodriguez-Flores, Robert P. Mohney, Hassina Zeriri, and Tala Odeh, et al. 2018. "Whole-Exome Sequencing Identifies Common and Rare Variant Metabolic QTLs in a Middle Eastern Population." *Nature Communications* 9(1): 1–13. https://doi.org/10.1038/s41467-017-01972-9

31. Rhee, Eugene P., Jennifer E. Ho, Ming Huei Chen, Dongxiao Shen, Susan Cheng, Martin G. Larson, Anahita Ghorbani, et al. 2013. "A Genome-Wide Association Study of the Human Metabolome in a Community-Based Cohort." *Cell Metabolism* 18(1): 130–43. https://doi.org/10.1016/J.CMET.2013.06.013

32. Hartiala, Jaana, Brian J. Bennett, W. H. Wilson Tang, Zeneng Wang, Alexandre F. R. Stewart, Robert Roberts, Ruth McPherson, et al. 2014. "Comparative Genome-Wide Association Studies in Mice and Humans for Trimethylamine *N*-Oxide, a Proatherogenic Metabolite of Choline and L-Carnitine." *Arteriosclerosis, Thrombosis, and Vascular Biology* 34(6): 1307–13. https://doi.org/10.1161/ATVBAHA.114.303252

33. Li, Xinmin S., Zeneng Wang, Tomas Cajka, Jennifer A. Buffa, Ina Nemet, Alex G. Hurd, Xiaodong Gu, et al. 2018. "Untargeted Metabolomics Identifies Trimethyllysine, a TMAO-Producing Nutrient Precursor, as a Predictor of Incident Cardiovascular Disease Risk." *JCI Insight* 3(6): e99096. https://doi.org/10.1172/JCI.INSIGHT.99096

34. Phillips, Ian R., and Elizabeth A. Shephard. 2008. "Flavin-Containing Monooxygenases: Mutations, Disease and Drug Response." *Trends in Pharmacological Sciences* 29(6): 294–301. https://doi.org/10.1016/J.TIPS.2008.03.004

35. Fan, Xin, Christopher D. Pericone, Elena Lysenko, Howard Goldfine, and Jeffrey N. Weiser. 2003. "Multiple Mechanisms for Choline Transport and Utilization In *Haemophilus influenzae*." *Molecular Microbiology* 50(2): 537–48. https://doi.org/10.1046/J.1365-2958.2003.03703.X

36. Org, Elin, Margarete Mehrabian, Brian W. Parks, Petia Shipkova, Xiaoqin Liu, Thomas A. Drake, and Aldons J. Lusis. 2016. "Sex Differences and Hormonal Effects on Gut Microbiota Composition in Mice." *Gut Microbes* 7(4): 313–22. https://doi.org/10.1080/19490976.2016.1203502

37. He, Maozhang, Jun Gao, Jinyuan Wu, Yunyan Zhou, Hao Fu, Shanlin Ke, Hui Yang, Congying Chen, and Lusheng Huang. 2019. "Host Gender and Androgen Levels Regulate Gut Bacterial Taxa in Pigs Leading to Sex-Biased Serum Metabolite Profiles." *Front Microbiol* 10, 451217. https://doi.org/10.3389/fmicb.2019.01359

38. Landfald, Bjarne, Jørgen Valeur, Arnold Berstad, and Jan Raa. 2017. "Microbial Trimethylamine-*N*-Oxide as a Disease Marker: Something Fishy?" *Microbial Ecology in Health and Disease* 28(1): 1327309. https://doi.org/10.1080/16512235.2017.1327309

39. Millard, Heather R., Solomon K. Musani, Daniel T. Dibaba, Sameera A. Talegawkar, Herman A. Taylor, Katherine L. Tucker, and Aurelian Bidulescu. 2018. "Dietary Choline and Betaine; Associations With Subclinical Markers of Cardiovascular Disease Risk and Incidence of CVD, Coronary Heart Disease and Stroke: The Jackson Heart Study." *European Journal of Nutrition* 57(1): 51–60. https://doi.org/10.1007/s00394-016-1296-8

40. Konstantinova, Svetlana V., Grethe S. Tell, Stein Emil Vollset, Ottar Nygård, Øyvind Bleie, and Per Magne Ueland. 2008. "Divergent Associations of Plasma Choline and Betaine With Components of Metabolic Syndrome in Middle Age and Elderly Men and Women." *The Journal of Nutrition* 138(5): 914–20. https://doi.org/10.1093/JN/138.5.914

41. Mazidi, Mohsen, Niki Katsiki, Dimitri P. Mikhailidis, and Maciej Banach. 2019. "Dietary Choline is Positively Related to Overall and Cause-Specific Mortality: Results From Individuals of the National Health and Nutrition Examination Survey and Pooling Prospective Data." *British Journal of Nutrition* 122(11): 1262–70. https://doi.org/10.1017/S0007114519001065

42. Jia, Jinzhu, Pan Dou, Meng Gao, Xuejun Kong, Changwei Li, Zhonghua Liu, and Tao Huang. 2019. "Assessment of Causal Direction Between Gut Microbiota-Dependent Metabolites

and Cardiometabolic Health: A Bidirectional Mendelian Randomization Analysis." *Diabetes* 68(9): 1747–55. https://doi.org/10.2337/DB19-0153

43. Liu, Yuan, Juan Qu, Junjie Xu, Aiming Gu, Dezhi Deng, Xiaodan Jia, and Baoxiang Wang. 2023. "Trimethylamine-*N*-oxide: A Potential Biomarker and Therapeutic Target in Ischemic Stroke." *Frontiers in Neurology* 14: 1156879. https://doi.org/10.3389/fneur.2023.1156879

44. Lee, Yujin, Ina Nemet, Zeneng Wang, Heidi T. M. Lai, Marcia C. de Oliveira Otto, Rozenn N. Lemaitre, and Amanda M. Fretts, et al. 2021. "Longitudinal Plasma Measures of Trimethylamine *N*-Oxide and Risk of Atherosclerotic Cardiovascular Disease Events in Community-Based Older Adults." *Journal of the American Heart Association* 10(17), e020646. https://doi.org/10.1161/JAHA.120.020646

45. Zhou, Xiang, Mengchao Jin, Lei Liu, Zongliang Yu, Xiang Lu, and Hao Zhang. 2020. "Trimethylamine *N*-Oxide and Cardiovascular Outcomes in Patients With Chronic Heart Failure After Myocardial Infarction." *ESC Heart Failure* 7(1): 188–93. https://doi.org/10.1002/EHF2.12552

46. Matsuzawa, Yasushi, Hidefumi Nakahashi, Masaaki Konishi, Ryosuke Sato, Chika Kawashima, Shinnosuke Kikuchi, and Eiichi Akiyama, et al. 2019. "Microbiota-Derived Trimethylamine *N*-Oxide Predicts Cardiovascular Risk After STEMI." *Scientific Reports* 9(1): 11647. https://doi.org/10.1038/s41598-019-48246-6

47. Ringel, Clemens, Julia Dittrich, Alexander Gaudl, Paul Schellong, Carl Friedrich Beuchel, Ronny Baber, Frank Beutner, et al. 2021. "Association of Plasma Trimethylamine *N*-Oxide Levels With Atherosclerotic Cardiovascular Disease and Factors of the Metabolic Syndrome." *Atherosclerosis* 335: 62–7. https://doi.org/10.1016/J.ATHEROSCLEROSIS.2021.09.026

48. Andraos, Stephanie, Beatrix Jones, Katherine Lange, Susan A. Clifford, Eric B. Thorstensen, Jessica A. Kerr, Melissa Wake, et al. 2021. "Trimethylamine *N*-Oxide (TMAO) is not Associated With Cardiometabolic Phenotypes and Inflammatory Markers in Children and Adults." *Current Developments in Nutrition* 5(1): nzaa179. https://doi.org/10.1093/CDN/NZAA179

49. Shin, So Youn, Eric B. Fauman, Ann Kristin Petersen, Jan Krumsiek, Rita Santos, Jie Huang, and Matthias Arnold, et al. 2014. "An Atlas of Genetic Influences on Human Blood Metabolites." *Nature Genetics* 46(6): 543–50. https://doi.org/10.1038/ng.2982

50. Hartiala, Jaana A., W. H. Wilson Tang, Zeneng Wang, Amanda L. Crow, Alexandre F. R. Stewart, Robert Roberts, Ruth McPherson, et al. 2016. "Genome-Wide Association Study and Targeted Metabolomics Identifies Sex-Specific Association of CPS1 With Coronary Artery Disease." *Nature Communications* 7: 10558. https://doi.org/10.1038/NCOMMS10558

51. Clare, Constance E., Amey H. Brassington, Wing Yee Kwong, and Kevin D. Sinclair. 2019. "One-Carbon Metabolism: Linking Nutritional Biochemistry to Epigenetic Programming of Long-Term Development." *Annual Review of Animal Biosciences* 7: 263–87. https://doi.org/10.1146/ANNUREV-ANIMAL-020518-115206

52. Costa, Kerry-Ann, Olga G. Kozyreva, Jiannan Song, Joseph A. Galanko, Leslie M. Fischer, and Steven H. Zeisel. 2006. "Common Genetic Polymorphisms Affect the Human Requirement for the Nutrient Choline." *FASEB Journal* 20(9): 1336–44. https://doi.org/10.1096/FJ.06-5734COM

53. Yu, Bing, Yan Zheng, Danny Alexander, Alanna C. Morrison, Josef Coresh, and Eric Boerwinkle. 2014. "Genetic Determinants Influencing Human Serum Metabolome Among African Americans." *PLoS Genetics* 10(3): e1004212. https://doi.org/10.1371/JOURNAL.PGEN.1004212

54. Koeth, Robert A., Bruce S. Levison, Miranda K. Culley, Jennifer A. Buffa, Zeneng Wang, Jill C. Gregory, Elin Org, et al. 2014. "γ-Butyrobetaine is a Proatherogenic Intermediate in Gut Microbial Metabolism of L-Carnitine To TMAO." *Cell Metabolism* 20(5): 799–812. https://doi.org/10.1016/j.cmet.2014.10.006

55. Zhu, Zhaozhong, Yanjun Guo, Huwenbo Shi, Cong Lin Liu, Ronald Allan Panganiban, Wonil Chung, Luke J. O'Connor, et al. 2020. "Shared Genetic and Experimental Links Between Obesity-Related Traits and Asthma Subtypes in UK Biobank." *Journal of Allergy and Clinical Immunology* 145(2): 537–49. https://doi.org/10.1016/J.JACI.2019.09.035

56. Evangelou, Evangelos, Helen R. Warren, David Mosen-Ansorena, Borbala Mifsud, Raha Pazoki, He Gao, Georgios Ntritsos, et al. 2018. "Genetic Analysis of Over 1 Million People Identifies 535 New Loci Associated With Blood Pressure Traits." *Nature Genetics* 50(10): 1412–25. https://doi.org/10.1038/s41588-018-0205-x

57. Kavousi, Maryam, Maxime M. Bos, Hanna J. Barnes, Christian L. Lino Cardenas, Doris Wong, Haojie Lu, Chani J. Hodonsky, et al. 2023. "Multi-Ancestry Genome-Wide Study Identifies Effector Genes and Druggable Pathways for Coronary Artery Calcification." *Nature Genetics* 55(10): 1651–64. https://doi.org/10.1038/S41588-023-01518-4

58. Lu, Feng, Yun Qian, Huizhang Li, Meihua Dong, Yudi Lin, Jiangbo Du, Yuan Lin, et al. 2012. "Genetic Variants on Chromosome 6p21.1 and 6p22.3 are Associated With Type 2 Diabetes Risk: A Case-Control Study in Han Chinese." *Journal of Human Genetics* 57(5): 320–5. https://doi.org/10.1038/JHG.2012.25

59. Holliday, Elizabeth G., Jane M. Maguire, Tiffany Jane Evans, Simon A. Koblar, Jim Jannes, Jonathan W. Sturm, Graeme J. Hankey, et al. 2012. "Common Variants at 6p21.1 are Associated With Large Artery Atherosclerotic Stroke." *Nature Geneteics* 44(10): 1147–51. https://doi.org/10.1038/NG.2397

60. Wang, Zeneng, Adam B. Roberts, Jennifer A. Buffa, Bruce S. Levison, Weifei Zhu, Elin Org, Xiaodong Gu, et al. 2015. "Non-Lethal Inhibition of Gut Microbial Trimethylamine Production for the Treatment of Atherosclerosis." *Cell* 163(7): 1585–95. https://doi.org/10.1016/J.CELL.2015.11.055

61. Chen, Ming Liang, Long Yi, Yong Zhang, Xi Zhou, Li Ran, Jining Yang, Jun-dong Zhu, Qian-yong Zhang, and Man-tian Mi. 2016. "Resveratrol Attenuates Trimethylamine-*N*-Oxide (TMAO)-Induced Atherosclerosis by Regulating TMAO Synthesis and Bile Acid Metabolism via Remodeling of the Gut Microbiota." *mBio* 7(2): e02210–5. https://doi.org/10.1128/MBIO.02210-15

62. Wu, Wei Kai, Chieh Chang Chen, Po Yu Liu, Suraphan Panyod, Ben Yang Liao, Pei Chen Chen, Hsien Li Kao, et al. 2019. "Identification of TMAO-producer Phenotype and Host-Diet-Gut Dysbiosis by Carnitine Challenge Test in Human and

Germe-Free Mice." *Gut* 68(8): 1439–49. https://doi.org/10.1136/GUTJNL-2018-317155

63. Wu, Wei Kai, Suraphan Panyod, Po Yu Liu, Chieh Chang Chen, Hsien Li Kao, Hsiao Li Chuang, Ying Hsien Chen, et al. 2020. "Characterization of TMAO Productivity from Carnitine Challenge Facilitates Personalized Nutrition and Microbiome Signatures Discovery." *Microbiome* 8(1): 162. https://doi.org/10.1186/s40168-020-00912-y

64. Huda, M. Nazmul, Myungsuk Kim, and Brian J. Bennett. 2021. "Modulating the Microbiota as a Therapeutic Intervention for Type 2 Diabetes." *Frontiers in Endocrinology* 12: 632335. https://doi.org/10.3389/FENDO.2021.632335

65. Gessner, Arne, Romina di Giuseppe, Manja Koch, Martin F. Fromm, Wolfgang Lieb, and Renke Maas. 2020. "Trimethylamine-*N*-oxide (TMAO) Determined by LC-MS/MS: Distribution and Correlates in the Population-Based PopGen Cohort." *Clinical Chemistry and Laboratory Medicine* 58(5): 733–40. https://doi.org/10.1515/cclm-2019-1146

66. Aadland, Eli Kristin, Charles Lavigne, Ingvild Eide Graff, Øyvin Eng, Martine Paquette, Asle Holthe, Gunnar Mellgren, Hélène Jacques, and Bjørn Liaset. 2015. "Lean-Seafood Intake Reduces Cardiovascular Lipid Risk Factors in Healthy Subjects: Results From a Randomized Controlled Trial With a Crossover Design." *The American Journal of Clinical Nutrition* 102(3): 582–92. https://doi.org/10.3945/AJCN.115.112086

67. Rath, Silke, Benjamin Heidrich, Dietmar H. Pieper, and Marius Vital. 2017. "Uncovering the Trimethylamine-Producing Bacteria of the Human Gut Microbiota." *Microbiome* 5(1): 1–14. https://doi.org/10.1186/s40168-017-0271-9

68. Kühn, Tilman, Sabine Rohrmann, Disorn Sookthai, Theron Johnson, Verena Katzke, Rudolf Kaaks, Arnold Von Eckardstein, and Daniel Müller. 2017. "Intra-Individual Variation of Plasma Trimethylamine-*N*-Oxide (TMAO), Betaine and Choline Over 1 Year." *Clinical Chemistry and Laboratory Medicine* 55(2): 261–8. https://doi.org/10.1515/CCLM-2016-0374

69. van der Laan, Tom, Tim Kloots, Marian Beekman, Alida Kindt, Anne-Charlotte Dubbelman, Amy Harms, Cornelia M van Duijn, P. Eline Slagboom, and Thomas Hankemeier. 2019. "Fast LC-ESI-MS/MS Analysis and Influence of Sampling Conditions for Gut Metabolites in Plasma and Serum." *Scientific Reports* 9(1): 12370. https://doi.org/10.1038/s41598-019-48876-w

70. Wieberdink, Renske G., M. Arfan Ikram, Albert Hofman, Peter J. Koudstaal, and Monique M. B. Breteler. 2012. "Trends in Stroke Incidence Rates and Stroke Risk Factors in Rotterdam, The Netherlands From 1990 to 2008." *European Journal of Epidemiology* 27(4): 287–95. https://doi.org/10.1007/S10654-012-9673-Y

71. Leening, Maarten J. G., Maryam Kavousi, Jan Heeringa, Frank J. A. Van Rooij, Jolande Verkroost-Van Heemst, Jaap W. Deckers, Francesco U. S. Mattace-Raso, et al. 2012. "Methods of Data Collection and Definitions of Cardiac Outcomes in the Rotterdam Study." *European Journal of Epidemiology* 27(3): 173–85. https://doi.org/10.1007/S10654-012-9668-8

72. Sajjad, A., R. L. Freak-Poli, A. Hofman, S. J. Roza, M. A. Ikram, and H. Tiemeier. 2017. "Subjective Measures of Health and All-Cause Mortality—The Rotterdam Study." *Psychological Medicine* 47(11): 1971–80. https://doi.org/10.1017/S0033291717000381

73. Odink, A. E., A. Van Der Lugt, A. Hofman, M. G. M. Hunink, M. M. B. Breteler, G. P. Krestin, and J. C. M. Witteman. 2010. "Risk Factors for Coronary, Aortic Arch and Carotid Calcification; The Rotterdam Study." *Journal of Human Hypertension* 24(2): 86–92. https://doi.org/10.1038/JHH.2009.42

74. Higgins, Julian P. T., Simon G. Thompson, Jonathan J. Deeks, and Douglas G. Altman. 2003. "Measuring Inconsistency in Meta-Analyses." *The BMJ* 327(7414): 557–60. https://doi.org/10.1136/BMJ.327.7414.557

75. Das, Sayantan, Lukas Forer, Sebastian Schönherr, Carlo Sidore, Adam E. Locke, Alan Kwong, Scott I. Vrieze, et al. 2016. "Next-Generation Genotype Imputation Service and Methods." *Nature Genetics* 48(10): 1284–7. https://doi.org/10.1038/ng.3656

76. Tigchelaar, Ettje F., Alexandra Zhernakova, Jackie A. M. Dekens, Gerben Hermes, Agnieszka Baranska, Zlatan Mujagic, Morris A. Swertz, et al. 2015. "Cohort Profile: LifeLines DEEP, a Prospective, General Population Cohort Study in the Northern Netherlands: Study Design and Baseline Characteristics." *BMJ Open* 5(8): e006772. https://doi.org/10.1136/bmjopen-2014-006772

77. Vojinovic, Dina, Djawad Radjabzadeh, Alexander Kurilshikov, Najaf Amin, Cisca Wijmenga, Lude Franke, M. Arfan Ikram, et al. 2019. "Relationship Between Gut Microbiota and Circulating Metabolites in Population-Based Cohorts." *Nature Communications* 10(1): 5813. https://doi.org/10.1038/s41467-019-13721-1

78. Kurilshikov, Alexander, Carolina Medina-Gomez, Rodrigo Bacigalupe, Djawad Radjabzadeh, Jun Wang, Ayse Demirkan, Caroline I. Le Roy, et al. 2021. "Large-Scale Association Analyses Identify Host Factors Influencing Human Gut Microbiome Composition." *Nature Genetics* 53(2): 156–65. https://doi.org/10.1038/s41588-020-00763-1

79. Cole, J. R., Q. Wang, E. Cardenas, J. Fish, B. Chai, R. J. Farris, A. S. Kulam-Syed-Mohideen, et al. 2009. "The Ribosomal Database Project: Improved Alignments and New Tools for rRNA Analysis." *Nucleic Acids Research* 37(suppl_1): D141–5. https://doi.org/10.1093/NAR/GKN879

80. Streppel, Martinette T., Jeanne Hm de Vries, Saskia Meijboom, Marian Beekman, Anton J. M. de Craen, P. Eline Slagboom, and Edith J. M. Feskens. 2013. "Relative Validity of the Food Frequency Questionnaire Used to Assess Dietary Intake in the Leiden Longevity Study." *Nutrition Journal* 12(1): 75. https://doi.org/10.1186/1475-2891-12-75

81. Voortman, Trudy, Jessica C. Kiefte-de Jong, M. Arfan Ikram, Bruno H. Stricker, Frank J. A. van Rooij, Lies Lahousse, Henning Tiemeier, et al. 2017. "Adherence to the 2015 Dutch Dietary Guidelines and Risk of Non-Communicable Diseases and Mortality in the Rotterdam Study." *European Journal of Epidemiology* 32(11): 993–1005. https://doi.org/10.1007/S10654-017-0295-2

82. Siebelink, Els, Anouk Geelen, and Jeanne H. M. De Vries. 2011. "Self-Reported Energy Intake by FFQ Compared With Actual Energy Intake to Maintain Body Weight in 516 Adults." *British Journal of Nutrition* 106(2): 274–81. https://doi.org/10.1017/S0007114511000067

83. Willer, Cristen J., Yun Li, and Gonçalo R Abecasis. 2010. "METAL: Fast and Efficient Meta-Analysis of Genomewide

Association Scans." *Bioinformatics* 26(17): 2190–1. https://doi.org/10.1093/bioinformatics/btq340

84. Li, J., and L. Ji. 2005. "Adjusting Multiple Testing in Multilocus Analyses Using the Eigenvalues of a Correlation Matrix." *Heredity* 95(3): 221–7. https://doi.org/10.1038/sj.hdy.6800717

85. Van Der Harst, Pim, and Niek Verweij. 2018. "Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease." *Circulation Research* 122(3): 433–43. https://doi.org/10.1161/CIRCRESAHA.117.312086

86. Winkler, Thomas W., Felix R. Day, Damien C. Croteau-Chonka, Andrew R. Wood, Adam E. Locke, Reedik Mägi, Teresa Ferreira, et al. 2014. "Quality Control and Conduct of Genome-Wide Association Meta-Analyses." *Nature Protocols* 9(5): 1192–212. https://doi.org/10.1038/nprot.2014.071

87. Watanabe, Kyoko, Erdogan Taskesen, Arjen Van Bochoven, and Danielle Posthuma. 2017. "Functional Mapping and Annotation of Genetic Associations With FUMA." *Nature Communications* 8(1): 1826. https://doi.org/10.1038/s41467-017-01261-5

88. Kircher, Martin, Daniela M. Witten, Preti Jain, Brian J. O'Roak, Gregory M. Cooper, and Jay Shendure. 2014. "A General Framework for Estimating the Relative Pathogenicity of Human Genetic Variants." *Nature Genetics* 46(3): 310–5. https://doi.org/10.1038/NG.2892

89. Ward, Lucas D., and Manolis Kellis. 2012. "HaploReg: A Resource for Exploring Chromatin States, Conservation, and Regulatory Motif Alterations Within Sets of Genetically Linked Variants." *Nucleic Acids Research* 40(D1): D930–4. https://doi.org/10.1093/NAR/GKR917

90. Boyle, Alan P., Eurie L. Hong, Manoj Hariharan, Yong Cheng, Marc A. Schaub, Maya Kasowski, Konrad J. Karczewski, et al. 2012. "Annotation of Functional Variation in Personal Genomes Using RegulomeDB." *Genome Res* 22(9): 1790–7. https://doi.org/10.1101/GR.137323.112

91. Bulik-Sullivan, Brendan K., Po Ru Loh, Hilary K. Finucane, Stephan Ripke, Jian Yang, Nick Patterson, Mark J. Daly, Alkes L. Price, and Benjamin M. Neale. 2015. "LD Score Regression Distinguishes Confounding from Polygenicity in Genome-Wide Association Studies." *Nature Genetics* 47(3): 291–5. https://doi.org/10.1038/NG.3211

92. Rajakovich, Lauren J., Beverly Fu, Maud Bollenbach, and Emily P. Balskus. 2021. "Elucidation of an Anaerobic Pathway for Metabolism of L-Carnitine-Derived γ-butyrobetaine to Trimethylamine in Human Gut Bacteria." *Proceedings of the National Academy of Sciences of the United States of America* 118(32): e2101498118. https://doi.org/10.1073/pnas.2101498118

93. Cai, Yuan Yuan, Feng Qing Huang, Xingzhen Lao, Yawen Lu, Xuejiao Gao, Raphael N. Alolga, Kunpeng Yin, et al. 2022. "Integrated Metagenomics Identifies a Crucial Role for Trimethylamine-Producing Lachnoclostridium in Promoting Atherosclerosis." *NPJ Biofilms Microbiomes* 8(1): 11. https://doi.org/10.1038/S41522-022-00273-4

94. Kaminski, James, Molly K. Gibson, Eric A. Franzosa, Nicola Segata, Gautam Dantas, and Curtis Huttenhower. 2015. "High-Specificity Targeted Functional Profiling in Microbial Communities With ShortBRED." *PLoS Computational Biology* 11(12): e1004557. https://doi.org/10.1371/journal.pcbi.1004557

95. Zhernakova, Alexandra, Alexander Kurilshikov, Marc Jan Bonder, Ettje F. Tigchelaar, Melanie Schirmer, Tommi Vatanen, Zlatan Mujagic, et al. 2016. "Population-Based Metagenomics Analysis Reveals Markers for Gut Microbiome Composition and Diversity." *Science* 352(6285): 565–9. https://doi.org/10.1126/science.aad3369

96. Langmead, Ben, and Steven L. Salzberg. 2012. "Fast Gapped-Read Alignment With Bowtie 2." *Nature Methods* 9(4): 357–9. https://doi.org/10.1038/nmeth.1923

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1: Phenotypic correlation between metabolites included in the analyses.

Figure S2: Quantile–quantile plots for genome-wide association meta-analyses.

Figure S3: Results of sex-stratified genome-wide association analysis for individual metabolites.

Figure S4: SNP-based heritability and overlap of lead genomic loci identified in the GWAS of TMAO and its precursors.

Figure S5: Top associations of microbial taxonomy with metabolites.

Figure S6: Results of association analysis between metabolites and gut microbial taxa in males and females.

Figure S7: Gender-heterogeneous associations of microbial taxonomy with metabolites.

Figure S8: Results of correlation analysis between metabolites and ratios and food groups.

Table S1: Descriptive characteristics of metabolites and quantification method.

Table S2: Results of association analysis between metabolites and incident cardiovascular disease events. The analyses were performed for overall sample and females and males separately. The associations that surpassed significance threshold (p value $< 8.33 \times 10^{-3}$) are highlighted in bold.

Table S3: Results of association analysis between metabolites and mortality. The analyses were performed for overall sample and females and males separately. The associations that surpassed significance threshold (p value $< 8.33 \times 10^{-3}$) are highlighted in bold.

Table S4: Mendelian randomization analysis evaluating TMAO and its precursors with respect to cardiovascular disease. b, effect; se, standard error; pval,

p value; nsnps, number of SNPs acting as instrumental variables.

Table S5: Genomic inflation factor for the individual studies.

Table S6: Genome-wide significant results (*p* value < 8.33×10^{-9}) for genome-wide association study of TMAO, its precursors betaine, carnitine, choline and deoxycarnitine and TMAO-to-precursor ratios. Associations with HetPval > 0.1 are listed in the table.

Table S7: Index of lead genetic variants and genomic loci these variants are assigned to (novel findings highlighted).

Table S8: All genetic variants in linkage disequilibrium with any of independent significant genetic variants and their functional annotation.

Table S9: Independent lead genetic variants and loci these variants are mapped to in TMAO-to-precursor ratio GWAS.

Table S10: Pleotropic associations for independent significant genetic variants and all genetic variants in LD with these.

Table S11: Genome-wide significant findings in sex-stratified analyses.

Table S12: Results of gene-based analysis. Listed associations passed gene-wide significance level (*p* value < 4.42×10^{-7}).

Table S13: Results of gene-set analysis that showed suggestive evidence of association with metabolites (*p* value = 1/(15485*6) = 1.08×10^{-5}).

Table S14: Results of genome-wide genetic overlap between metabolites and metabolite ratios. Associations with evidence of significant genome-wide genetic overlap were shown in bold.

Table S15: The results of association analyses between metabolites and gut microbial taxa.

Table S16: The results of association analysis between metabolites and gut microbial taxa in females.

Table S17: The results of association analysis between metabolites and gut microbial taxa in males.

Table S18: Association between TMAO metabolic pathway abundances, using shotgun sequencing data from LLD and metabolite concentrations.

Table S19: Mendelian randomization analysis evaluating TMAO and its precursors with regard to gut microbiota. b, effect; se, standard error; pval, *p* value; nsnps, number of SNPs acting as instrumental variables.

Table S20: Results of correlation analyses between TMAO, its precursors, TMAO-to-precursor ratios and food groups.

Table S21: Information on genotyping and analysis.