

1 **Title: Evidence of a causal and modifiable relationship between kidney function and circulating**
2 **trimethylamine *N*-oxide with implications for heart and kidney disorders.**

3
4 **Authors:** ^{†,*}Petros Andrikopoulos^{1,2}, ^{*}Judith Aron-Wisnewsky^{3,4}, ^{*}Rima Chakaroun^{5,6}, ^{*}Antonis
5 Myridakis^{1,7}, ^{*}Sofia K. Forslund^{8,9,10,11,12}, Trine Nielsen¹³, Solia Adriouch³, Bridget Holmes¹⁴, Julien
6 Chilloux¹, Sara Vieira-Silva^{15,16}, Gwen Falony^{15,16}, Joe-Elie Salem⁴, Fabrizio Andreelli^{3,4}, Eugeni
7 Belda^{3,17}, Julius Kieswich^{18,19}, Kanta Chechi^{1,2}, Francesc Puig-Castellvi²⁰, Mickael Chevalier²⁰,
8 Emmanuelle Le Chatelier²¹, Michael T. Olanipekun¹, Lesley Hoyles²², Renato Alves⁸, Gerard
9 Helft^{4,23}, Richard Isnard⁴, Lars Køber²⁴, Luis Pedro Coelho⁸, Christine Rouault³, Dominique
10 Gauguier²⁵, Jens Peter Gøtze²⁴, Edi Prifti^{3,17}, MetaCardis Consortium, Jean-Daniel Zucker^{3,17}, Fredrik
11 Bäckhed^{6,13}, Henrik Vestergaard^{13,26}, Torben Hansen¹³, Jean-Michel Oppert⁴, Matthias Blüher⁵, Jens
12 Nielsen²⁷, Jeroen Raes^{15,16}, Peer Bork^{8,10,28,29}, [#]Muhammad M. Yaqoob^{18,19}, [#]Michael Stumvoll⁵, [#]Oluf
13 Pedersen^{13,30}, [#]S. Dusko Ehrlich³¹, ^{†,#}Karine Clément^{3,4}, ^{†,#}Marc-Emmanuel Dumas^{1,2,20,32}.

14

15 **Affiliations**

16 ¹Section of Biomolecular Medicine, Department of Metabolism, Digestion and Reproduction,
17 Imperial College London, London, UK

18 ²Section of Genomic & Environmental Medicine, National Heart & Lung Institute, Imperial College
19 London, London, UK

20 ³Sorbonne Université, INSERM, Nutrition and obesities; systemic approaches (NutriOmics), Paris,
21 France

22 ⁴Assistance Publique Hôpitaux de Paris, Pitie-Salpêtrière Hospital, Paris France

23 ⁵Medical Department III - Endocrinology, Nephrology, Rheumatology, University of Leipzig Medical
24 Center, Leipzig, Germany

25 ⁶The Wallenberg Laboratory, Department of Molecular and Clinical Medicine, Institute of Medicine,
26 Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

27 ⁷Department of Surgery and Cancer, Imperial College London, London W12 0HS, United Kingdom

28 ⁸Structural and Computational Biology, European Molecular Biology Laboratory, Heidelberg,
29 Germany

30 ⁹Experimental and Clinical Research Center, a cooperation of Charité-Universitätsmedizin and the
31 Max-Delbrück Center, Berlin, Germany

32 ¹⁰Max Delbrück Center for Molecular Medicine (MDC), Berlin, Germany

33 ¹¹Charité University Hospital, Berlin, Germany

34 ¹²DZHK (German Centre for Cardiovascular Research), Partner Site Berlin, Berlin, Germany

35 ¹³Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical
36 Sciences, University of Copenhagen, Copenhagen, Denmark

37 ¹⁴Danone Research, Palaiseau France

38 ¹⁵Laboratory of Molecular Bacteriology, Department of Microbiology and Immunology, Rega
39 Institute, KU Leuven, Leuven, Belgium

40 ¹⁶Center for Microbiology, VIB, Leuven, Belgium

41 ¹⁷Sorbonne Université, IRD, Unité de Modélisation Mathématique et Informatique des Systèmes
42 Complexes, UMMISCO, F-93143, Bondy, France

43 ¹⁸Diabetic Kidney Disease Centre, Renal Unit, Barts Health National Health Service Trust, The Royal
44 London Hospital, London, UK

45 ¹⁹Centre for Translational Medicine and Therapeutics, William Harvey Research Institute, Barts and
46 the London School of Medicine and Dentistry, Queen Mary University of London, London, UK

47 ²⁰European Genomics Institute for Diabetes, EGENODIA, INSERM U1283, CNRS UMR8199, Institut
48 Pasteur de Lille, Lille University Hospital, University of Lille, Lille, France

49 ²¹Metagenopolis, INRA, AgroParisTech, Université Paris-Saclay, Paris, France

50 ²²Department of Bioscience, School of Science and Technology, Nottingham Trent University,
51 Nottingham, UK

52 ²³Institute of Cardiometabolism and Nutrition, ICAN, INSERM, 1166, Paris, France

53 ²⁴Department of Cardiology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

54 ²⁵Université de Paris Cité, INSERM UMR 1124, Paris, France

55 ²⁶Department of Medicine, Bornholms Hospital, Rønne, Denmark

56 ²⁷Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg,
57 Sweden

58 ²⁸Department of Bioinformatics, Biocenter, University of Würzburg, Würzburg, Germany.

59 ²⁹Yonsei Frontier Lab (YFL), Yonsei University, Seoul 03722, South Korea.

60 ²³⁰Center for Clinical Metabolic Research, Gentofte University Hospital, Copenhagen, Denmark

61 ³¹Department of Clinical and Movement Neurosciences, University College London, London NW3
62 2PF, UK

63 ³²McGill Genome Centre, McGill University and Genome Québec Innovation Center, Montréal,
64 Quebec, Canada.

65 ***Joint first authors.**

66 **#Joint Senior authors.**

67 **+Corresponding authors**

68 **Correspondence to** Petros Andrikopoulos (p.andrikopoulos04@imperial.ac.uk), Karine Clément

69 (karine.clement@inserm.fr) or Marc-Emmanuel Dumas (m.dumas@imperial.ac.uk)

70

71 **Abstract**

72 **Objectives:** The host-microbiota co-metabolite trimethylamine *N*-oxide (TMAO) is linked to
73 increased thrombotic and cardiovascular risks. Here we, sought to i) characterize which host
74 variables contribute to fasting serum TMAO levels in real-life settings ii) identify potential actionable
75 therapeutic means related to circulating TMAO.

76 **Design:** We applied “explainable” machine learning, univariate-, multivariate- and mediation
77 analyses of fasting plasma TMAO concentration and a multitude of bioclinical phenotypes in 1,741
78 adult Europeans of the MetaCardis study. We expanded and validated our epidemiological findings
79 in mechanistic studies in human renal fibroblasts and a murine model of kidney fibrosis following
80 TMAO exposure.

81 **Results:** Next to age, kidney function was the primary variable predicting circulating TMAO in
82 MetaCardis, with microbiota composition and diet playing minor, albeit significant roles. Mediation
83 analysis revealed a causal relationship between TMAO and kidney function decline that
84 strengthened at more severe stages of cardiometabolic disease. We corroborated our findings in
85 preclinical models where TMAO exposure augmented conversion of human renal fibroblasts into
86 myofibroblasts and increased kidney scarring *in vivo*. Mechanistically, TMAO aggravated kidney
87 fibrosis due to ERK1/2 hyperactivation synergistically with TGF- β 1 signaling. Consistent with our
88 findings, patients receiving next-generation glucose-lowering drugs with reno-protective properties,
89 had significantly lower circulating TMAO when compared to propensity-score matched control
90 individuals.

91 **Conclusion:** After age, kidney function is the major determinant of fasting circulating TMAO in
92 adults. Our findings of lower TMAO levels in individuals medicated with reno-protective anti-diabetic
93 drugs suggests a clinically actionable intervention for decreasing TMAO-associated excess
94 cardiovascular risk that merits urgent investigation in human trials.

95

96

97 **Data availability statement**

98 Raw shotgun sequencing data that support the findings of this study have been deposited in the
99 European Nucleotide Archive with accession codes PRJEB37249, PRJEB38742, PRJEB41311 and
100 PRJEB46098. Serum NMR and urine NMR metabolome data have been uploaded to Metabolights
101 with accession number MTBLS3429; serum GC-MS and isotopically quantified serum metabolites
102 (UPLC-MS/MS) are available from MassIVE with accession numbers MSV000088042 and
103 MSV000088043, respectively.

104

105 **Introduction**

106 Over the past two decades, the central role of the commensal gut microbiota in pathologies such as
107 atherosclerosis and type-2 diabetes (T2D) has gained prominence¹. The microbiota can influence
108 host pathophysiology by producing molecules that directly alter metabolism and/or modulate
109 cellular signaling either locally in the gut or systemically via the circulation^{1,2}.

110 Trimethylamine *N*-oxide (TMAO) is the phase-one liver *N*-oxide of trimethylamine (TMA).
111 TMA is a product of the microbial^{3,4,5} (predominately Firmicutes) metabolism of
112 phosphatidylcholine^{6,7}, choline⁷ and *L*-carnitine^{8,9,10}, components of the high-fat, high-red meat
113 western diet. TMA is taken up from the gut via the hepatic portal vein and *N*-oxidised into TMAO by
114 host flavin mono-oxygenase 3. High circulating levels of TMAO have been linked to increased
115 thrombotic and cardiovascular risk in animal and human studies, even after adjustment for known
116 cardiovascular risk factors^{6,7,8,11}. TMAO is therefore proposed to mediate the higher cardiovascular
117 risk associated with high red meat and fat intake¹². Fish, the consumption of which is associated with
118 reduced incidence of cardiovascular disease¹², is also a rich source of TMAO¹³. Dietary TMAO is
119 subject to retro-conversion: *i.e.*, it can undergo microbial reduction to TMA by *Enterobacteriaceae*
120 followed by hepatic conversion back to TMAO¹⁴. Concurrent with diet and microbiota composition,
121 TMAO plasma levels also reflect age¹⁵, sex¹⁶, kidney function^{17,18,19} and chronic diseases^{5,20}. To date,
122 the relative contribution of each of these factors to circulating TMAO levels and, therefore, elevated
123 cardiovascular risk remains unclear. Understanding how serum TMAO levels are regulated in real-life
124 settings could uncover host-TMAO mechanistic targets and identify modifiable and actionable
125 therapeutic factors to lower circulating TMAO levels.

126 Here, by using a data-driven “explainable” machine learning (ML) strategy²¹, multivariate
127 and univariate analyses of epidemiological data and mechanistic studies in cultured cells and
128 rodents, we sought to objectively identify variables influencing serum TMAO levels in participants of
129 the European multi-center MetaCardis study. Moreover, by taking advantage of the unique cross-
130 sectional MetaCardis design, we queried how variables influencing circulating TMAO manifest at

131 different stages of cardiometabolic disease. Capitalizing on the ML analysis, we aimed to identify (i)
132 novel host TMAO-related mechanistic targets and (ii) a rationale for future interventions that could
133 reduce circulating TMAO levels and thereby decrease associated excess cardiovascular risk
134 **(Suppl.Fig.1)**.

135 In epidemiological studies, we observed that kidney function is the main modifiable factor
136 consistently regulating fasting serum TMAO levels and our preclinical studies align with the
137 suggestion that elevated circulating TMAO adversely affects kidney function, by increasing kidney
138 fibrotic injury. Interestingly, patients with T2D in the cohort prescribed new-generation anti-
139 diabetics (GLP-1 Receptor Agonists²²; GLP-1RAs) with evidenced reno-protective effects²³ had lower
140 serum circulating TMAO levels when compared to propensity-score matched controls.

141 Results

142 In a first analysis, we investigated how TMAO levels change depending on disease classification in
143 the MetaCardis population. We confirmed circulating TMAO significantly increased with
144 cardiometabolic disease severity, in line with previous reports^{4,20} (**Suppl.Fig.2A**). To identify
145 determinants of circulating TMAO manifesting at prodromal stages of cardiometabolic disease, we
146 focused on the subset of the MetaCardis cohort termed MetaCardis Body Mass Index Spectrum
147 subset (BMIS²³; N=837) comprising obese/overweight individuals presenting with a range of
148 metabolic syndrome features but not overt T2D or ischaemic heart disease (IHD) (**Suppl.Table.1**).

149 Fasting serum concentration of TMAO is associated with worse cardiometabolic profiles in BMIS 150 MetaCardis population

151 We explored how TMAO correlated with bioclinical variables related to cardiometabolic
152 health in BMIS individuals. In this group, circulating TMAO was associated with reduced values of
153 estimated glomerular filtration rate (eGFR, Spearman rho=-0.124, pFDR=0.06) and higher fasting
154 concentrations of uric acid (rho=0.114, pFDR=0.09) after adjustment for age, sex, country of
155 recruitment (demographics thereafter) and body mass index (BMI) (**Suppl.Fig.2B, Suppl.Table.2**).
156 Higher TMAO also positively associated with indicators of central adiposity, including BMI
157 (rho=0.107, pFDR=0.09), visceral body fat rating (rho=0.122, pFDR=0.09) and waist circumference
158 (rho=0.119, pFDR=0.09), after demographics adjustment (**Suppl.Fig.2C, Suppl.Table.3**). Moreover,
159 BMIS individuals with hypertension (systolic blood pressure over 140mmHg, or receiving therapy for
160 high blood pressure; Methods) had higher plasma TMAO ($P=0.01$, Mann-Whitney test; **Suppl.Fig.2D**).
161 In line with previous studies^{13,14,20}, objectively dividing BMIS participants into TMAO clusters using k-
162 means²⁵ revealed that those in the cluster with the highest TMAO levels had consistently worse
163 cardiometabolic traits and were significantly older when compared to those in the cluster with
164 lowest TMAO levels (**Suppl.Fig.2E, Suppl.Table.4**). Traits of cardiometabolic risk included altered
165 eGFR, elevated liver enzymes and systolic blood pressure (**Suppl.Fig.2F-I**).

166 **Age and altered kidney function variables are the main drivers of circulating TMAO in BMIS.**

167 To better understand which variables (**Suppl.Table.5**) affect circulating TMAO most, we
168 trained extreme gradient-boosted decision-tree models. We used five-fold cross-validation to
169 predict the Explained Variance (EV) of each variable group on plasma TMAO in the left-out BMIS
170 participants after 100 iterations (**Figure 1A, Suppl.Table.6**). Microbiota composition alone
171 performed poorly (EV 2%) whilst diet, another purported major contributor to TMAO production,
172 explained less than 5% of TMAO variance. Serum metabolomics, excluding TMAO and its precursor
173 TMA, was the best predictor (EV 12%) with demographics second, explaining 10% of circulating
174 TMAO variance (**Suppl.Table.7**). The full model, containing all variable groups, accounted, on
175 average, for 18.4% of TMAO variability. The averaged predicted TMAO values by the full model
176 significantly correlated with measured TMAO values ($\rho=0.473$, $P<2.2\times 10^{-16}$; **Suppl.Fig.2J**,
177 **Suppl.Table.8**).

178 We next assessed the independent contribution of each variable group to the predictive
179 power of the full model by training algorithms as above but removing one feature group at a time
180 for 100 iterations. Almost 40% of the explainable variance of the full model (set to 100%) was
181 contributed by serum metabolomic variables, with biological, dietary and microbiota taxonomic
182 variables adding 5.5%, 4% and 3.4%, respectively, independently explained variance (**Figure 1B**).
183 Other variable categories displayed negligible contribution to prediction suggesting considerable
184 information overlap with the metabolomic, metagenomic, biological and dietary datasets
185 (**Suppl.Table.9**).

186 Using feature attribution analysis (SHapley Additive exPlanations; SHAP^{21,26}), we assessed
187 how individual variables drive TMAO-predicting models. SHAP analysis identified 24 variables that
188 contributed, more than 4% of the regularized TMAO standard deviation (SD) to model outcomes. Of
189 those, age affected predictions the most, followed by eGFR, urinary betaine, percentage of visceral
190 body fat and serum butyryl-carnitine (**Figure 1C, Suppl.Table.10**). Besides eGFR, additional variables

191 associated with kidney function including plasma urea²⁷, the uremic toxin *p-cresol*²⁸ or kidney
192 function decline consequences: i.e. serum albumin²⁷ were among those affecting model outcomes
193 the most. This analysis suggests that kidney function is a major determinant of circulating TMAO.
194 The impact of eGFR on model outcomes for BMIS individuals was bimodal with values over 90
195 mL/min/1.73m², the clinical cut-off value for normal kidney function in adults²⁵, predicting reduced
196 plasma TMAO and lower values resulting in increased predicted circulating TMAO (**Figure 1D**).

197 We next trained algorithms predicting TMAO using the 24 variables identified by our SHAP
198 analysis (“top SHAP” model) and compared it to models trained by traditional clinical risk factors²⁹ or
199 the full model (**Figure 1E**). The “top SHAP” model significantly ($P < 2.2 \times 10^{-16}$, Mann-Whitney test)
200 improved predictions when compared to the full model (EV 21% vs. 18%, on average), presumably
201 by removing noise, supporting the importance of the variables identified by the SHAP analysis.

202 **Diet and microbiota composition have a modest impact on fasting serum TMAO levels in the BMIS** 203 **cohort**

204 Our ML models suggested that diet plays a modest role in determining circulating TMAO. To
205 corroborate this, we assessed how consumption of food items rich in TMAO dietary precursors³
206 affected circulating TMAO in BMIS participants with a Food Frequency Questionnaire (FFQ; N=763).
207 After correcting for demographics and BMI, with the exception of oily fish ($\rho=0.125$, $p\text{FDR}=0.03$),
208 circulating TMAO neither significantly associated with habitual consumption of red meat, eggs or
209 milk (**Suppl.Table.11**) nor with the estimated dietary intake of the micronutrients choline, carnitine
210 or betaine (**Suppl.Table.12**), broadly in agreement with a previous study³ (**Figure 2A**). Additionally,
211 no significant correlations were found between consumption of these food items and the serum
212 levels of TMAO precursors choline, betaine or γ -butyrobetaine³⁰. Conversely, circulating TMAO
213 positively correlated with plasma choline, betaine and γ -butyrobetaine, implying an association
214 between the serum concentrations of TMAO precursors.

215 To identify putative microbial taxa influencing serum plasma TMAO, we performed
216 multivariate and univariate analyses. Principal coordinates analysis of Bray-Curtis dissimilarity
217 matrices at the species level for BMIS individuals stratified into TMAO clusters revealed a significant
218 difference in microbiota composition between clusters ($P=0.033$; **Figure 2B**). Consistent with
219 previous reports^{3,31} and our ML models, multivariate PERMANOVA analysis uncovered a significant,
220 albeit weak, association between circulating TMAO and microbiota composition ($P=0.001$; $R^2 =$
221 0.009 ; **Figure 2B**) after demographics adjustment. TMAO levels significantly associated with absolute
222 abundance of 65 bacterial species (corrected for bacterial load) in BMIS individuals after correcting
223 for demographics and BMI, primarily (44/65) of the Firmicutes phylum (**Suppl.Table.13, Figure 2C**).
224 In agreement with Li and colleagues³, we did not identify any overlap between bacteria associated
225 with circulating TMAO and higher red meat, milk or eggs consumption in BMIS participants ($N=761$;
226 **Figure 2C, Suppl.Table.13**). Contrasting low and high TMAO clusters revealed 215 differentially
227 abundant bacterial species between these two groups (**Figure 2D, Suppl.Table.14**).

228 We further analyzed the impact on circulating TMAO of the only species, an unknown
229 bacterium taxonomically closely related to ruminococci (CAG01909), that contributed at least 4% of
230 the TMAO SD in our prediction models and was significant ($pFDR<0.05$) for both correlation and
231 differential abundance analyses (**Figure 2E**). CAG01909 was more abundant and prevalent in the
232 high TMAO cluster when compared to the low (**Suppl.Fig.3A-B**). Finally, BMIS individuals harboring
233 CAG01909 had significantly higher circulating TMAO than those that lacked it (**Suppl.Fig.3C**), the
234 CAG01909 abundance significantly correlating with TMAO concentration (**Suppl.Fig.3D**).

235 This analysis corroborated the ML models suggesting a small, albeit significant, contribution
236 of microbiota variations on circulating TMAO and identified a bacterium newly associated with
237 higher serum TMAO in BMIS.

238 **Signatures predicting circulating TMAO shift in different disease groups**

239 To identify putative common variables driving circulating TMAO levels in different disease
240 groups, we trained ML algorithms for T2D (N=561) and IHD sub-cohorts (N=356). Similar to BMIS,
241 microbiota composition and dietary variables alone performed poorly, explaining on average 2% and
242 1% of TMAO variance, respectively, in T2D and IHD individuals (**Figure 3A and Suppl.Fig.4A**
243 respectively). For the T2D cohort, serum metabolomics was again the best predictor (EV 12.8%)
244 followed by bioclinical (EV 9.6%) and demographic variables (EV 8.7%) with the full model
245 accounting for 16.2% of TMAO variance (**Suppl.Table.17**). For IHD, individual feature categories,
246 except for serum metabolomics (EV 19.5%) performed poorly, with the full model accounting for
247 16.1% of circulating TMAO variance (**Suppl.Fig.4A, Suppl.Table.18**), possibly reflecting reduced
248 power due to smaller sample size and/or IHD group heterogeneity. Feature attribution analysis
249 revealed 19 variables in the T2D cohort and 10 variables in the IHD cohort that contributed more
250 than 4% of regularized TMAO SD to model outcome (**Figure 3B and Suppl.Fig.4B** respectively;
251 **Suppl.Table.19, Suppl.Table.20**). In patients with T2D, model outcome was mostly affected by eGFR
252 followed by age and the serum concentrations of the uremic toxin *p*-cresol²⁸ and *D*-threitol, also
253 indicative of kidney function³². In patients with IHD, serum butyryl-carnitine, followed by age,
254 alternative healthy eating score (aHEI) related to the ratio of white to red meat intake and the levels
255 of the proinflammatory cytokine IP10, a marker of adverse cardiac remodelling³³ were the top
256 variables. As in BMIS, the “top SHAP” models significantly improved predictions (EV 24% *versus*
257 16.2% for T2D; **Figure 3C** and EV 21.3% *versus* 16.1% for IHD; **Suppl.Fig.4C**), when compared to
258 models trained with clinical risk factors or the full model.

259 Eight variables contributing more than 4% of TMAO SD were shared between BMIS and T2D,
260 including, eGFR, age, and *p*-cresol whilst there were six common features between BMIS and IHD
261 (including eGFR, urea, age and butyryl-carnitine). Only three variables (eGFR, age and butyryl-
262 carnitine) strongly contributed as predictors across all three disease groups (**Figure 3D**).

263 We thus identified age and kidney function as the prominent variables influencing fasting
264 circulating TMAO levels. To further substantiate this relationship, we evaluated whether data fit a
265 model where eGFR causally mediates the increase of TMAO concentration with age. Under this
266 model, kidney function achieved significance as mediating ~20% of the positive association between
267 age and circulating TMAO levels in BMIS with its impact further strengthening in the T2D (55%) and
268 IHD (51%) disease groups (**Suppl.Fig.4D**), thus corroborating the ML analysis.

269 **Microbiota composition modestly affects plasma TMAO levels in T2D.**

270 In patients with T2D, ML and SHAP analyses revealed an inverse association between fecal
271 *Romboutsia timonensis* abundance and circulating TMAO (**Figure 3B**), whilst no taxa strongly
272 influenced TMAO predictions in the IHD sub-cohort (**Suppl.Fig.4B**). *R. timonensis* inversely associated
273 with TMAO concentration ($\rho=-0.140$, $P=0.0009$; **Suppl.Fig.5A**), T2D individuals with detectable *R.*
274 *timonensis* had significantly lower TMAO levels (**Suppl.Fig.5B**) and *R. timonensis* was depleted in T2D
275 compared to BMIS participants (**Suppl.Fig.5C**), in line with circulating TMAO (**Suppl.Fig.2A**).
276 Conversely, in T2D and IHD CAG01909 abundance or presence, unlike BMIS, did not associate with
277 higher circulating TMAO (**Suppl.Fig.5D-5F**). Further investigation revealed that CAG01909 abundance
278 was inversely associated with metformin intake ($\rho=-0.141$, $pFDR=0.042$; **Suppl.Fig.5G**,
279 **Suppl.Table.21**), in accordance with the well-documented effect of drugs on the gut microbiome³⁴.
280 Irrespective of TMAO associations with individual species, similarly to BMIS, multivariate
281 PERMANOVA analysis, uncovered a significant, albeit weak, association between circulating TMAO
282 and microbiota composition ($P=0.001$; $R^2=0.005$) after adjustment for demographics.

283 Collectively, these analyses suggest that the composition of the gut microbiota may only
284 modestly influence circulating levels of TMAO in patients with T2D and in subjects with IHD or T2D
285 abundance of the CAG01909 taxon is not linked with serum TMAO.

286 **Inference analysis suggests that TMAO may causally mediate the decline of eGFR with age**

287 ML analysis identified age and kidney function as the most prominent variables influencing
288 TMAO. Therefore, we tested the inverse relationship: *i.e.*, whether TMAO may mediate eGFR decline
289 with age³⁵. In BMIS, TMAO modestly, albeit significantly, modulated the inverse relationship
290 between age and eGFR (mediation effect 3%, $P=0.018$), whilst in T2D and IHD the impact of TMAO
291 on kidney function decline with age strengthened (mediation effect 13%, $P<2.2\times 10^{-16}$ and 7.1%,
292 $P<2.2\times 10^{-16}$; **Figure 3E**). Mediation analysis thus indicated that TMAO, far from being a bystander,
293 directly and adversely affects kidney function. Our finding is in agreement with a prospective study
294 reporting that baseline TMAO positively associated with rates of eGFR decline¹⁹ and with studies in
295 animal models of chronic kidney disease (CKD) where TMAO diet-supplementation increased kidney
296 injury^{36,37} and reduced eGFR.

297 Interestingly, mediation analysis showed that TMAO's adverse effect on kidney function
298 strengthened at more severe stages of cardiometabolic disease, implying TMAO synergy with
299 existing pathology (thereby providing "2-hit model"). Accordingly, we next employed preclinical
300 models of kidney injury to gain mechanistic insights into the potential interplay between TMAO and
301 kidney function and the putative molecular nature of the 2nd hit.

302 **TMAO increases trans-differentiation of human primary renal fibroblasts into myofibroblasts in** 303 **conjunction with TGF- β 1 signaling**

304 Based on TMAO's detrimental association with kidney damage of diverse aetiologies^{17,18}
305 similarly to renal fibrosis³⁸, we investigated the impact of TMAO on human primary renal fibroblasts
306 (HRFs). In HRFs, unlike platelets³⁹, TMAO stimulation resulted in rapid $[Ca^{2+}]_i$ increase (**Figure 4A**). In
307 endothelial cells the ERK1/2 pathway is activated in a Ca^{2+} -dependent manner^{40,41} and ERK1/2
308 activation exacerbates renal fibrosis⁴². Therefore, we investigated whether a similar pathway
309 operates in HRFs. TMAO challenge increased phospho-ERK1/2 levels in a time- (**Figure 4B**) and dose-
310 dependent (**Figure 4C, Suppl.Fig6A**) manner, at concentrations (10-100 μ M) relevant to human
311 disease (up to 150 μ M in patients with CKD)¹⁷. TMAO-induced ERK1/2 activation was inhibited by a

312 MEK inhibitor, suggesting the activation occurs downstream of Ras-Raf-MEK⁴². Additionally, ERK1/2
313 activation in response to TMAO was suppressed when intracellular Ca²⁺ was chelated (**Suppl.Fig.6B**)
314 or extracellular Ca²⁺ was removed (**Suppl.Fig.6C**). Moreover, Ca²⁺ influx was sufficient to activate
315 ERK1/2 in HRFs (**Suppl.Fig.6D**); indicating that TMAO-induced Ca²⁺ influx was required for ERK1/2
316 activation. We have reported that ERK1/2 phosphorylation was required for HRF trans-
317 differentiation to myofibroblasts in response to TGF- β 1⁴². Unlike short-term application (up-to
318 30min), TMAO stimulation for 48h had minimal effect on ERK1/2 or SMAD3 phosphorylation (**Figure**
319 **4D, Suppl.Fig.6E-F**) or on the expression of the myofibroblast marker α SMA (**Figure 4E,**
320 **Suppl.Fig.6G**). Conversely, TMAO dose-dependently augmented ERK1/2 pathway activation and
321 myofibroblast trans-differentiation, when co-administered with TGF- β 1 in comparison to TGF- β 1 or
322 TMAO alone, without affecting SMAD3 phosphorylation (**Figure 4D-E, Suppl.Fig.6E-F**).

323 **TMAO increases renal fibrosis: evidence from an intervention in mice**

324 To corroborate our *in vitro* findings suggesting that TMAO directly increases myofibroblast
325 trans-differentiation and establish *in vivo* relevance, we performed Unilateral Ureter Obstruction
326 (UUO) surgery in mice, a model assessing renal fibrosis in the absence of other co-morbidities⁴³ that
327 could be also impacted by TMAO. Mice were fed a choline (1% w/w)- or TMAO (0.12% w/w)-
328 supplemented diet for 6 weeks¹⁸. Subsequently, the ureter of one kidney was ligated whilst the
329 other kidney remained unobstructed (**Suppl.Fig.7A**). In the unobstructed (control) kidneys α SMA
330 staining was not affected by diet. As expected, five days of UUO resulted in a significant increase in
331 renal α SMA staining of the injured kidney. However, unlike the unobstructed kidneys, TMAO or
332 choline diet supplementation resulted in significantly more myofibroblast expansion (**Figure 4F-G**).
333 Western blotting of kidney lysates corroborated augmented α SMA and vimentin (another marker of
334 myofibroblasts⁴²) expression in injured kidneys of mice receiving TMAO or choline diets (**Figure 4H-**
335 **J**). Similar to α SMA, collagen deposition or macrophage infiltration, another indicator of fibrotic
336 kidney damage³⁸, were not affected by the TMAO or choline diets in the unobstructed kidneys
337 (**Suppl.Fig.7B-E**). Conversely, collagen and macrophage staining of kidney slides from UUO kidneys of

338 mice consuming TMAO or choline diets were significantly enhanced compared to controls. Thus, in
339 our experiments in a murine model of kidney fibrosis, TMAO or choline diet supplementation
340 resulted in “hyperactivation” of ERK1/2, mTORC1 and SMAD3 pro-fibrotic signalling^{38,42}
341 (**Suppl.Fig.7F**).

342 Collectively, our *in vivo* and *in vitro* findings are consistent with TMAO aggravating kidney
343 fibrosis due to ERK1/2 hyperactivation synergistically probably with activation of the TGF- β 1-
344 mediated SMAD3 pathway causing a secondary hit in our disease model.

345 **Glucagon-Like Peptide-1 Receptor analogues (GLP-1RAs) intake associates with lower serum**
346 **TMAO concentration in MetaCardis T2D participants.**

347 Given the strong bi-directional connection we uncovered between serum TMAO and eGFR,
348 we hypothesized that use of reno-protective medication could be linked with lower circulating
349 TMAO. To identify suitable drugs, we trained algorithms to predict eGFR in MetaCardis patients with
350 T2D, where TMAO appears to have the biggest impact on kidney function, using prescribed
351 medication as input. SHAP analysis revealed that anti-hypertensive and anti-cholesterol treatments
352 had a negative impact on predicted eGFR, probably reflecting more advanced disease (**Figure 5A**).
353 Conversely, anti-diabetic treatments had a positive effect, with GLP-1RAs being the drugs with the
354 biggest positive impact on predicted eGFR (**Figure 5A, Suppl.Table.22**). In accordance with their
355 documented reno-protective effect even in T2D patients with no preexisting nephropathy²³, patients
356 with T2D receiving GLP-1RAs had significantly lower circulating TMAO (**Figure 5B**) than individuals
357 matched for age, sex, hypertension and disease group (**Suppl.Fig.8A-E, Suppl.Table.23**). This
358 exploratory analysis suggests that GLP-1RAs may reduce serum TMAO concentration and thereby
359 the associated higher risk of IHD.

360

361 **Discussion**

362 With our approach combining epidemiological studies, explorative cellular experiments and murine
363 interventions, we demonstrate that kidney function is the main modifiable factor consistently
364 regulating fasting serum TMAO concentrations and that TMAO adversely impacts on eGFR, at least
365 partially, by increasing kidney scarring synergistically with existing pathophysiology elevating kidney
366 tissue TGF- β 1 signaling. Consistent with our findings, use of reno-protective drugs, GLP-1RAs²³, was
367 associated with lower circulating TMAO in MetaCardis participants with T2D (**Figure 5C**)

368 Irrespective of cardiometabolic disease severity, gut microbiota composition had a modest
369 ($R^2 < 0.01$), albeit significant, association with serum TMAO levels in the MetaCardis population, in
370 agreement with recent human studies^{3,31}. We replicated associations of Firmicutes^{3,4,5} with
371 circulating TMAO and uncovered novel associations between an unknown bacterium (CAG01909)
372 and *R. timonensis*, a marker of less diverse diet in autistic children⁴⁴. However, taxa associated with
373 TMAO diverged between disease modalities. This could at least, partially, be due to medication,
374 suggesting that targeting TMAO production at the species level would be ineffective.

375 In agreement with a recent study³, we did not find any significant association between
376 habitual consumption of red meat and fasting serum levels of TMAO. With a few notable
377 exceptions³, the contribution of diet and in particular red meat and *L*-carnitine to TMAO levels,
378 showing an increase following intervention, has predominately been examined in metabolically
379 healthy volunteers^{8,9,10}. In such interventions, often *L*-carnitine has been provided as a dietary
380 supplement, which is poorly absorbed in the small intestine (~12%), as opposed to dietary *L*-
381 carnitine (~71%)⁴⁵, and therefore may be more available for microbial catabolism in the upper gut
382 and large intestine, leading to overestimating its role in TMA, and thereby TMAO production. Our
383 observations, similar to the report by Li and colleagues³, suggests that in real-life settings where
384 individuals habitually consume meat (75 to 233 g/day in European adults⁴⁶), this contributes
385 minimally to fasting circulating TMAO variability, possibly limiting the isolated effect of dietary
386 manipulation on TMAO levels in real-life, aside strict vegans or vegetarians.

387 Further, our analyses revealed that the multi-omics signature associated with higher
388 circulating TMAO concentrations and, therefore, elevated cardiovascular risk, shifted as disease
389 progressed with only three variables (age, eGFR and serum butyryl-carnitine) consistently being
390 strong contributors across models built with the BMIS, T2D and IHD disease groups. This is in
391 agreement with our recent report that the majority of markers of dysmetabolism manifest early
392 during cardiometabolic disease development²⁹. Accumulation of butyryl-carnitine has been
393 associated with abnormal mitochondrial lipid β -oxidation in T2D⁴⁷, suggesting a link between TMAO
394 and mitochondrial dysfunction consistent with the finding that TMAO binds to protein kinase R-like
395 endoplasmic reticulum kinase and increases mitochondrial stress⁴⁸.

396 Far from being a bystander, mediation analysis indicated that TMAO significantly accelerated
397 the rate of kidney function decline by age, with its effect increasing in more severe disease stages.
398 Consistent with this, our pre-clinical work uncovered that TMAO primed renal fibroblasts for
399 conversion to myofibroblast, the primary collagen-producing cells in the kidney³⁸ and contributed to
400 renal fibrosis, a hallmark of kidney damage irrespective of the underlying cause that significantly
401 contributes to eGFR decline³⁸. However, TMAO insult alone was not sufficient to convert renal
402 fibroblasts to myofibroblasts requiring synergy with a pre-existing pathological state, *i.e.*, pro-fibrotic
403 signaling (“second-hit”) in the form of TGF- β 1 stimulation, the most prominent pro-fibrotic
404 cytokine³⁸. TGF- β 1 expression is increased in ligated UOO mouse kidneys⁴⁹ or kidneys of patients
405 with diabetic nephropathy⁵⁰ and circulating TGF- β 1 is elevated in conditions that are risk factors for
406 CKD including hypertension, dyslipidemia and T2D⁵¹. Therefore, our results suggest that TMAO
407 accelerates eGFR decline in concert with existing pro-fibrotic signaling (*i.e.* TGF- β 1; “second-hit”),
408 consistent with observations in animal models of CKD^{36,37} and humans^{18,19}.

409 Supporting our assertion that TMAO accelerates cardiovascular disease progression together
410 with existing pathology, at least partially by impacting the kidney, TMAO increased all-cause
411 mortality only in individuals with eGFR<90 mL/min/1.73m²⁵² and circulating TMAO levels in healthy
412 adults were not indicative of future atherosclerotic burden⁵³.

413 TMAO is an independent risk factor of cardiovascular morbidity and mortality in patients
414 with established IHD or CKD^{6,17,18} with accumulating evidence suggesting direct causal effects⁴⁰. Our
415 work suggests that reno-protective strategies could potentially lower circulating TMAO and
416 therefore preserve kidney function in individuals with high circulating TMAO in the presence of risk
417 factors known to increase TGF- β 1 signaling (i.e hypertension and T2D⁵¹). Indeed, we observed that
418 patients with T2D prescribed GLP-1RAs, drugs with documented reno-protective and beneficial
419 cardiovascular effects^{22,23}, had significantly lower fasting plasma TMAO levels than propensity-score
420 matched controls.

421 **Conclusion**

422 Collectively, our findings demonstrate that eGFR, irrespective of disease stage, is the primary
423 modifiable modulator of circulating TMAO, which then by promoting renal fibrosis in conjunction
424 with established pathophysiology (“second-hit” model) further negatively impacts kidney function.
425 Accordingly, we observe that intake of GLP-1RAs is associated with lower circulating TMAO levels,
426 thereby potentially moderating the adverse effect of TMAO on kidney function and suggesting a
427 putative mechanism for these drugs observed reno-protection in large pharmaceutical trials²³.

428 Our work, conceptually advances understanding of how TMAO levels (and therefore
429 associated cardiovascular risk) are regulated in humans with a wide range of Cardiometabolic
430 disease burden in real-life settings. Additionally, we uncover a direct mechanistic link between
431 TMAO and renal fibrosis in conjunction with existing co-morbidities (known to elevate TGF- β 1
432 signaling) such as hypertension and T2D. Furthermore, our findings suggest that therapeutic
433 modalities preserving kidney function could markedly benefit and reduce cardiovascular risk in
434 individuals with high circulating TMAO in the presence of risk factors (T2D, hypertension or
435 metabolic syndrome). This merits urgent testing in a longitudinal independent clinical trial.

436

437 **Author contribution**

438 P.A., K.C. and M.E.-D. developed the present project concept and protocol. K.C. (coordinator), O.P.,
439 M.S., S.D.E., P.B., J.R., M.-E.D., F.B. and J.N. conceived the overall objectives and study design of the
440 MetaCardis initiative. MetaCardis cohort recruitment, phenotyping and lifestyle: J.A.-W., R.C., T.N.,
441 J.-E.S., F.A, L.K., H.V., T.H., J.-M.O. and M.B. and supervised by K.C., M.S. and O.P. Data curation: R.C.,
442 S.K.F., J.A.-W. and T.N. Bacterial cell count measurement: S.V.S. and G.F. Serum and urine
443 metabolome profiling: A.M., J.C., M.T.O. and L.H. Biochemical analyses: J.-P.G. Bioinformatics and
444 statistical analyses: P.A., S.K.F., G.F, S.V.S., R.A., E.L.C., L.P.C., E.P. and E.B. Dietary FQQ analyses: S.A.,
445 and B.H. Modelling of phenotypic data: P.A. Cellular in vitro experiments: P.A. Murine in vivo
446 experiments P.A. and J.K. with supervision from M.M.Y. The manuscript was written by P.A., KC and
447 M-E.D. with input from J.-A.W., R.C., S.K.F., M.M.Y., M.S., O.P. and S.D.E. All authors participated in
448 project development, discussion of results and revision of the article and approved the final version
449 for submission.

450 **Sources of funding**

451 This work was supported by European Union's Seventh Framework Program for research,
452 technological development and demonstration under grant agreement HEALTH-F4-2012-305312
453 (METACARDIS). Assistance Publique-Hôpitaux de Paris (AP-HP) is the promoter of the clinical
454 investigation (MetaCardis). Partial funding supports to K.C. and J.A.-W were also obtained from
455 Leducq Foundation (TransAtlantic grant), SFN (Société Française de Nutrition), F-CRIN-FORCE
456 network for support, INSERM via ITMO and JPI-Microdiet study and Novo Nordisk foundation
457 (Jacobaus prize). S.K.F. received support from Deutsche Forschungsgesellschaft SFB1365
458 ("RENOPROTECTION") and SFB1470: "HFpEF". P.A. is an ISSF Fellow supported by the Wellcome
459 Trust and Imperial College London (Grant No. PSN104). K.Ch. is an ISSF Fellow supported by the
460 Wellcome Trust and Imperial College London (Grant No. PSM279). M.-E.D. is supported by the NIHR
461 Imperial Biomedical Research Centre, GutsUK, Diabetes UK and by grants from the French National

462 Research Agency (ANR-10-LABX-46, European Genomics Institute for Diabetes), from the National
463 Center for Diabetes Precision Medicine – PreciDIAB, which is jointly supported by the French
464 National Agency for Research (ANR-18-IBHU-0001), by the European Union (FEDER), by the Hauts-
465 de-France Regional Council (Agreement 20001891/NP0025517) and by the European Metropolis of
466 Lille (MEL, Agreement 2019_ESR_11) and by Isite ULNE (R-002-20-TALENT-DUMAS), also jointly
467 funded by ANR (ANR16-IDEX-0004-ULNE), the Hauts-de-France Regional Council (20002845) and by
468 the European Metropolis of Lille (MEL). This research was conducted within the context of the
469 CNRS–Imperial International Research Project METABO-LIC. The Novo Nordisk Foundation Center for
470 Basic Metabolic Research is an independent research institution at the University of Copenhagen
471 partially funded by an unrestricted donation from the Novo Nordisk Foundation.

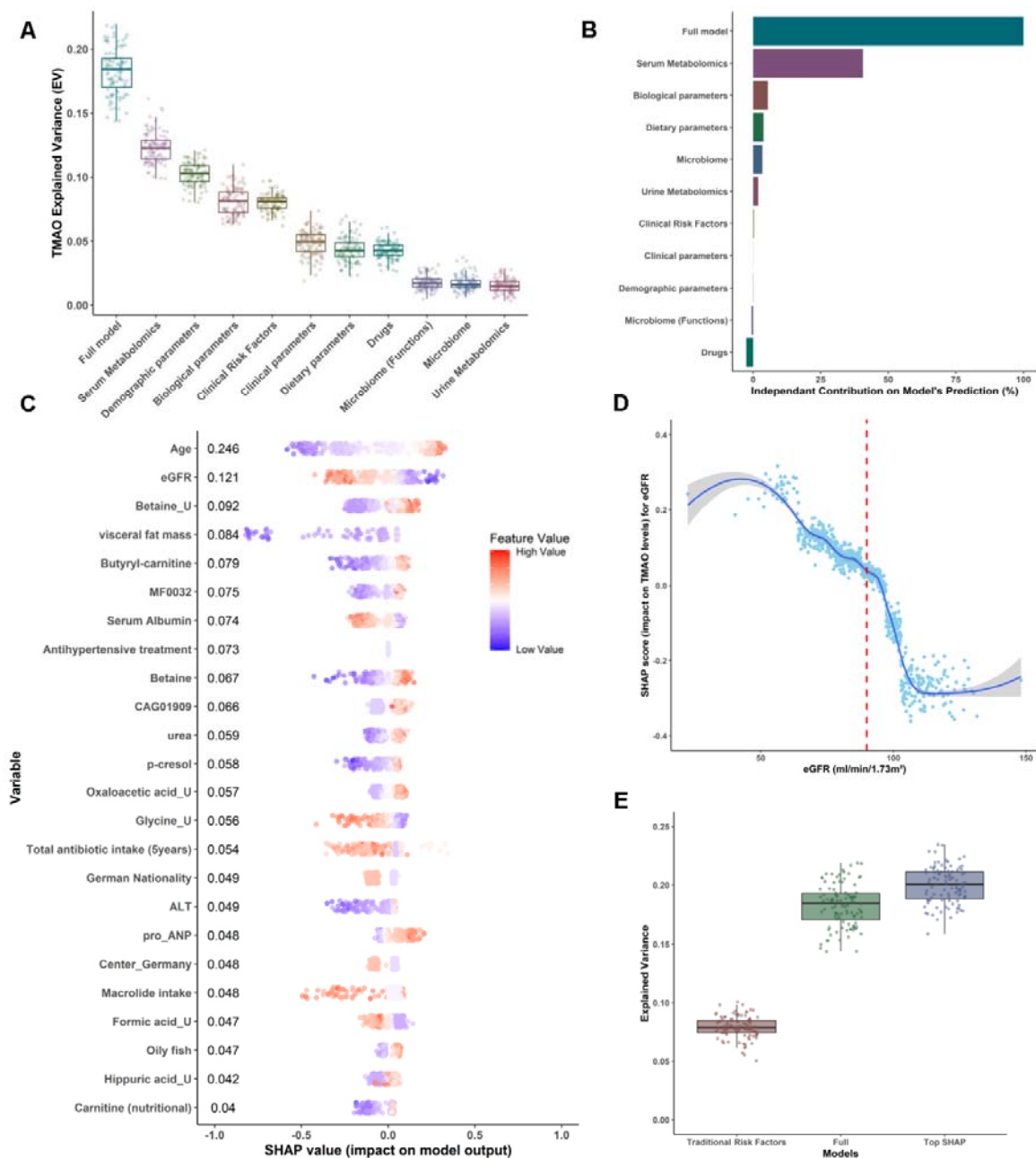
472 **Acknowledgements**

473 We thank the subjects for their participation in the MetaCardis study and particularly patient
474 associations (Alliance du Coeur and CNAO) for their input and interface, as well as Dr Dominique
475 Bonnefont-Rousselot (Department of Metabolic Biochemistry, Pitié-Salpêtrière hospital) for the
476 analysis of plasma lipid profiles. We thank the nurses, technicians, clinical research assistants and
477 data managers from the Clinical investigation platform at the Institute of Cardiometabolism and
478 Nutrition for patient investigations, at the CRNH (Centre de recherche en Nutrition Humaine CRNH-
479 Ile de France) and, the Clinical Investigation Center (CIC) from Pitié-Salpêtrière Hospital for
480 investigation of healthy controls. Quanta Medical provided regulatory oversight of the clinical study
481 and contributed to the processing and management of electronic data. Similarly, we are indebted to
482 the MetaCardis consortium (<http://www.metacardis.net/>) collaborators for contributions at multiple
483 levels since the consortium start in 2012. A full list of collaborators is given in the Supplementary
484 Section.

485 **Competing Interests.** KC is a consultant for Danone Research, Ysopia and CONFO therapeutics for
486 work not associated with this study. KC held a collaborative research contract with Danone Research
487 in the context of MetaCardis project. FB is a shareholder of Implexion pharma AB. MB received

488 lecture and/or consultancy fees from AstraZeneca, Boehringer-Ingelheim, Lilly, Novo Nordisk,
489 Novartis and Sanofi.

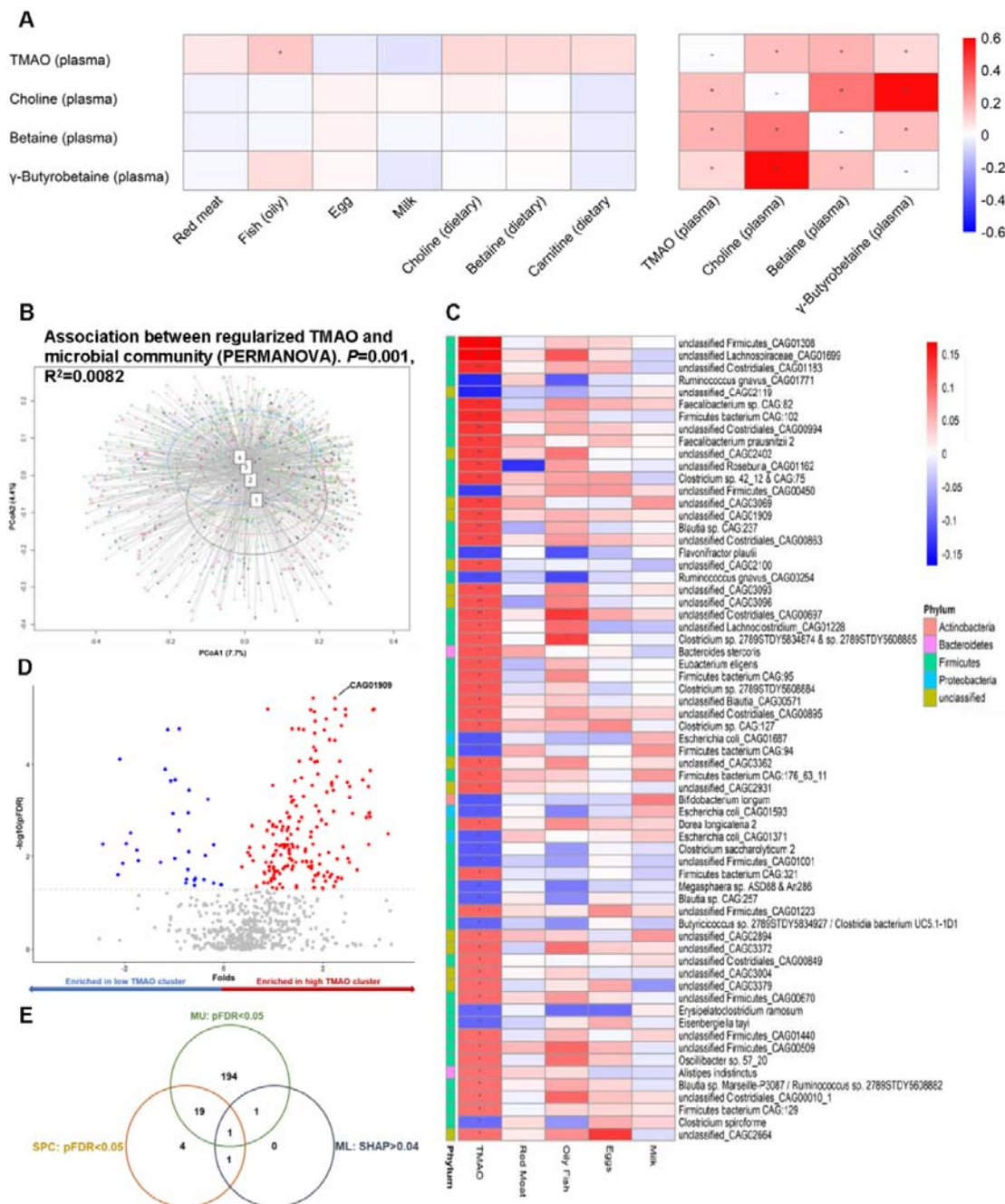
490 Main Figures



491

492 **Figure 1. Age and parameters associated with kidney function are the main drivers of circulating**
 493 **TMAO in BMIS MetaCardis participants. (A)** Coefficients of determination (Explained Variance; EV)
 494 of predicted circulating TMAO levels determined by xgboost algorithms after 5-fold cross-validation
 495 in the left-out group (**Suppl.Table.6** for N numbers and optimized xgboost parameters per group),
 496 trained exclusively on variables from each feature category (**Suppl.Table.5** for a list of variables

497 included in each group), or the full model (containing all variables), after 100 iterations
498 (**Suppl.Table.7**). (B) Averaged independent predictive contribution of each feature category to full
499 model predictions of plasma TMAO, trained as in (A), calculated as the average reduction of the
500 coefficient of determination achieved in relation to the full model (set to 100%) after removing all
501 variables belonging to each feature group after 100 iterations (**Suppl.Table.8**). (C) Swarm plots of
502 impact on model output (SHAP values; **Suppl.Table.10**) for each BMIS individual with complete
503 phenotypic data (N=582) for all variables contributing more than 4% to model predictions of
504 regularized TMAO standard deviation, as determined by xgboost algorithms trained on each feature
505 category. Mean absolute SHAP values from all BMIS participants (N=582) are shown (in descending
506 order) next to each variable. Individual dots, representing each participant, are colored by the
507 inverse-normalized value of the corresponding variable. U denotes urinary metabolites. (D)
508 Dependence plot of eGFR values (*x-axis*) versus their impact on model outcome (*y-axis*, SHAP values)
509 calculated for each individual in the BMIS cohort (N=837) from algorithms trained exclusively on
510 bioclinical variables (**Suppl.Table.5**), vertical red line indicates 90mL/min/1.73m². The curve was
511 drawn using locally weighted scatterplot smoothing (LOWESS). (E) Boxplots depicting Explained
512 Variance (EV; R^2) of circulating TMAO for BMIS participants computed from algorithms trained on
513 clinical risk factors²⁹, the full model containing all variables or all 24 variables contributing more than
514 4% of regularized TMAO standard deviation to model predictions, as determined by SHAP analysis,
515 after 100 iterations.



516

517 **Figure 2. Modest impact of diet and microbiota composition on circulating TMAO in BMIS**

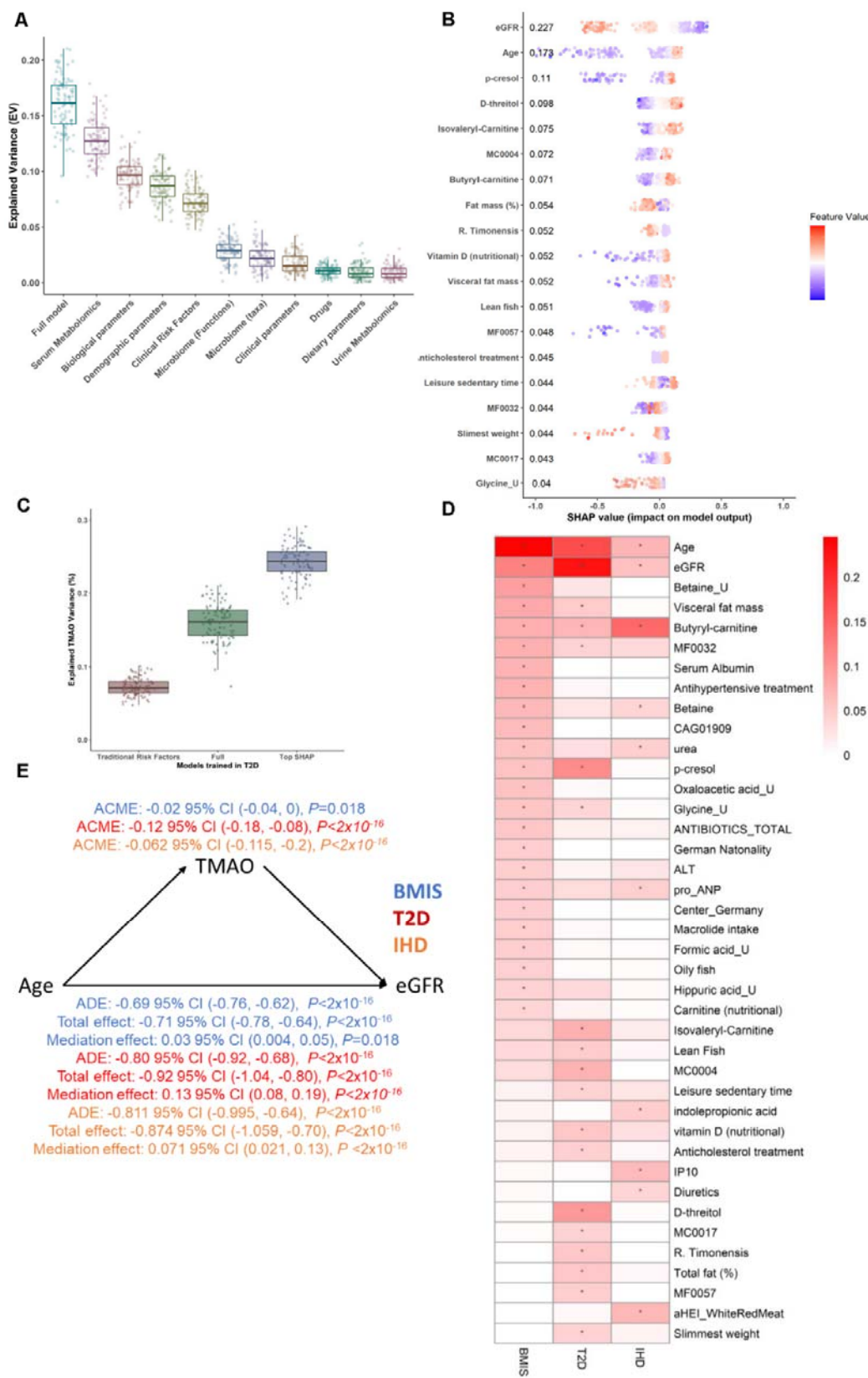
518 **MetaCardis subjects. (A)** Associations between circulating TMAO and its precursors with habitual

519 consumption of food items rich in TMAO precursors (N=763; left panel) or with its precursors

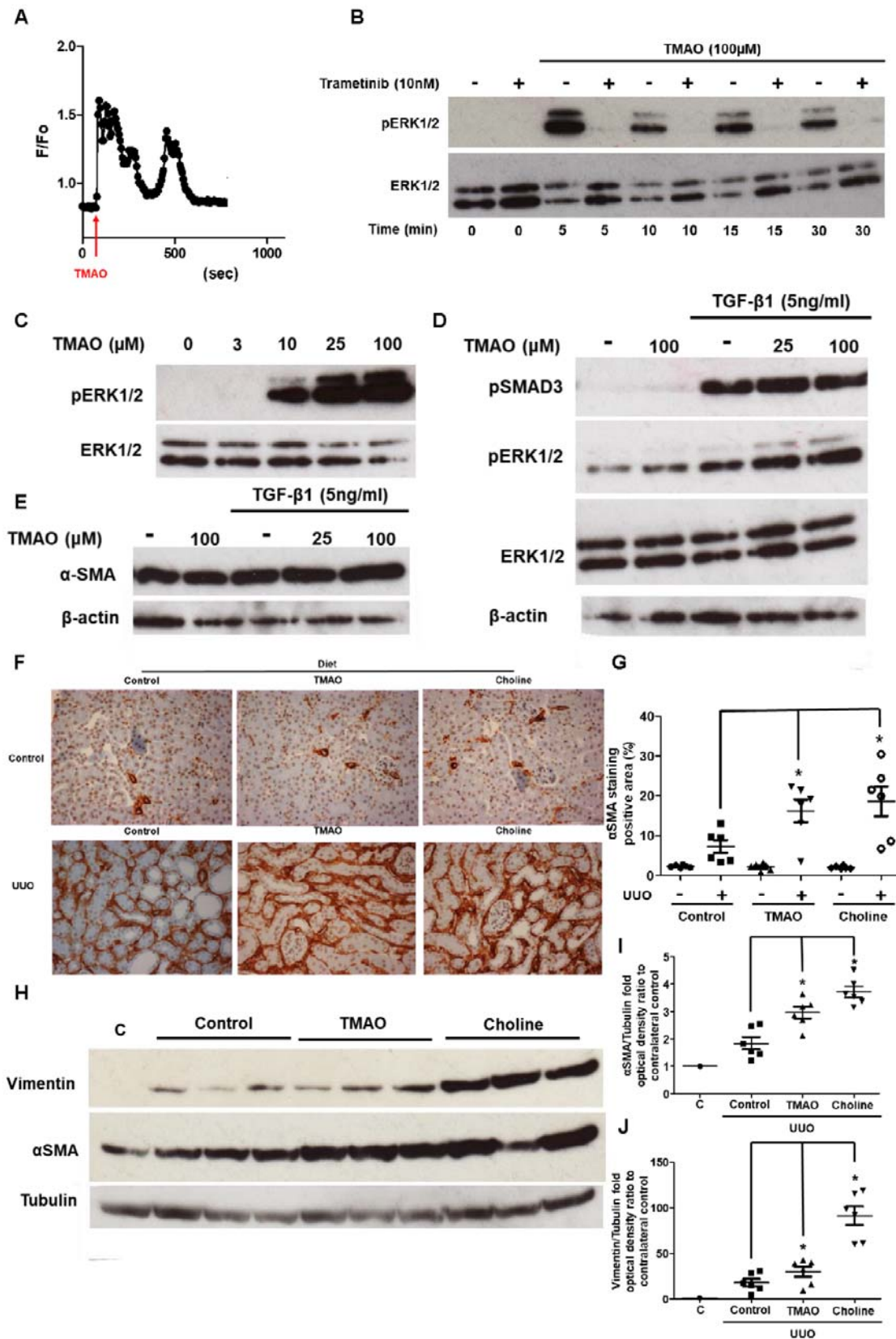
520 themselves (right panel) (**Suppl.Table.11, Suppl.Table.12**).

521 **(B)** Principal Coordinates analysis of Bray-Curtis dissimilarity matrices of participants (N=834) stratified in TMAO clusters by the k-means

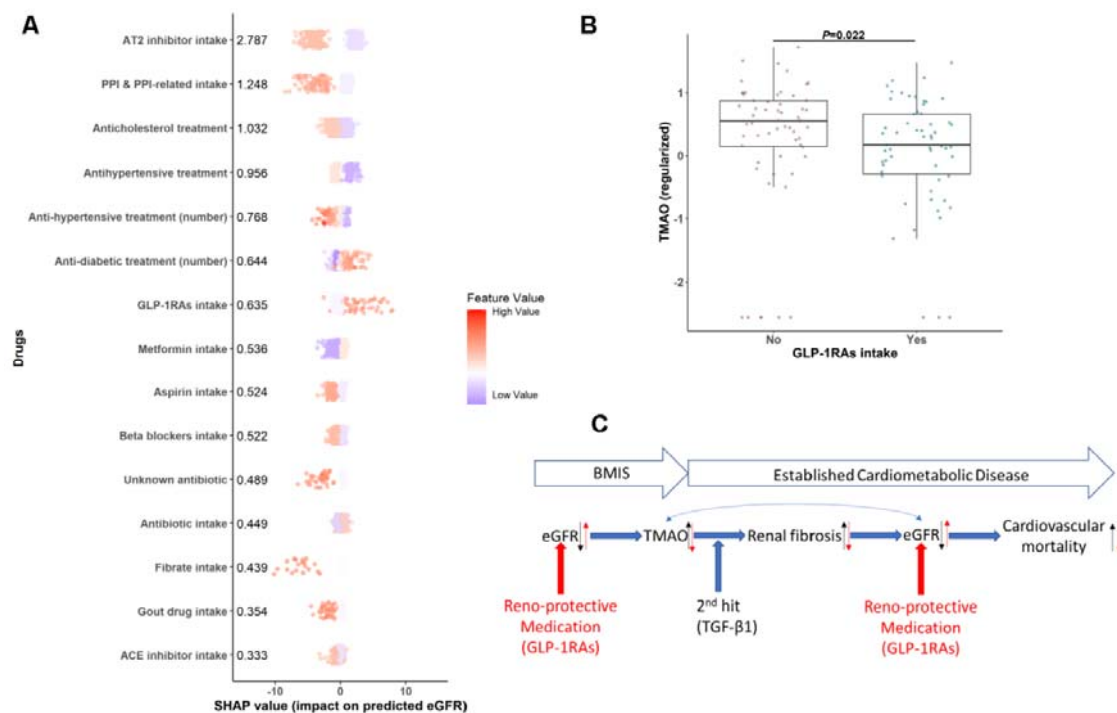
522 algorithm (1 the lowest, 4 the highest) at the species level (input; 699 species present in at least 20%
523 of the BMIS population). Insert; PERMANOVA (999 iterations) of taxonomic Bray-Curtis dissimilarity
524 matrices association with regularized TMAO levels with age, sex, and country of recruitment as
525 covariates. (C) Overlap of microbiome taxa significantly associated with circulating TMAO (Spearman
526 partial correlations adjusted for age, sex country of recruitment and BMI) and the consumption of
527 food items rich in TMAO precursors in BMIS participants (N=763; **Suppl.Table.13**). (D) Volcano plot
528 of differential bacterial species abundances between BMIS participants in the lowest (N=101) and
529 highest (N=147) TMAO clusters (blue; taxa significantly depleted, red; taxa significantly enriched in
530 the high TMAO cluster respectively, $pFDR < 0.05$; **Suppl.Table.14**). (E) Venn diagram summarizing the
531 overlap between taxa associating with circulating TMAO according to our three complimentary
532 analyses (SPC: Spearman correlations; ML: Machine learning and feature attribution analysis; MU:
533 Mann-Whitney U test between high and low TMAO clusters). For all * $pFDR < 0.05$, ** $pFDR < 0.01$.



535 **Figure 3. Signatures predicting circulating TMAO shift in different disease groups and TMAO**
536 **causally mediates eGFR decline with age.** (A) Explained Variance (EV) of predicted serum TMAO
537 levels (**Suppl.Table.17**) determined by boosted decision trees (**Suppl.Table.15** for N numbers and
538 optimized xgboost parameters per variable group), trained exclusively on variables from each
539 variable category (**Suppl.Table.5**, for a list of variables included in each group), or the full model
540 (containing all variables), after 100 iterations in T2D MetaCardis patients. (B) Swarm plots of SHAP
541 values (impact on model outcome; **Suppl.Table.19**) for each T2D MetaCardis participant with
542 complete phenotypic data (N=387); represented by individual dots, for all variables contributing
543 more than 4% to model predictions of regularized TMAO standard deviation, computed from
544 xgboost algorithms trained on each feature category. Numbers denote mean absolute SHAP values
545 from all T2D participants (in descending order) next to their corresponding variable. Dots are colored
546 by the inverse-normalized value of their corresponding variable. (C) Boxplots depicting Explained
547 Variance (EV; R^2) of circulating TMAO in T2D individuals calculated by algorithms trained on clinical
548 risk factors³¹, the full model containing all variables or all the variables contributing more than 4% of
549 regularized TMAO standard deviation to T2D model predictions, as determined by SHAP analysis,
550 after 100 iterations. (D) Heatmap depicting all the variables contributing at least 4% of regularized
551 TMAO standard deviation in model predictions as determined by SHAP analysis in at least one of the
552 MetaCardis disease groups. *Mean absolute SHAP value>0.04. (E) Mediation analysis (see Methods)
553 computing the direct effect of TMAO on eGFR decline with age in BMIS (blue), T2D (red) or IHD
554 (orange) MetaCardis participants. ADE: Average direct effect (of Age on eGFR); ACME: average
555 causal mediated effect (of TMAO on eGFR); Total effect: (cumulative effect of age and TMAO on
556 eGFR (ADE + ACME)); Mediation effect: (% of the effect of age on eGFR attributed to TMAO).



558 **Figure 4. TMAO promotes myofibroblast differentiation and exacerbates renal fibrotic injury. (A)**
559 Representative ratiometric traces (340/380nm) from Human Renal Fibroblasts (HRFs) loaded with
560 the Ca²⁺ indicator Fura-2 and stimulated with 100μM TMAO. **(B)** Serum-starved HRFs were pre-
561 incubated with the MEK inhibitor trametinib (10nM) for 30min prior to stimulation with 100μM
562 TMAO for the indicated times. Phospho-ERK1/2 levels were probed by Western blot; membranes
563 were stripped and re-probed for total-ERK1/2. **(C)** Serum-starved HRFs were stimulated with the
564 indicated concentrations of TMAO and phospho-ERK1/2 and total-ERK1/2 levels were determined as
565 in **(B)**. HRFs in complete medium were preincubated with the indicated concentrations of TMAO for
566 30min and stimulated with TGF-β1 (5nM) or vehicle for 24h. Phospho-ERK1/2, phospho-SMAD3 **(D)**
567 and alpha-smooth muscle actin (αSMA) **(E)** levels were probed with Western blot. **(F)**
568 Immunostaining with αSMA of kidney sections (20x magnification) from obstructed (UUO; 5days
569 post-surgery) or contralateral sham-operated (control) kidneys. Animals were fed normal chow
570 (control), a diet containing 0.12% w/w TMAO (TMAO) or 1% choline w/w (Choline) for 6weeks prior
571 to surgery, as indicated. N=6 per group. **(G)** Quantification of positive αSMA staining as (%) of
572 stained area/field of view averaged from 5 images per animal. **(H)** Western blot of whole-kidney
573 lysates for αSMA and vimentin expression. Membranes were subsequently stripped and re-probed
574 for tubulin, as loading control. A representative photomicrograph from n=2 Western blots with n=2–
575 6 animals in each group is shown. OD of the **(I)** αSMA and **(J)** vimentin bands in **(H)** normalized
576 against tubulin. The normalized density of the sham-control samples was arbitrarily set to 1. For all
577 graphs, error bars represent the mean±SEM of data from n=4–6 animals per group. *P<0.05 versus
578 the UUO control.



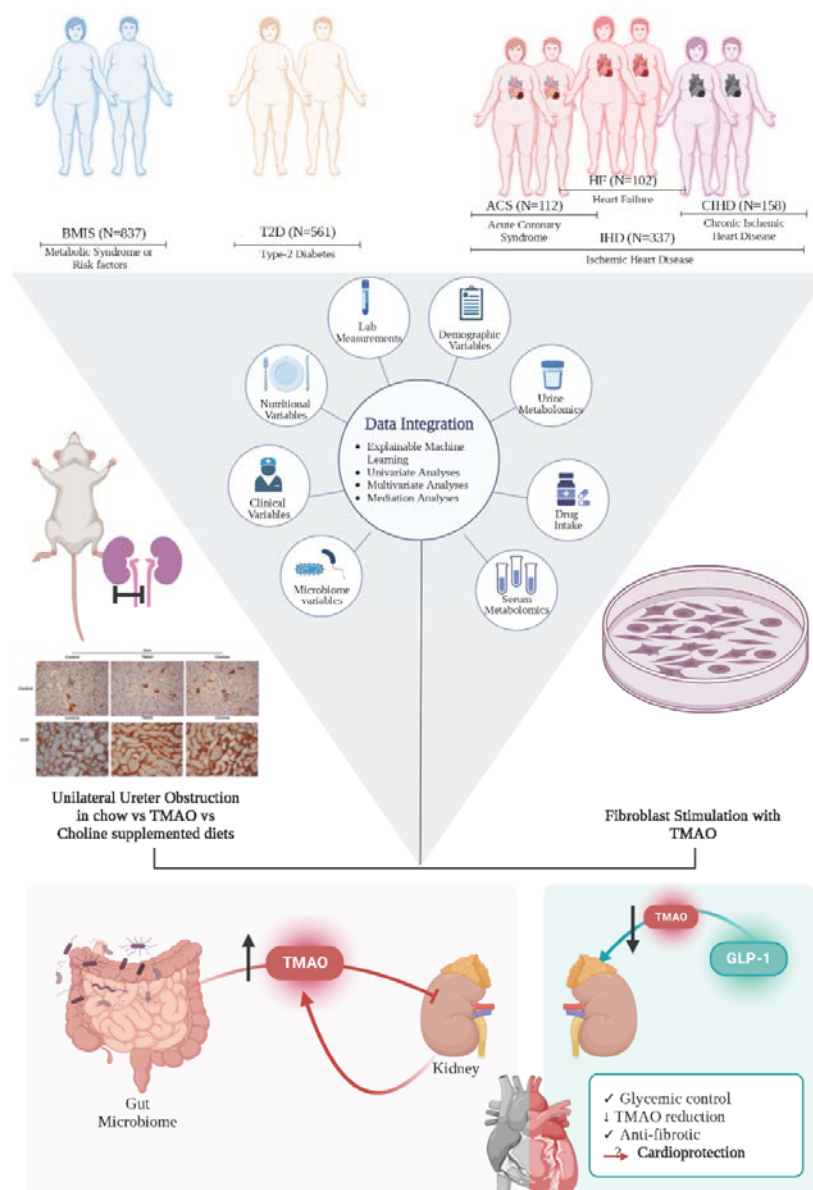
579

580 **Figure 5. Reno-protective medication is associated with reduced circulating TMAO in MetaCardis**
 581 **participants with T2D. (A)** Swarm plots of impact on model eGFR predictions (SHAP values;
 582 **Suppl.Table.22)** for MetaCardis T2D individuals (N=561) for the top 15 drugs, as determined by
 583 xgboost algorithms trained exclusively on prescribed medication. Mean absolute SHAP values from
 584 participants with T2D are shown (in descending order) next to each variable. Individual dots,
 585 representing each participant, are colored by the inverse-normalized value of the corresponding
 586 drug variable. **(B)** Comparison of circulating regularized TMAO levels between subjects with T2D
 587 prescribed GLP-1 receptor agonists (GLP-1Ras; N=59) and non-GLP-1Ras treated subjects with T2D
 588 propensity-matched for age, sex, BMI, disease group and hypertension status (N=59)
 589 **(Suppl.Table.23** for group characteristics). *P*-value determined by Mann-Whitney U test. **(C)**
 590 Summary of the main findings of our study. We demonstrate that eGFR, irrespective of disease
 591 stage, is the primary modifiable modulator of circulating TMAO. Far from being a bystander, TMAO
 592 significantly accelerates the rate of renal output decline by age, with its effect increasing at
 593 advanced stages of disease. TMAO promotes renal fibrosis in conjunction with established

594 pathophysiology (two-hit” model) further negatively impacting renal clearance. Accordingly,
595 medication with reno-protective properties (red arrows), such as GLP-1RAs, reduce circulating TMAO
596 levels thereby potentially moderating its adverse effect on kidney function.

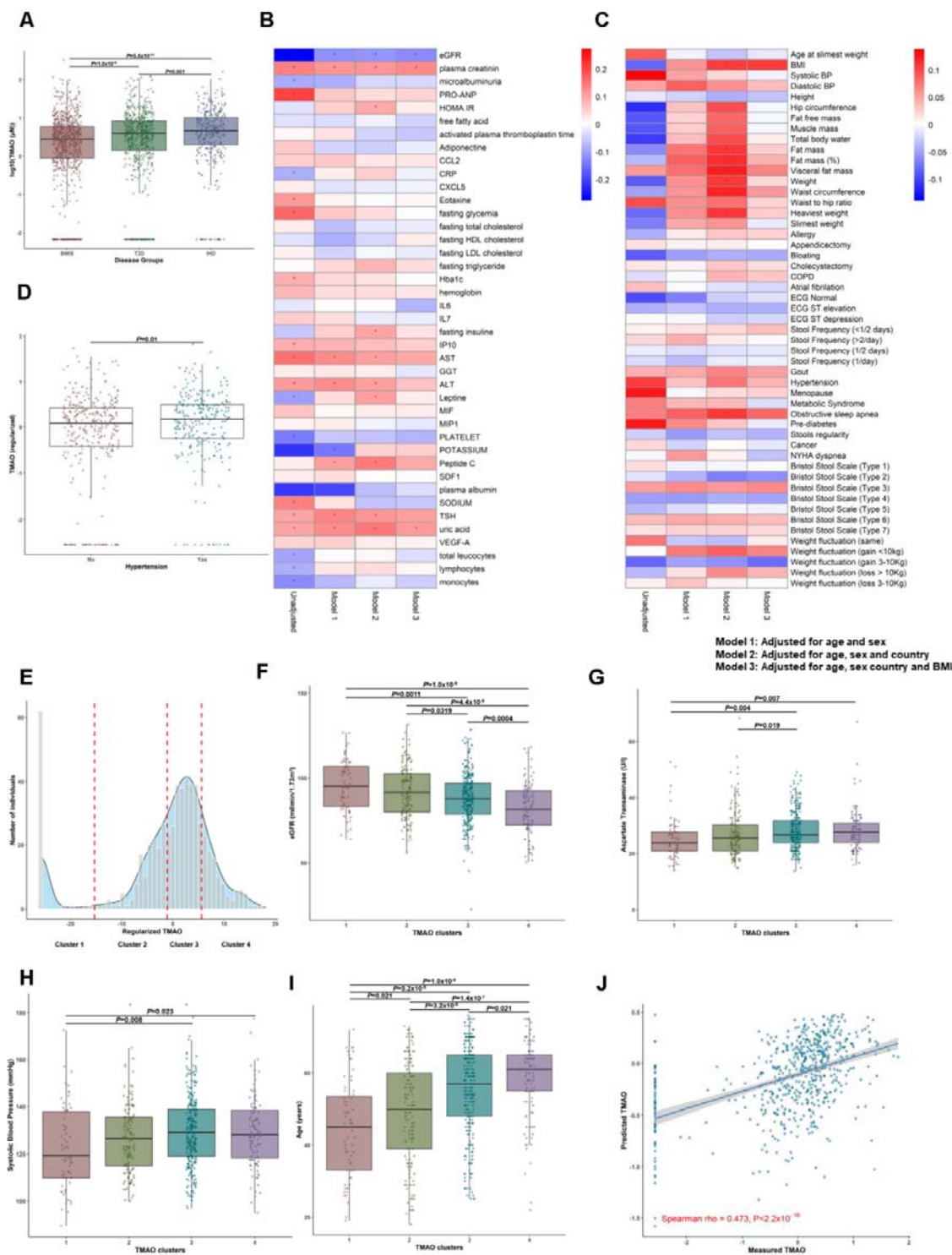
597

598 Supplemental Figures



599

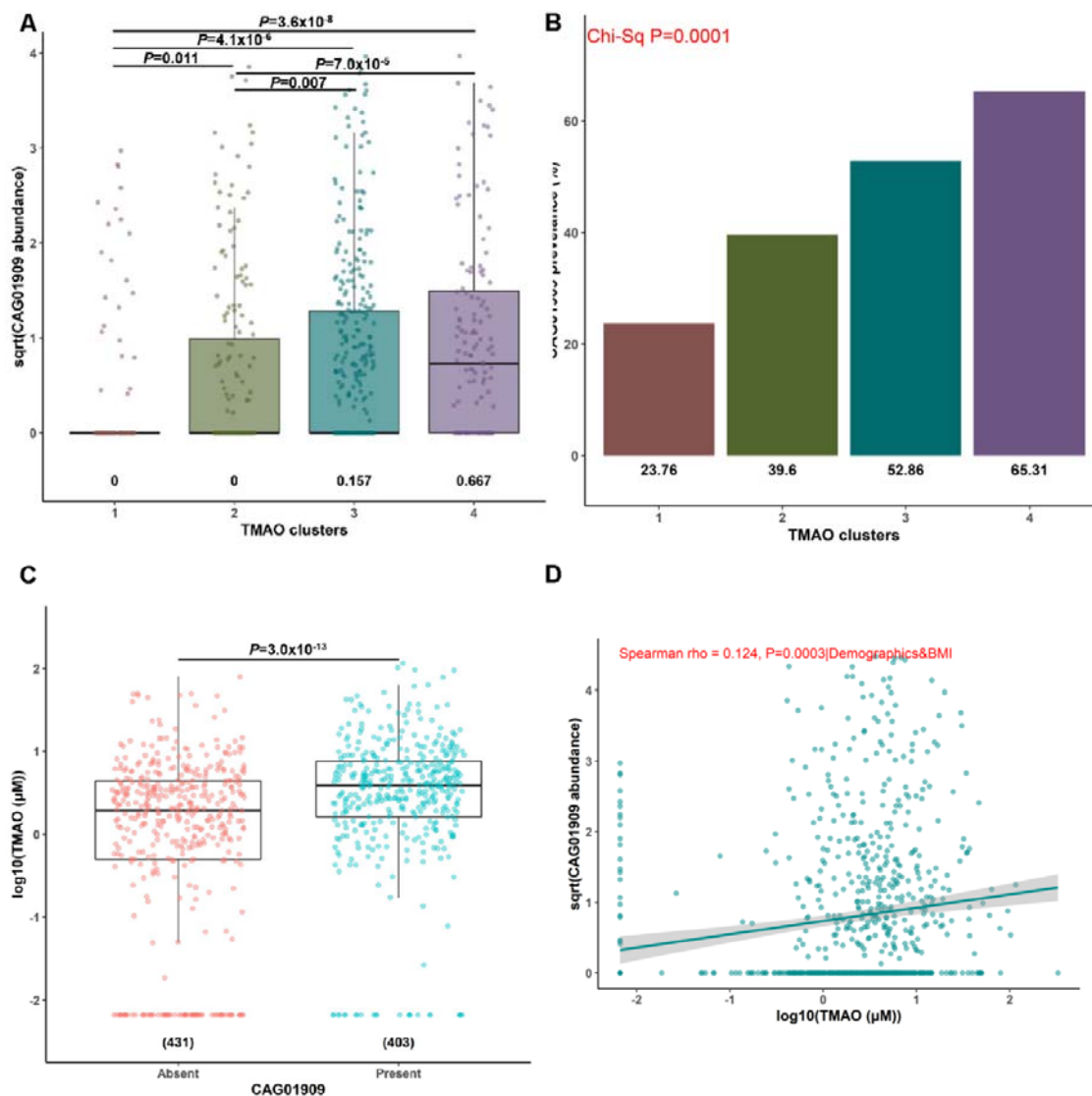
600 **Supplemental Figure 1. Overview of study design and aims.** Here we used an integrated approach
601 comprising Machine Learning (ML), multivariate, univariate and mediation analyses to objectively
602 characterize which host parameters contribute to plasma TMAO levels in the multi-center European
603 MetaCardis study. Capitalizing on the bioinformatics analysis we aimed to i) uncover novel host-
604 TMAO mechanistic targets and ii) identify actionable approaches (e.g. precision nutrition and/or
605 pharmacology interventions) to reduce systemic TMAO levels.



606

607 Supplemental Figure 2. TMAO is associated with worse cardiometabolic profiles in study
 608 participants with metabolic syndrome-related risk factors or pathologies but not over
 609 cardiometabolic disease (BMIS subgroup N=581). (A) Boxplot illustrating differences in fasting

610 circulating TMAO levels in the three MetaCardis patient groups (**Suppl.Table.1** for sub-cohort
611 characteristics). *P* determined by Mann-Whitney U test. **(B)** Spearman correlations between
612 circulating TMAO levels and bioclinical variables unadjusted, adjusted for age and sex (Model 1), age
613 sex and country of recruitment (Model 2) or age, sex country of recruitment and BMI (Model 3),
614 * $p_{FDR} < 0.1$. (**Suppl.Table.2**). **(C)** Spearman correlations between plasma TMAO and clinical variables,
615 corrected as in **(A)**, (**Suppl.Table.3**). **(D)** Comparison of circulating TMAO levels between BMIS
616 participants classed as hypertensives, *P* determined by Mann-Whitney U test. **(E)** Density plot
617 illustrating the distribution of regularized circulating TMAO levels in BMIS (N=582). The population
618 was split into four TMAO clusters using the k-means algorithm. Dashed red lines denote the cut-off
619 values for this distribution. Between group comparisons of selected variables for BMIS participants
620 split into clusters according to their circulating TMAO levels for eGFR (**F**; mL/min/1.73m²), aspartate
621 transaminase (**G**; U/l), systolic blood pressure (**H**; mmHg) and age at the time of recruitment (**I**,
622 years). *P* values were determined with pairwise Mann-Whitney U tests corrected for multiple
623 comparisons with the Benjamini-Hochberg method. **(J)** Predicted (averaged after 100 iterations; y-
624 axis) regularized plasma TMAO of BMIS participants by the BMIS-trained full-model *versus* actual
625 measured regularized TMAO values (*x-axis*) for BMIS individuals (N=582) (**Suppl.Table.8**); insert
626 Spearman rho and *P*-value of predicted *versus* measured values.



627

628 **Supplemental Figure 3. CAG01909, associates with higher circulating TMAO in BMIS (N=834)**

629 **individuals.** Comparisons of normalized to bacterial load expression levels (square-root transformed)

630 of (A) CAG01909, an unknown bacterium in BMIS participants divided into clusters according to

631 circulating TMAO levels with the k-means algorithm. Median abundance level for each cluster is

632 shown below the corresponding box plot, all pairwise comparisons performed with Mann-Whitney U

633 tests (corrected for multiple testing with the Benjamini-Hochberg method). (B) Prevalence ((%); in

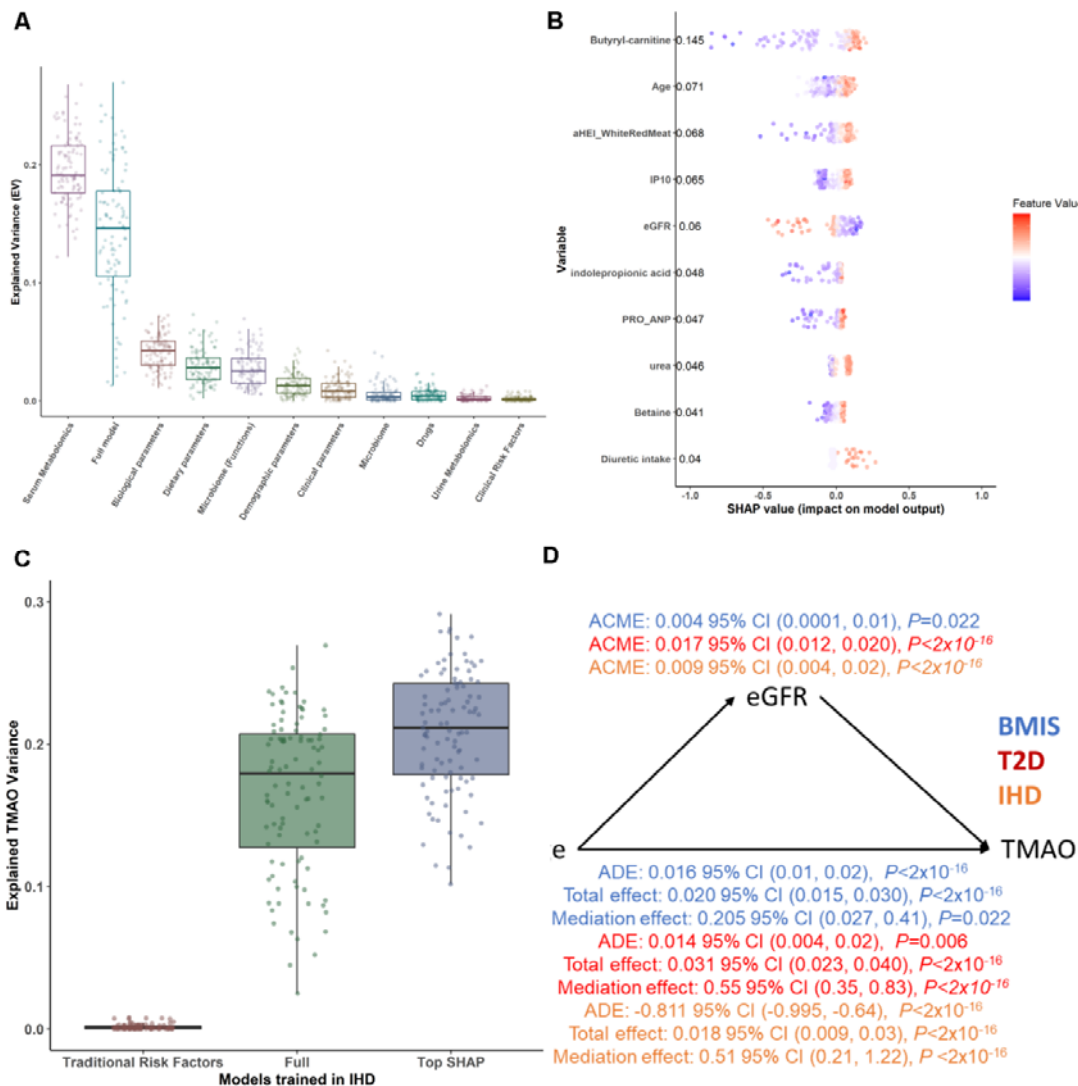
634 parentheses under corresponding boxplots) of CAG01909 in BMIS participants (N=834) split into

635 TMAO clusters as in (A). *P* value determined by Chi-square test. (C) BMIS participants were split into

636 those where CAG01909 is present in their gut microbiota (N=401) and those where it is absent
637 (N=433) and circulating log-transformed TMAO levels between the two groups were compared
638 (Mann-Whitney U test). (D) Linear-regression-based scatterplot showing correlation between
639 normalized by bacterial load CAG01909 abundance levels (square-root transformed for visualization
640 purposes) and TMAO (log₁₀-transformed). Adjusted (age, sex, country of recruitment and BMI)
641 Spearman rho=0.123 and pFDR=0.032 for CAG01909 (Suppl.Table.13).

642

643



644

645 **Supplemental Figure 4. Boosted decision tree algorithms predict circulating TMAO in IHD**

646 **MetaCardis participants. (A)** Explained Variance (EV) of predicted serum TMAO levels

647 **(Suppl.Table.18)** determined by boosted decision trees (methods), trained exclusively on variables

648 from each feature category **(Suppl.Table.5** for a list of variables included in each group), or the full

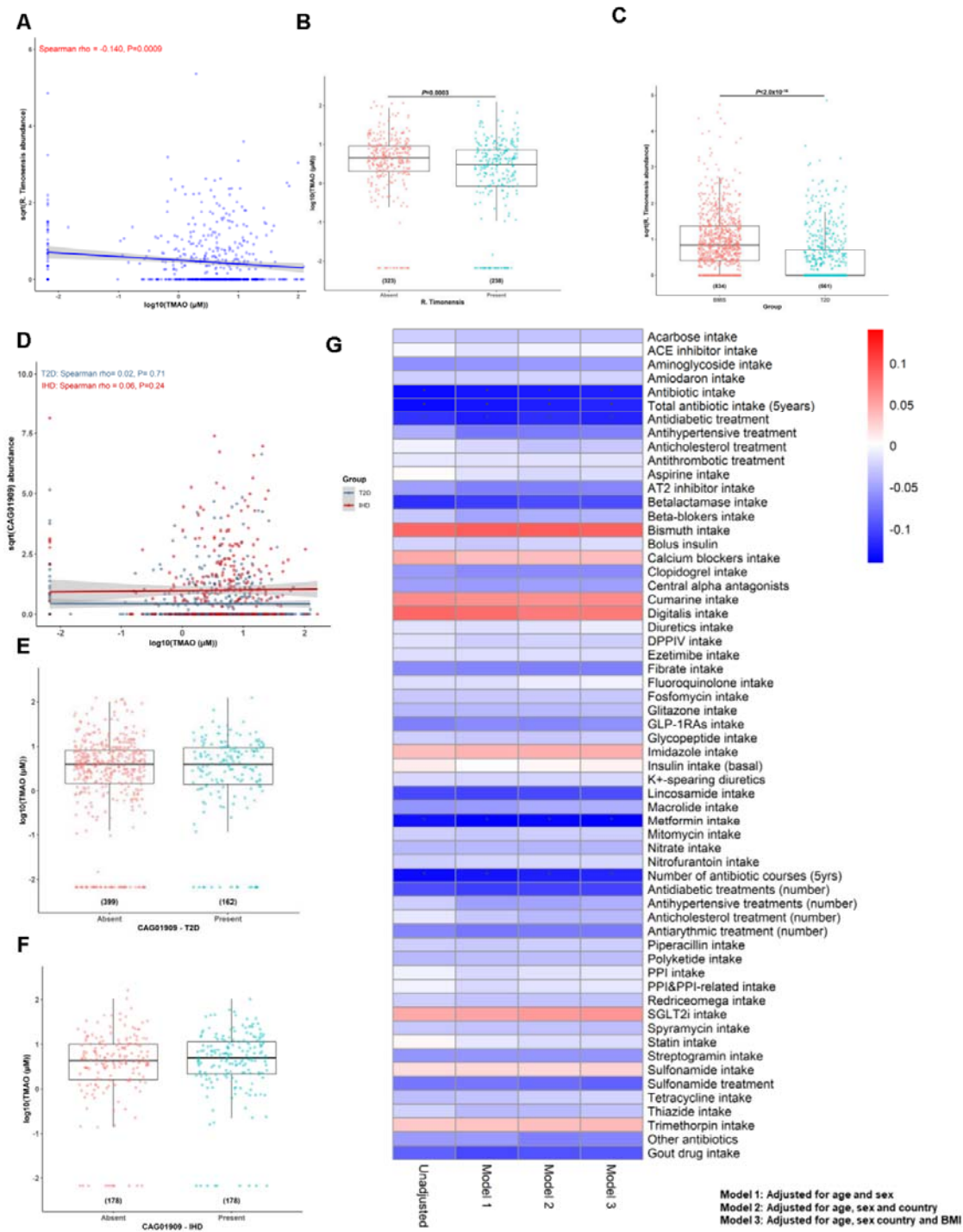
649 model (containing all variables), after 100 iterations **(Suppl.Table.16** for N numbers and optimized

650 xgboost parameters per variable group) in IHD MetaCardis individuals. **(B)** Swarm plots of SHAP

651 values (impact on model outcome; *x*-axis) for each IHD MetaCardis participant with complete

652 phenotypic data (N=221), for all variables contributing to model predictions more than 4% of

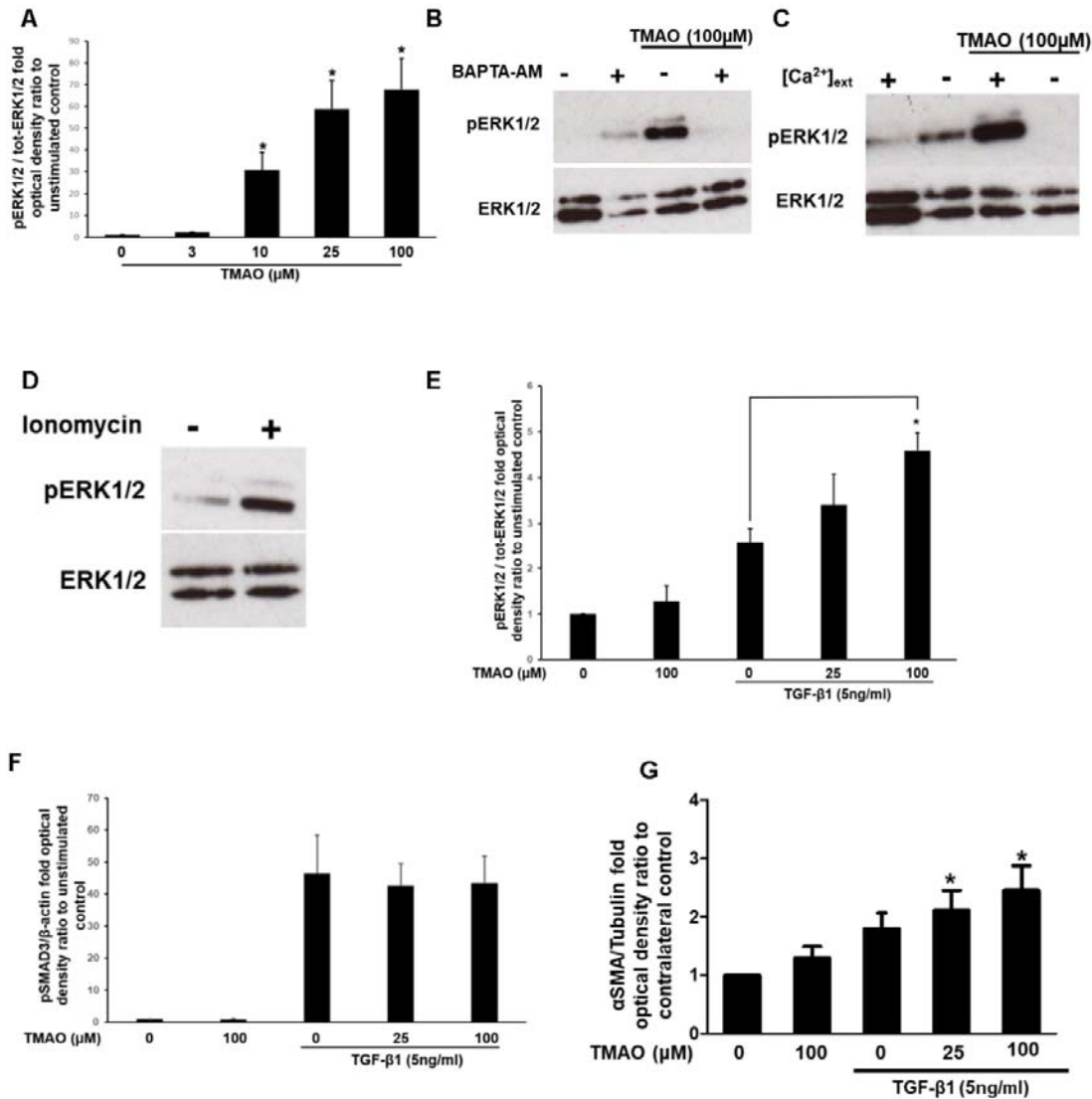
653 regularized TMAO standard deviation, computed from xgboost algorithms trained on each feature
654 category. Numbers denote mean absolute SHAP values from all IHD participants (in descending
655 order) next to their corresponding variable (**Suppl.Table.20**). Dots, representing IHD individuals, are
656 colored by the inverse-normalized value of their corresponding variable. (C) Boxplots depicting
657 Explained Variance (EV; R^2) of circulating TMAO in IHD individuals computed by algorithms trained
658 on clinical risk factors²⁹, the full model containing all variables or all the variables contributing more
659 than 4% of regularized TMAO standard deviation to IHD model predictions, as determined by SHAP
660 analysis, after 100 iterations. (D) Mediation analysis computing the direct effect of eGFR on TMAO
661 increase with age in BMIS (blue), T2D (red) or IHD (orange) MetaCardis participants. ADE: Average
662 direct effect (of Age on TMAO); ACME: average mediation effect (of eGFR on TMAO); Total effect:
663 (cumulative effect of age and eGFR on TMAO (ADE + ACME)); Mediation effect: (% of the effect of
664 age on circulating TMAO attributed to eGFR).



665

666 Supplemental Figure 5. In MetaCardis T2D participants *R. timonensis* inversely associates with
 667 circulating TMAO whilst BMIS-identified CAG01909 does not in overt disease. (A) Linear-
 668 regression-based scatterplot showing associations between normalized by bacterial load *R.*
 669 *timonensis* abundance levels (square-root transformed for visualization purposes) and TMAO (\log_{10} -

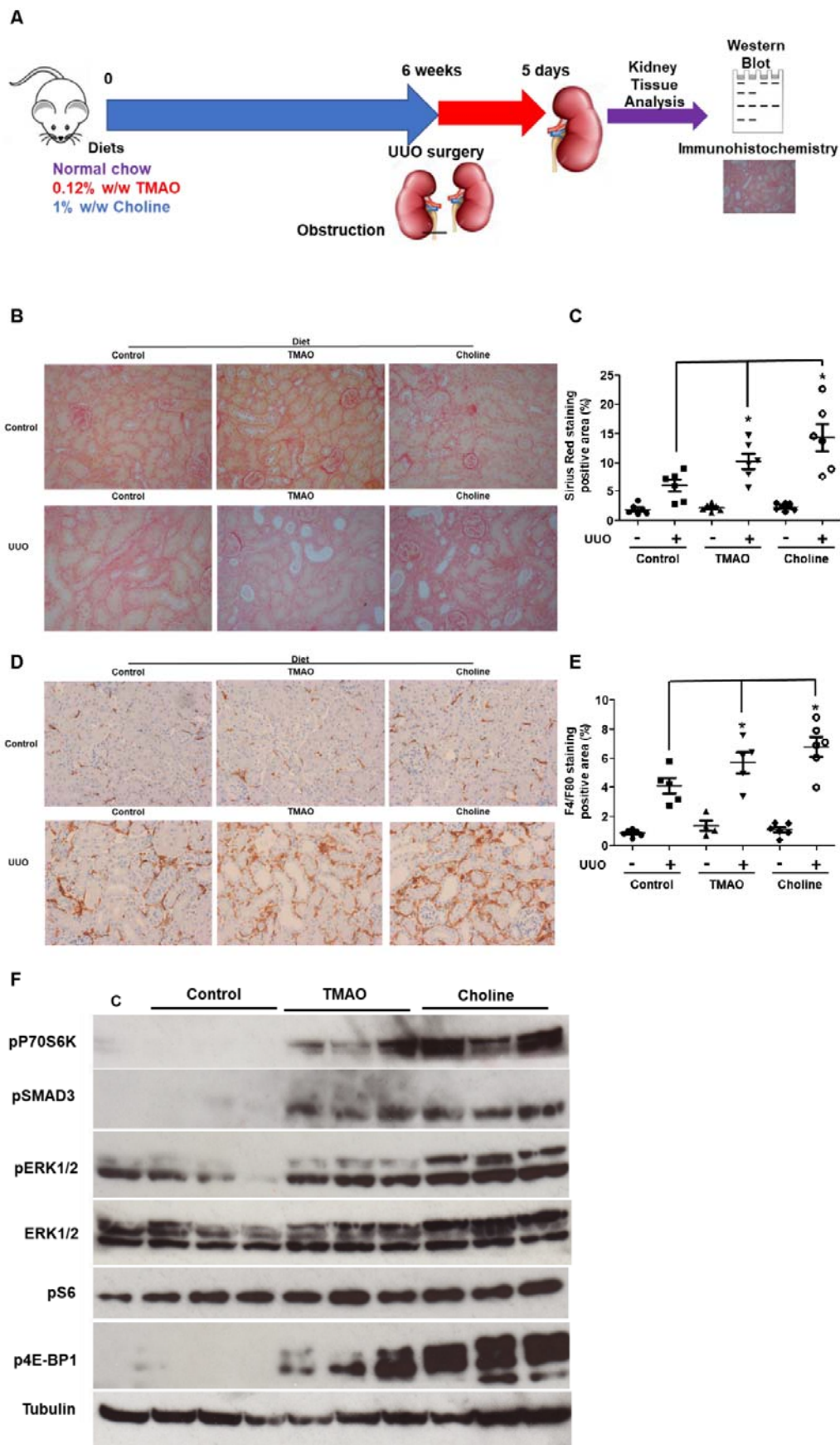
670 transformed). Unadjusted spearman rho=-0.140 and P=0.009. **(B)** Comparison of log-transformed
671 circulating TMAO between T2D individuals with detectable (present) and undetectable (absent) fecal
672 *R. timonensis* (Mann-Whitney test). Number of individuals in each group shown in parenthesis. **(C)**
673 Boxplot illustrating differences in *R. timonensis* fecal abundance in the BMIS and T2D MetaCardis
674 patient groups (group sizes in parentheses), *P* determined by Mann-Whitney U test. **(D)** Linear-
675 regression-based scatterplot showing correlation between normalized by bacterial load CAG01909
676 expression levels (square-root transformed for visualization purposes) and TMAO (\log_{10} -
677 transformed) in T2D and IHD MetaCardis subjects. Comparison of log-transformed circulating TMAO
678 between T2D **(E)** and IHD **(F)** MetaCardis subjects with detectable (present) or undetectable (absent)
679 fecal CAG01909 (Mann-Whitney test). Number of individuals in each group shown in parenthesis. **(G)**
680 Heatmap illustrating Spearman correlations between bacterial load-normalized abundance of
681 CAG01909 and intake of medication in T2D MetaCardis individuals (N=561) unadjusted, adjusted for
682 age and sex (Model 1), age sex and country of recruitment (Model 2) or age, sex country of
683 recruitment and BMI (Model 3), *pFDR<0.1. **(Suppl.Table.21)**.



684

685 **Supplemental Figure 6. TMAO activates ERK1/2 via a Ca²⁺-sensitive pathway in human renal**
 686 **fibroblasts. (A)** Optical density (OD) of pERK1/2 levels normalized against total ERK1/2 for
 687 experiments represented in **Figure 4C**. The normalized density of the control unstimulated samples
 688 was arbitrarily set to 1. **(B)** Serum-starved Human Renal Fibroblasts (HRFs) were loaded with the Ca²⁺
 689 chelator BAPTA-AM (20 μM ; 20min) and then stimulated with TMAO (100 μM) for 10min. ERK1/2
 690 activation was probed by Western blot and membranes were stripped and re-probed for total
 691 ERK1/2, a representative image of N=3 independent experiments is shown. **(C)** HRFs were serum-
 692 starved in complete physiological medium for 45min, the medium was aspirated and cells were

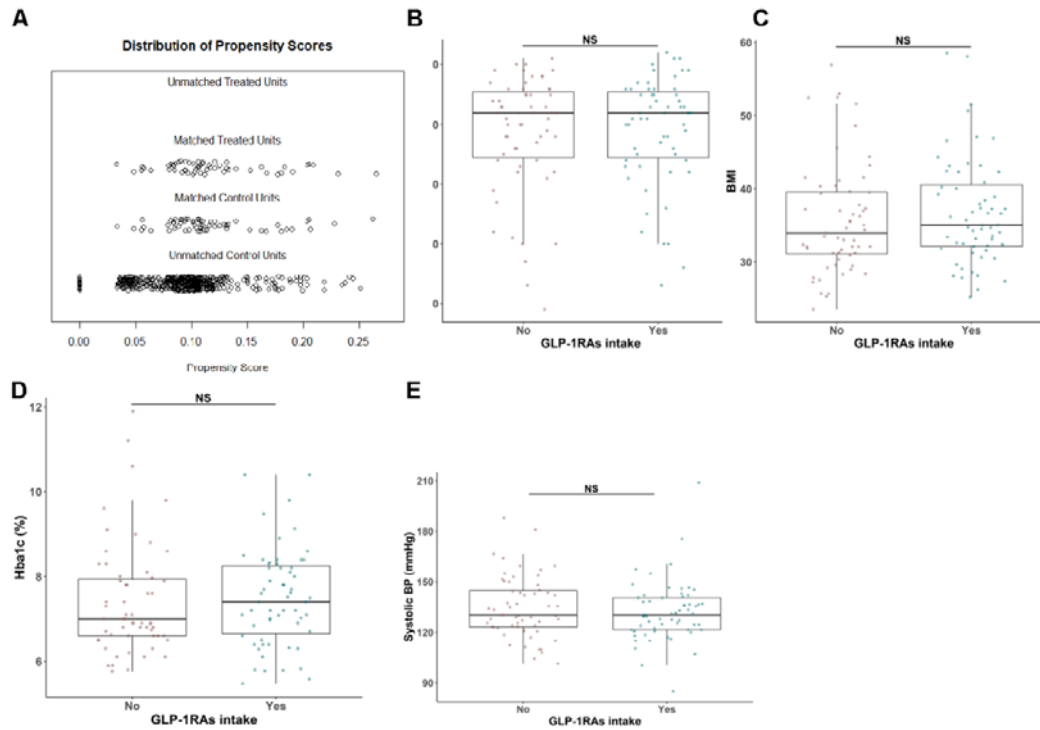
693 incubated for further 15min in physiological medium with Ca^{2+} omitted from the buffer.
694 Subsequently, cells were stimulated with 100 μM TMAO for 10min and ERK1/2 activation was probed
695 as in (B). (D) Serum-starved HRFs were stimulated with the Ca^{2+} ionophore ionomycin (100 μM) and
696 phospho-ERK1/2 levels were probed as in (B). OD of pERK1/2 (E) or pSMAD3 (F) normalized by total
697 ERK1/2 or β -actin respectively for experiments represented in **Figure 4D**. The normalized density of
698 the control unstimulated samples was arbitrarily set to 1. (G) OD of αSMA normalized by β -actin
699 levels for experiments represented in **Figure 4E**. For all N=3, error bars represent $\pm\text{SEM}$.



701 **Supplemental Figure 7. TMAO or choline diet exacerbates collagen deposition, macrophage**
702 **infiltration and activate pro-fibrotic signaling in kidneys of mice that underwent Unilateral Ureter**
703 **Obstruction (UUO) surgery. (A)** Outline of the animal experiments. **(B)** Sirius red staining of kidney
704 sections (20x magnification) from obstructed (UUO; 5days post-surgery) or contralateral sham-
705 operated (control) kidneys. Animals were fed normal chow (control), a diet containing 0.12% w/w
706 TMAO (TMAO) or 1% choline w/w (Choline) for 6weeks prior to surgery, as indicated. N=6 per group.
707 **(C)** Quantification of collagen deposition staining as (%) of positive Sirius red area/field of view
708 averaged from 5 images per animal. **(D)** Immunostaining of kidney sections (20x magnification) from
709 UUO or control kidneys as in **(B)** with the macrophage marker F4/F80. **(E)** Quantification of
710 macrophage infiltration as % of F4/F80 staining/ field of view of images from **(D)**; N=6
711 animals/group. **(F)** Western blot for profibrotic signaling pathways activation from UUO kidneys from
712 animals treated as in **(A)**. C: contralateral kidney from a control-diet fed animal; TMAO: 0.12% w/w
713 TMAO diet; Choline: 1% w/w Choline diet.

714

715



716

717 **Supplemental Figure 8. Propensity-score matching of MetaCardis patients with T2D.** (A) Propensity
718 scores of MetaCardis participants with T2D taking GLP-1 Receptor agonists (GLP-1RAs); N=59) and
719 matched controls (N=59) using nearest neighbors matching with age, sex, disease severity group and
720 hypertension status as covariates (Suppl.Table.23). Comparisons of age (B), BMI (C), glycated
721 hemoglobin (D) and systolic blood pressure (E), between MetaCardis patients with T2D diabetics
722 prescribed GLP-1RAs (N=59) and matched controls (as in A; N=59). Mann-Whitney U tests were used
723 to determine statistical significance for B-E. NS: non-significant ($P>0.05$).

724

725 **Supplemental Tables**

726 **Supplemental Table 1.** Characteristics of MetaCardis sub-cohorts (BMIS, T2D, IHD).

727 **Supplemental Table 2.** Spearman correlations between circulating TMAO and biological variables in
728 BMIS.

729 **Supplemental Table 3.** Spearman correlations between circulating TMAO and clinical variables in
730 BMIS.

731 **Supplemental Table 4.** BMIS TMAO cluster characteristics.

732 **Supplemental Table 5.** List of all the variables per Group variable category.

733 **Supplemental Table 6.** Model hyperparameters and N numbers for each variable group in BMIS.

734 **Supplemental Table 7.** Explained Variance (EV) of models trained with each variable group in BMIS
735 for each iteration.

736 **Supplemental Table 8.** Predicted *versus* measured circulating TMAO in BMIS subjects.

737 **Supplemental Table 9.** Contributions to predictions of the full model by each variable group.

738 **Supplemental Table 10.** Average SHAP values for all variables for each BMIS subject.

739 **Supplemental Table 11.** Spearman correlations between circulating TMAO and habitual intake of 37
740 food categories in BMIS.

741 **Supplemental Table 12.** Spearman correlations between circulating TMAO and calculated habitual
742 intake of micronutrients in BMIS.

743 **Supplemental Table 13.** Spearman correlations between circulating TMAO and bacterial species
744 abundance in BMIS.

745 **Supplemental Table 14.** Differential taxon abundance between high *versus* low TMAO clusters in
746 BMIS.

747 **Supplemental Table 15.** Model hyperparameters and N numbers for each variable group in T2D.

748 **Supplemental Table 16.** Model hyperparameters and N numbers for each variable group in IHD.

749 **Supplemental Table 17.** Explained Variance (EV) of models trained with each variable group in T2D

750 for each iteration.

751 **Supplemental Table 18.** Explained Variance (EV) of models trained with each variable group in IHD

752 for each iteration.

753 **Supplemental Table 19.** Average impact on TMAO predictions (SHAP values) for all variables for each

754 T2D subject.

755 **Supplemental Table 20.** Average impact on TMAO predictions (SHAP values) for all variables for each

756 IHD subject.

757 **Supplemental Table 21.** Spearman correlations between CAG01909 abundance and intake of drugs

758 in T2D.

759 **Supplemental Table 22.** Average impact on predicted eGFR (SHAP values) for prescribed medication

760 for MetaCardis T2D subjects.

761 **Supplemental Table 23.** Characteristics of MetaCardis participants prescribed GLP-1Ras and

762 matched controls.

763 **Supplemental Table 24.** Materials.

764

765

766

767

768

769

770 **References**

- 771 1. Fan Y, Pedersen O. Gut microbiota in human metabolic health and disease. *Nat Rev*
772 *Microbiol* 2021;**19**:55-71. doi: 10.1038/s41579-020-0433-9
773 pmid:<https://pubmed.ncbi.nlm.nih.gov/32887946>
- 774 2. Agus A, Clément K, Sokol H. Gut microbiota-derived metabolites as central regulators in
775 metabolic disorders. *Gut* 2021;**70**:1174-1182. doi: 10.1136/gutjnl-2020-323071
776 pmid:<https://pubmed.ncbi.nlm.nih.gov/33272977>
- 777 3. Li J, Li Y, Ivey KL, et al. Interplay between diet and gut microbiome, and circulating
778 concentrations of trimethylamine N-oxide: findings from a longitudinal cohort of US men. *Gut*
779 2022;**71**:724-733..doi: 10.1136/gutjnl-2020-322473
780 pmid:<https://pubmed.ncbi.nlm.nih.gov/33926968>
- 781 4. Romano KA, Vivas EI, Amador-Noguez D, et al. Intestinal microbiota composition modulates
782 choline bioavailability from diet and accumulation of the proatherogenic metabolite trimethylamine-
783 N-oxide. *mBio* 2015;**6**:e02481.doi: 10.1128/mBio.02481-14
784 pmid:<https://pubmed.ncbi.nlm.nih.gov/25784704/>
- 785 5. Rath S, Heidrich B, Pieper DH, et al. Uncovering the trimethylamine-producing bacteria of
786 the human gut microbiota. *Microbiome* 2017;**5**:54.doi: 10.1186/s40168-017-0271-9.
- 787 6. Wang Z, Klipfell E, Bennett BJ, et al. Gut flora metabolism of phosphatidylcholine promotes
788 cardiovascular disease. *Nature* 2011;**472**:57-63.doi:10.1038/nature09922
789 pmid:<https://pubmed.ncbi.nlm.nih.gov/21475195>
- 790 7. Tang WH, Wang Z, Levison BS, et al. Intestinal microbial metabolism of phosphatidylcholine
791 and cardiovascular risk. *N Engl J Med* 2013;**368**:1575-84.doi:10.1056/NEJMoa1109400
792 pmid:<https://pubmed.ncbi.nlm.nih.gov/23614584>
- 793 8. Koeth RA, Wang Z, Levison BS, et al. Intestinal microbiota metabolism of L-carnitine, a
794 nutrient in red meat, promotes atherosclerosis. *Nat Med* 2013;**19**:576-85.doi: 10.1038/nm.3145
795 pmid:<https://pubmed.ncbi.nlm.nih.gov/23563705>

- 796 9. Wang Z, Bergeron N, Levison BS, et al. Impact of chronic dietary red meat, white meat, or
797 non-meat protein on trimethylamine N-oxide metabolism and renal excretion in healthy men and
798 women. *Eur Heart J* 2019;40:583-594.doi: 10.1093/eurheartj/ehy799
799 pmid:<https://pubmed.ncbi.nlm.nih.gov/30535398>
- 800 10. Wu WK, Chen CC, Liu PY, et al. Identification of TMAO-producer phenotype and host–diet–
801 gut dysbiosis by carnitine challenge test in human and germ-free mice. *Gut* 2019;68:1439–1449.doi:
802 10.1136/gutjnl-2018-317155 pmid:<https://pubmed.ncbi.nlm.nih.gov/30377191>
- 803 11. Li XS, Obeid S, Klingenberg R, et al. Gut microbiota-dependent trimethylamine N-oxide in
804 acute coronary syndromes: a prognostic marker for incident cardiovascular events beyond
805 traditional risk factors. *Eur Heart J* 2017;38:814-824.doi: 10.1093/eurheartj/ehw582
806 pmid:<https://pubmed.ncbi.nlm.nih.gov/28077467>
- 807 12. Key TJ, Appleby PN, Bradbury KE, et al. Consumption of Meat, Fish, Dairy Products, and Eggs
808 and Risk of Ischemic Heart Disease. *Circulation* 2019;139:2835-2845.
809 doi:10.1161/CIRCULATIONAHA.118.038813 pmid:<https://pubmed.ncbi.nlm.nih.gov/31006335/>
- 810 13. Gibson R, Lau CH, Loo RL, et al. The association of fish consumption and its urinary
811 metabolites with cardiovascular risk factors: the International Study of Macro-/Micronutrients and
812 Blood Pressure (INTERMAP). *Am J Clin Nutr* 2020;111:280-290.doi:10.1093/ajcn/nqz293
813 pmid:<https://pubmed.ncbi.nlm.nih.gov/31782492>
- 814 14. Hoyles L, Jiménez-Pranteda ML, Chilloux J, et al. Metabolic retroconversion of
815 trimethylamine N-oxide and the gut microbiota. *Microbiome* 2018;6:73.doi: 10.1186/s40168-018-
816 0461-0 pmid:<https://pubmed.ncbi.nlm.nih.gov/29678198>
- 817 15. Brunt VE, Gioscia-Ryan RA, Casso AB, et al. Trimethylamine-N-Oxide Promotes Age-Related
818 Vascular Oxidative Stress and Endothelial Dysfunction in Mice and Healthy Humans. *Hypertension*
819 2020;76:101-112.doi: 10.1161/HYPERTENSIONAHA.120.14759

- 820 16. Bennett BJ, Vallim TQ, Wang Z, *et al.* Trimethylamine-N-oxide, a metabolite associated with
821 atherosclerosis, exhibits complex genetic and dietary regulation. *Cell Metab* 2013;**17**:49-60.doi:
822 10.1016/j.cmet.2012.12.011 pmid:<https://pubmed.ncbi.nlm.nih.gov/23312283>
- 823 17. Stubbs JR, House JA, Ocque AJ, *et al.* Serum Trimethylamine-N-Oxide is Elevated in CKD and
824 Correlates with Coronary Atherosclerosis Burden. *J Am Soc Nephrol* 2016;**27**:305-13.
825 doi:[10.1681/ASN.2014111063](https://doi.org/10.1681/ASN.2014111063) pmid:<https://pubmed.ncbi.nlm.nih.gov/26229137>
- 826 18. Tang WHW, Wang Z, Kennedy DJ, *et al.* Gut microbiota-dependent trimethylamine N-oxide
827 (TMAO) pathway contributes to both development of renal insufficiency and mortality risk in chronic
828 kidney disease. *Circ Res* 2015;**116**:448-55.doi: 10.1161/CIRCRESAHA.116.305360
829 pmid:<https://pubmed.ncbi.nlm.nih.gov/25599331>
- 830 19. Manor O, Zubair N, Conomos MP, *et al.* A Multi-omic Association Study of Trimethylamine N-
831 Oxide. *Cell Rep* 2018;**24**:935-946.doi: 10.1016/j.celrep.2018.06.096
832 pmid:<https://pubmed.ncbi.nlm.nih.gov/30044989>
- 833 20. Zhuang R, Ge X, Han L, *et al.* Gut microbe-generated metabolite trimethylamine N-oxide and
834 the risk of diabetes: A systematic review and dose-response meta-analysis. *Obes Rev* 2019;**20**:883-
835 894. doi: [10.1111/obr.12843](https://doi.org/10.1111/obr.12843) pmid:<https://pubmed.ncbi.nlm.nih.gov/30868721>
- 836 21. Bar N, Korem T, Weissbrod O, *et al.* A reference map of potential determinants for the
837 human serum metabolome. *Nature* 2020;**588**:135-140.doi:[10.1038/s41586-020-2896-2](https://doi.org/10.1038/s41586-020-2896-2)
838 pmid:<https://pubmed.ncbi.nlm.nih.gov/33177712>
- 839 22. Wilding JPH, Batterham RL, Calanna S, *et al.* Once-Weekly Semaglutide in Adults with
840 Overweight or Obesity. *N Engl J Med* 2021;**384**:989-1002. doi:[10.1056/NEJMoa2032183](https://doi.org/10.1056/NEJMoa2032183)
841 pmid:<https://pubmed.ncbi.nlm.nih.gov/33567185>
- 842 23. Shaman AM, Bain SC, Bakris GL, *et al.* Effect of the Glucagon-Like Peptide-1 Receptor
843 Agonists Semaglutide and Liraglutide on Kidney Outcomes in Patients With Type 2 Diabetes: Pooled
844 Analysis of SUSTAIN 6 and LEADER. *Circulation* 2022;**145**:575-585.

- 845 24. Vieira-Silva S, Falony G, Belda E, *et al.* Statin therapy is associated with lower prevalence of
846 gut microbiota dysbiosis. *Nature* 2020;**581**:310-315.doi:10.1038/s41586-020-2269-x.
847 pmid:<https://pubmed.ncbi.nlm.nih.gov/32433607>
- 848 25. Brial F, Chilloux J, Nielsen T, *et al.* Human and preclinical studies of the host-gut microbiome
849 co-metabolite hippurate as a marker and mediator of metabolic health. *Gut* 2021; gutjnl-2020-
850 323314.doi: 10.1136/gutjnl-2020-323314 pmid:<https://pubmed.ncbi.nlm.nih.gov/33975870>
- 851 26. Lundberg SM, Erion G, Chen H, *et al.* From Local Explanations to Global Understanding with
852 Explainable AI for Trees. *Nat Mach Intell* 2020;2:56-67.doi: 10.1038/s42256-019-0138-9
853 pmid:<https://pubmed.ncbi.nlm.nih.gov/32607472>
- 854 27. Levey AS, Coresh J, Balk E, *et al.* National Kidney Foundation practice guidelines for chronic
855 kidney disease: evaluation, classification, and stratification. *Ann Intern Med* 2003;**139**:137-47.
856 doi:10.7326/0003-4819-139-2-200307150-00013 pmid:<https://pubmed.ncbi.nlm.nih.gov/12859163>
- 857 28. Vanholder R, Schepers E, Pletinck A, *et al.* The uremic toxicity of indoxyl sulfate and p-cresyl
858 sulfate: a systematic review. *J Am Soc Nephrol* 2014;**25**:1897-90.doi:10.1681/ASN.2013101062
859 pmid:<https://pubmed.ncbi.nlm.nih.gov/24812165>
- 860 29. Fromentin S, Forslund SK, Chechi K, *et al.* Microbiome and metabolome features of the
861 cardiometabolic disease spectrum. *Nat Med* 2022;**28**:303-314.doi:10.1038/s41591-022-01688-4.
862 pmid:<https://pubmed.ncbi.nlm.nih.gov/35177860>
- 863 30. Koeth RA, Levison BS, Culley MK, *et al.* γ -Butyrobetaine is a proatherogenic intermediate in
864 gut microbial metabolism of L-carnitine to TMAO. *Cell Metab* 2014;**20**:799-812.
865 doi:10.1016/j.cmet.2014.10.006 pmid:<https://pubmed.ncbi.nlm.nih.gov/25440057>
- 866 31. Ferrell M, Bazeley P, Wang Z, *et al.* Fecal Microbiome Composition Does Not Predict Diet-
867 Induced TMAO Production in Healthy Adults. *J Am Heart Assoc* 2021;**10**: e021934.
868 doi:10.1161/JAHA.121.021934 pmid: <https://pubmed.ncbi.nlm.nih.gov/34713713>

- 869 32. Titan SM, Venturini G, Padilha K, *et al.* Metabolites related to eGFR: Evaluation of candidate
870 molecules for GFR estimation using untargeted metabolomics. *Clin Chim Acta* 2019;**489**:242-248.
871 doi:10.1016/j.cca.2018.08.037 pmid:<https://pubmed.ncbi.nlm.nih.gov/30153452>
- 872 33. Altara R, Manca M, Hessel MH, *et al.* CXCL10 Is a Circulating Inflammatory Marker in Patients
873 with Advanced Heart Failure: a Pilot Study. *J Cardiovasc Transl Res* 2016;**9**:302-14.
874 doi:10.1007/s12265-016-9703-3 pmid:<https://pubmed.ncbi.nlm.nih.gov/27271043>
- 875 34. Forslund SK, Chakaroun R, Zimmermann-Kogadeeva M, *et al.* Combinatorial, additive and
876 dose-dependent drug-microbiome associations. *Nature* 2021;**600**:500-505.doi:10.1038/s41586-021-
877 04177-9. pmid:<https://pubmed.ncbi.nlm.nih.gov/34880489>
- 878 35. Schmitt R, Melk A. Molecular mechanisms of renal aging. *Kidney Int* 2017;**92**:569-
879 579.doi:10.1016/j.kint.2017.02.036 pmid:<https://pubmed.ncbi.nlm.nih.gov/28729036>
- 880 36. Gupta N, Buffa JA, Roberts AB, *et al.* Targeted Inhibition of Gut Microbial Trimethylamine N-
881 Oxide Production Reduces Renal Tubulointerstitial Fibrosis and Functional Impairment in a Murine
882 Model of Chronic Kidney Disease. *Arterioscler Thromb Vasc Biol* 2020;**40**:1239-1255.
883 doi:10.1161/ATVBAHA.120.314139 pmid:<https://pubmed.ncbi.nlm.nih.gov/32212854>
- 884 37. Fang Q, Zheng B, Liu N, *et al.* Trimethylamine N-Oxide Exacerbates Renal Inflammation and
885 Fibrosis in Rats With Diabetic Kidney Disease. *Front Physiol* 2021;**12**:682482.
886 doi:10.3389/fphys.2021.682482 pmid:<https://pubmed.ncbi.nlm.nih.gov/34220546>
- 887 38. Duffield JS. Cellular and molecular mechanisms in kidney fibrosis. *J Clin Invest*
888 2014;**124**:2299-306.doi:10.1172/JCI72267 pmid:<https://pubmed.ncbi.nlm.nih.gov/24892703>
- 889 39. Zhu W, Gregory JC, Org E, *et al.* Gut Microbial Metabolite TMAO Enhances Platelet
890 Hyperreactivity and Thrombosis Risk. *Cell* 2016;**165**:111-124.doi:10.1016/j.cell.2016.02.011
891 pmid:<https://pubmed.ncbi.nlm.nih.gov/26972052>
- 892 40. Andrikopoulos P, Eccles SA, Yaqoob MM. Coupling between the TRPC3 ion channel and the
893 NCX1 transporter contributed to VEGF-induced ERK1/2 activation and angiogenesis in human

- 894 primary endothelial cells. *Cell Signal* 2017;**37**:12-30 doi:10.1016/j.cellsig.2017.05.013
895 pmid:<https://pubmed.ncbi.nlm.nih.gov/28535874>
- 896 41. Andrikopoulos P, Kieswich J, Harwood SM, *et al.* Endothelial Angiogenesis and Barrier
897 Function in Response to Thrombin Require Ca²⁺ Influx through the Na⁺/Ca²⁺ Exchanger. *J Biol Chem*
898 2015;**290**:18412-28.doi:10.1074/jbc.M114.628156
899 pmid:<https://pubmed.ncbi.nlm.nih.gov/25979335>
- 900 42. Andrikopoulos P, Kieswich J, Pacheco S, *et al.* The MEK Inhibitor Trametinib Ameliorates
901 Kidney Fibrosis by Suppressing ERK1/2 and mTORC1 Signaling. *J Am Soc Nephrol* 2019;**30**:33-49.
902 doi:10.1681/ASN.2018020209 pmid:<https://pubmed.ncbi.nlm.nih.gov/30530834>
- 903 43. Chevalier RL, Forbes MS, Thornhill BA. Ureteral obstruction as a model of renal interstitial
904 fibrosis and obstructive nephropathy. *Kidney Int* 2009;**75**:1145-1152.doi:10.1038/ki.2009.86
905 pmid:<https://pubmed.ncbi.nlm.nih.gov/19340094>
- 906 44. Yap CX, Henders AK, Alvares GA, *et al.* Autism-related dietary preferences mediate autism-
907 gut microbiome associations. *Cell* 2021;**184**:5916-5931.e17.doi:10.1016/j.cell.2021.10.015
908 pmid:<https://pubmed.ncbi.nlm.nih.gov/34767757>
- 909 45. Rebouche CJ. Kinetics, pharmacokinetics, and regulation of L-carnitine and acetyl-L-carnitine
910 metabolism. *Ann N Y Acad Sci* 2004;**1033**:30-41. doi:10.1196/annals.1320.003
911 pmid:<https://pubmed.ncbi.nlm.nih.gov/15591001>
- 912 46. Cocking C, Walton J, Kehoe L, *et al.* The role of meat in the European diet: current state of
913 knowledge on dietary recommendations, intakes and contribution to energy and nutrient intakes
914 and status. *Nutr Res Rev* 2020;**33**:181-189. doi:10.1017/S0954422419000295
915 pmid:<https://pubmed.ncbi.nlm.nih.gov/31918784>
- 916 47. Zhao S, Feng XF, Huang T, *et al.* The Association Between Acylcarnitine Metabolites and
917 Cardiovascular Disease in Chinese Patients With Type 2 Diabetes Mellitus. *Front Endocrinol*
918 (*Lausanne*) 2020;**11**:212. doi:10.3389/fendo.2020.00212
919 pmid:<https://pubmed.ncbi.nlm.nih.gov/32431666>

- 920 48. Chen S, Henderson A, Petriello MC, *et al.* Trimethylamine N-Oxide Binds and Activates PERK
921 to Promote Metabolic Dysfunction. *Cell Metab* 2019;**30**:1141-1151.e5.
922 doi:10.1016/j.cmet.2019.08.021 pmid:<https://pubmed.ncbi.nlm.nih.gov/31543404>
- 923 49. Martinez-Klimova E, Aparicio-Trejo OE, Tapia E, *et al.* Unilateral Ureteral Obstruction as a
924 Model to Investigate Fibrosis-Attenuating Treatments. *Biomolecules* 2019;**9**:141.
925 doi:10.3390/biom9040141 pmid:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6523883>
- 926 50. Sureshbabu A, Muhsin SA, Choi ME. TGF- β signaling in the kidney: profibrotic and protective
927 effects. *Am J Physiol Renal Physiol* 2016;**310**:F596–F606. doi:10.1152/ajprenal.00365.2015
928 pmid:<https://pubmed.ncbi.nlm.nih.gov/26739888/>
- 929 51. Suthanthiran M, Gerber LM, Schwartz JE, *et al.* Circulating transforming growth factor-beta1
930 levels and the risk for kidney disease in African Americans. *Kidney Int* 2009;**76**:72-80.doi:
931 10.1038/ki.2009.66 pmid:<https://pubmed.ncbi.nlm.nih.gov/19279557/>
- 932 52. Gruppen EG, Garcia E, Connely MA, *et al.* TMAO is Associated with Mortality: Impact of
933 Modestly Impaired Renal Function. *Sci Rep* 2017;**7**:13781.doi:10.1038/s41598-017-13739-9
934 pmid:<https://pubmed.ncbi.nlm.nih.gov/29061990>
- 935 53. Meyer KA, Benton TZ, Bennett BJ, *et al.* Microbiota-Dependent Metabolite Trimethylamine
936 N-Oxide and Coronary Artery Calcium in the Coronary Artery Risk Development in Young Adults
937 Study (CARDIA). *J Am Heart Assoc* 2016;**5**:e003970.doi:10.1161/JAHA.116.003970
938 pmid:<https://pubmed.ncbi.nlm.nih.gov/27792658>
- 939

940 **Methods**

941 **MetaCardis study design and recruitment**

942 Study design, recruitment and exclusion criteria has been extensively described^{24,29,34,54}. Patients
943 were subclassified in three groups: BMI-spectrum patients (BMIS²⁹; N=837), encompassing
944 MetaCardis participants presenting with metabolic syndrome-related risk factors or conditions
945 (hypertension⁵⁵, obesity⁵⁶ and metabolic syndrome⁵⁷) and patients diagnosed with type-2 diabetes
946 (T2D⁵⁸; N=561) or ischaemic heart disease (IHD; N=356). The IHD group comprised patients with
947 Acute (<15days) Coronary Syndrome (ACS; N=106), Chronic IHD (CIHD; N=157) with normal Left
948 Ventricular Ejection Fraction (LVEF) determined by echocardiography and Heart Failure patients (HF;
949 N=93, LVEF<45%). Ethical approval was obtained from the Ethics Committee CPP Ile-de France, the
950 Ethical Committees of the Capital Region of Denmark (H-3-2013-145) and Ethics Committee at the
951 Medical Faculty at the University of Leipzig (047-13-28012013). Study participants provided written
952 informed consent and the study was undertaken according to Helsinki Declaration-II.

953 **Sample and phenotypic information collection**

954 Biofluid and biomatter collection has been described elsewhere^{24,29,34,54}. Briefly, blood samples were
955 collected in the morning after overnight fasting and fecal samples were collected at home by
956 participants, frozen immediately and transferred to study centers on dry ice within 48h. All samples
957 were stored at -80°C until use. Clinical history, medication and phenotypic information were
958 acquired as described^{23,28,34,54} with standardized procedures across centers. Participants reported
959 habitual food intake through a customized Food Frequency Questionnaire (FQQ)⁵⁹. Bioclinical
960 variables were measured in a single center according to standard procedures²⁸. Estimated
961 glomerular filtration rate (eGFR) was calculated with the CKD-EPI formula without ethnicity
962 adjustment⁶¹.

963 **Metabolic profiling**

964 **¹H-Nuclear Magnetic Resonance (¹H NMR) spectroscopy**

965 Spectra acquisition, using an Avance spectrometer (Bruker) at 600 MHz; and structural assignments
966 have been extensively described previously^{25,29}. Absolute quantifications were derived using the “*In*
967 *Vitro* Diagnostics for research” (IVDr) quantification BI-LISA and BI-Quant algorithms (Bruker).

968 **Gas Chromatography coupled Mass Spectrometry (GC-MS)**

969 Serum samples (100µl) were prepared, analyzed and processed as described^{29,35}. Briefly, protein was
970 methanol-precipitated, methanol was evaporated to dryness and subsequent to derivatization
971 samples were injected to an Agilent 7890B-5977B Inert Plus GC-MS system. The chromatographic
972 column was an Agilent ZORBAX DB5- MS (30m x 250µm x 0.25µm + 10m Duragard). The
973 temperature gradient was 37.5min long and the mass analyzer was operated in full scan mode
974 between 50 and 600 m/z. Peaks were annotated with the use of the Fiehn library (Agilent G1676AA
975 Fiehn GC/MS Metabolomics RTL Library, User Guide, Agilent Technologies,
976 https://www.agilent.com/cs/library/usermanuals/Public/G1676-90001_Fiehn.pdf). Metabolic
977 features with low reproducibility or linearity were removed from the dataset, resulting in 102
978 annotated metabolic features.

979 **Ultra-Performance Liquid Chromatography-Tandem Mass Spectrometry (UPLC-MS/MS)**

980 UPLC-MS/MS performed on a Waters Acquity UPLC-Xevo TQ-S UPLC-MS/MS system equipped with
981 an Acquity BEH HILIC (2.1×100mm, 1.7µm) chromatographic column was employed to determine
982 TMA, TMAO, choline, betaine, γ-butyrobetaine, betaine aldehyde, butyryl-carnitine, isovaleryl-
983 carnitine, OH-isovaleryl-carnitine, stearoyl-carnitine, oleoyl-carnitine, linoleoyl-carnitine, myristoyl-
984 carnitine, lauroyl-carnitine and decanoyl-carnitine as described previously^{28,34}.

985 TMAO circulating values of all MetaCardis participants were log₁₀-transformed and subsequently the
986 median was subtracted and divided by the standard deviation²¹ (SD; regularized TMAO values
987 thereafter).

988 **Metagenomic analysis.**

989 Phylogenetic microbiota profiles were built after correction for bacterial load as extensively
990 described^{24,29,35,54} using a protocol devised by Vandeputte *et al*⁶¹ with modifications. Briefly, total
991 faecal DNA was extracted following the International Human Microbiome Standards (IHMS)
992 guidelines (SOP 07 V2 H) and samples were sequenced using ion-proton technology (ThermoFisher
993 Scientific). Gene abundance profiling was performed using the 9.9 million gene integrated reference
994 catalog of the human microbiome, as described^{24,29,34,54}.

995 **Customized microbial module analysis (GMM)**

996 Manually-curated customized module sets focusing on anaerobic bacterial and archaeal
997 fermentation processes relevant to the human gut microbiota were assembled as previously
998 extensively described^{24,29,34,54}.

999 **Statistical analyses**

1000 Statistical analysis was undertaken using R (v4.03)⁶². For comparing two groups we used Mann-
1001 Whitney U and for multiple group comparisons Kruskal-Wallis tests. Unadjusted Spearman
1002 correlations were computed using R, whilst adjusted Spearman correlations with ppcor (v1.1). *P*-
1003 values were corrected for multiple comparisons using the Benjamini-Hochberg method.

1004 **Machine Learning (ML) analysis**

1005 **Variable groups**

1006 Patient phenotypic variables were separated into 10 groups (**Suppl.Table.5**). Specifically, biological
1007 parameters included biochemical and clinical serum laboratory tests including lipids, glycated
1008 haemoglobin, creatinine and eGFR. Clinical parameters consisted of clinical history, BMI, systolic and
1009 diastolic blood pressure, anthropometric variables and stool frequency and type. Demographic
1010 information comprised age, physical activity, educational and income levels, smoking status,
1011 ethnicity and country of recruitment. Drug variables included intake of common medication as

1012 described³⁴, number of antibiotic courses in the last 5 years and number of antihypertensive, anti-
1013 diabetic and lipid-lowering treatments. Dietary parameters included habitual consumption of 37
1014 food items, daily nutrients intake derived from these food items calculated as in⁵⁹, alternative
1015 Healthy Eating⁶³ (aHEI), Dietary Approaches to Stop Hypertension⁶⁴ (DASH) and Dietary Diversity⁶⁵
1016 (DDS) scores. Clinical risk factors variables included age, systolic and diastolic blood pressure,
1017 glycated haemoglobin levels, fasting cholesterol levels, smoking status and waist circumference as
1018 described²⁹. Serum metabolomics comprised the absolute or relative levels of 116 circulating
1019 metabolites determined by GC-MS or UPLC-MS/MS^{29,34}. Urine metabolomics included absolute
1020 quantification of 47 urine metabolites from ¹H-NMR spectra calculated with the IVDr algorithm.
1021 Microbiota variables included abundance of 699 bacterial species present in at least 20% of
1022 MetaCardis patients corrected for microbial load and the first 10 principal components from a PCA
1023 of relative microbial gene abundances^{29,34}. Microbiome (Modules) group included abundance of 116
1024 manually curated bacterial modules²⁹. In all cases categorical variables were converted into dummy
1025 variables using caret (v6.0.86).

1026 **Boosted decision trees (Xgboost)**

1027 We predicted regularized circulating TMAO levels by using gradient boosting decision trees based on
1028 the xgboost algorithm (v1.3.2.1)⁶⁶, co-opting a strategy from Bar *et. al*²¹. Xgboost consistently
1029 outperforms other algorithms in Kaggle competitions for tabular data. For each of our 10 variable
1030 groups (**Suppl.Table.5**) we optimized xgboost models using 5-fold cross-validation and two
1031 sequential hyperparameter grids searches (972 different parameter combinations for each feature
1032 group in total) to predict mean-centred and unit variance-scaled (regularized) TMAO levels in the
1033 left-out group using root-mean-square error (RMSE) to evaluate model outcomes. After parameter
1034 optimization, we predicted circulating TMAO using 5-fold cross-validation for 100 iterations using as
1035 input the variables of each group or all variables (full-model). For each round, we calculated the
1036 coefficient of determination using the rsq function from yardstick (v0.0.7) and the predicted

1037 regularized TMAO values. Xgboost models trained with 80% of each patient group participants
1038 during cross-validation (5 for each round) were saved and used for feature attribution analysis in the
1039 left-out group (see below).

1040 **SHapley Additive exPlanations (SHAP) analysis**

1041 We interpreted our ML models and assigned relative importance to variables influencing circulating
1042 TMAO levels by co-opting SHAP values, as expanded for tree-based ML models²⁶ and recently used
1043 to objectively evaluate factors driving metabolite plasma levels in humans²¹. Briefly, for each
1044 prediction the SHAP value of a particular variable is the difference in the model output when this
1045 variable is included *versus* when excluded. Variables are added to the model in all possible orderings
1046 and the SHAP value is computed from the average of model outcomes²⁶. Variable SHAP values for
1047 each individual in the 5 left-out groups were extracted from corresponding xgboost models trained
1048 on the remaining 80% participants for each variable group using SHAPforxgboost (0.1.0) and
1049 averaged over 100 iterations. We assigned relative importance to each variable by computing the
1050 mean of the absolute SHAP value for all individuals in each disease group similarly to Bat *et al.*²¹. For
1051 swarm-plots depicting each individual participants SHAP values, variable values were inverse-
1052 normalized using the RNOmni (v1.0.0) package.

1053 **Clustering**

1054 Clustering was performed using the built-in R k-means function with the Hartigan-Wong algorithm
1055 using 25 random sets, as described²⁵.

1056 **Multivariate analysis**

1057 To identify microbiota composition differences between individuals split in TMAO clusters we
1058 performed multivariate homogeneity analysis of Bray-Curtis dissimilarity matrices at the species
1059 level (699 species present in at least 20% of our population) and determined statistical significance
1060 with a permutation ANOVA test (999 iterations) using the vegan (v2.5.7) package. Permutation

1061 analysis of variance (PERMANOVA) of regularized circulating TMAO versus Bray-Curtis taxonomic
1062 dissimilarity matrices with age, sex and country of recruitment as co-variates for 999 iterations were
1063 performed with vegan (v2.5.7).

1064 **Mediation analysis**

1065 We performed mediation analysis⁶⁷ to assess the putative impact of eGFR on plasma TMAO increase
1066 with age and conversely the impact of TMAO on eGFR decline with age. We used the Preacher and
1067 Hayes bootstrapping method, as implemented in the mediation (v4.5.0) R package⁶⁸, using general
1068 linear models with sex and country of recruitment as covariates. Confidence intervals and Bayesian *P*
1069 values were computed after 999 simulations.

1070 **Propensity score matching**

1071 MetaCardis participants with T2D were propensity-score matched using the R package MatchIt⁶⁹
1072 (v.4.1.0) with age, sex, disease group and hypertension status as covariates using nearest neighbor
1073 matching determined by generalized linear models. All covariates were given equal weight.

1074 **Preclinical models**

1075 A complete list of antibodies and reagents used in these studies are appended (**Suppl.Table.24**)

1076 ***In vitro* experiments**

1077 **Cell Culture and Protein Extraction**

1078 Primary human adult kidney fibroblasts from a single donor (Cambridge Bioscience) were cultured in
1079 10% FCS low-Glucose DMEM with 1% Pen/Strep antibiotics (Sigma). For acute (up-to 30min) TMAO
1080 stimulation, cells (100,000/condition) were serum-starved for 1h in a physiologic serum-free buffer
1081 as previously described⁴³. Fibroblasts were preincubated with Tramenitib (10nM, Shelleckchem),
1082 BAPTA-AM (20μM, Molecular Probes) or vehicle for 30min before TMAO (Sigma) stimulation. For
1083 longer-term (24h) stimulation, fibroblasts were in serum-free low-glucose DMEM overnight and

1084 subsequently stimulated with TMAO, TGF- β 1 (5ng/ml, R&D), or their combination. At the end of the
1085 experiment fibroblasts were lysed and stored at -80°C until further use, as described⁴².

1086 **Cytosolic [Ca²⁺] Measurements**

1087 Fibroblasts were serum-starved overnight in low-glucose DMEM and subsequently loaded with 5 μM
1088 fura-2-AM (Molecular Probes) in pH 7.4 Hanks Balanced Salt Solution (HBSS; Sigma) containing Ca²⁺
1089 and Mg²⁺. Measurements were obtained on an epifluorescence inverted microscope equipped with
1090 a $\times 20$ fluorite objective. Single cell intracellular Ca²⁺ ([Ca²⁺]_i) was monitored using excitation at 340
1091 and 380nm, through a monochromator (Cairn Research). Emitted light was reflected through a
1092 515nm filter to a QImaging Retiga CCD camera (Cairn Research) and digitized to 12-bit resolution. All
1093 imaging data were collected and analyzed using software from Andor.

1094 **Animal Procedures**

1095 All animal experiments were conducted in accordance with the United Kingdom Home Office
1096 Animals 1986 Scientific Procedures, with local ethical committee approval (Project License 70/8356).
1097 Male C57BL/6J mice (Charles River) at 6-8weeks of age were fed a control, 0.12% w/w TMAO- or 1%
1098 w/w choline-containing diets (Teklad; 6animals/group) for 6weeks¹⁸. Subsequently, Unilateral Ureter
1099 Obstruction (UUO) was established as previously described⁴². Briefly, mice were anesthetized with
1100 isoflurane and the abdominal cavity was exposed using midline laparotomy, the right ureter was
1101 isolated and tied off 0.5cm from the pelvis. The left ureter was left unclamped and served as the
1102 sham-operated control. At day 5 the kidneys were harvested. Both kidneys from each animal (UUO
1103 and sham) were cut in half longitudinally. One half of each kidney was snap-frozen in liquid nitrogen
1104 and subsequently stored at -80°C for Western blotting. The other half was fixed with 10% Formalin
1105 (Sigma) for 16h at 4 $^{\circ}\text{C}$, transferred to a 70% v/v ethanol solution for a further 24h and was then
1106 paraffin-embedded for immunohistochemistry.

1107 **Western Blot**

1108 We used the NuPAGE electrophoresis and buffer system (Invitrogen) for immunoblot analysis of
1109 kidney samples or cellular lysates as described^{40,41}. Proteins were visualized with ECL Prime (GE
1110 Healthcare). Optical densities of bands of interest were determined using ImageJ 1.46r (NIH) and
1111 normalized against loading controls. The value of the normalized control sample was arbitrarily set
1112 to 1. Membranes were stripped using the Restore Plus reagent (Fisher Scientific), and re-probed with
1113 appropriate loading controls.

1114 **Immunohistochemical Staining**

1115 Formalin-fixed kidneys were embedded in paraffin and 4µm sections were cut by the Barts Cancer
1116 Institute Pathology Unit. Staining for Sirius red, α-Smooth Muscle Actin (αSMA) and F4/F80 were
1117 performed as described⁴² with the DISCOVERY XT (Ventana) automated slide processing instrument
1118 using the OmniMap reagents (Ventana), according to the manufacturer's recommendations. Images
1119 were captured at ×20 magnification, using a Zeiss AxioPhot microscope with an AxioCam HRc
1120 camera. Five kidney cortex images were captured per mouse and staining was quantified as
1121 percentage of total area, using ImageJ 1.46r (NIH).

1122 **References**

- 1123 54. Belda E, Voland L, Tremaroli V, *et al.* Impairment of gut microbial biotin metabolism and host
1124 biotin status in severe obesity: effect of biotin and prebiotic supplementation on improved
1125 metabolism. *Gut* 2022; gutjnl-2021-325753. doi: 10.1136/gutjnl-2021-325753
1126 pmid:<https://pubmed.ncbi.nlm.nih.gov/35017197>
- 1127 55. Yancy CW, Jessup M, Bozkurt B, *et al.* 2017 ACC/AHA/HFSA Focused Update of the 2013
1128 ACCF/AHA Guideline for the Management of Heart Failure: A Report of the American College of
1129 Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart
1130 Failure Society of America. *Circulation* 2017;**136**:e137-e161. doi: 10.1161/CIR.0000000000000509
1131 pmid:<https://pubmed.ncbi.nlm.nih.gov/28455343>

- 1132 56. The challenge of obesity in the WHO European region and the strategies for response:
1133 summary. (World Health Organization, Regional Office for Europe, 2007)
1134 <https://www.euro.who.int/en/publications/abstracts/challenge-of-obesity-in-the-who-european->
1135 [region-and-the-strategies-for-response-the](https://www.euro.who.int/en/publications/abstracts/challenge-of-obesity-in-the-who-european-region-and-the-strategies-for-response-the)
- 1136 57. Alberti KG, Zimmet P, Shaw J, *et al.* The metabolic syndrome--a new worldwide definition.
1137 *Lancet* 2005;**366**:1059-62.doi: 10.1016/S0140-6736(05)67402-8
1138 pmid:<https://pubmed.ncbi.nlm.nih.gov/16182882>
- 1139 58. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes*
1140 *Care* 2004;**27** Suppl 1:S5-S10.doi: 10.2337/diacare.27.2007.s5
1141 pmid:<https://pubmed.ncbi.nlm.nih.gov/14693921>
- 1142 59. Verger EO, Armstrong P, Nielsen T, *et al.* Dietary Assessment in the MetaCardis Study:
1143 Development and Relative Validity of an Online Food Frequency Questionnaire. *J Acad Nutr Diet*
1144 2017;**117**:878-888.doi:10.1016/j.jand.2016.10.030 pmid:<https://pubmed.ncbi.nlm.nih.gov/2802480>
- 1145 60. Levey AS, Stevens LA, Schmidt CH, *et al.* A new equation to estimate glomerular filtration
1146 rate. *Ann Intern Med* 2009;**150**:604-12.doi: 10.7326/0003-4819-150-9-200905050-00006
1147 pmid:<https://pubmed.ncbi.nlm.nih.gov/19414839>
- 1148 61. Vandeputte D, Kathagen G, D'hoel K. *et al.* Quantitative microbiome profiling links gut
1149 community variation to microbial load. *Nature* 2017;**551**:507-511.doi:10.1038/nature24460
1150 pmid:<https://pubmed.ncbi.nlm.nih.gov/29143816>
- 1151 62. R Core Team (2020). R: A language and environment for statistical computing. R Foundation
1152 for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- 1153 63. InterAct Consortium. Adherence to predefined dietary patterns and incident type 2 diabetes
1154 in European populations: EPIC-InterAct Study. *Diabetologia* 2014;**57**:321-33.doi:10.1007/s00125-
1155 013-3092-9 pmid:<https://pubmed.ncbi.nlm.nih.gov/24196190>

- 1156 64. Tyson CC, Lin PH, Corsino L, *et al.* Short-term effects of the DASH diet in adults with
1157 moderate chronic kidney disease: a pilot feeding study. *Clin Kidney J* 2016;**9**:592-8.
1158 doi:10.1093/ckj/sfw046 pmid:<https://pubmed.ncbi.nlm.nih.gov/27478603>
- 1159 65. Jeurnink SM, Büchner FL, Bueno-de-Mesquita HB, *et al.* Variety in vegetable and fruit
1160 consumption and the risk of gastric and esophageal cancer in the European Prospective Investigation
1161 into Cancer and Nutrition. *Int J Cancer* 2012;**131**:E963-73.doi:10.1002/ijc.27517
1162 pmid:<https://pubmed.ncbi.nlm.nih.gov/22392502>
- 1163 66. Chen T., & Guestrin C. (2016). XGBoost: A Scalable Tree Boosting System. In Proceedings of
1164 the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–
1165 794). New York, NY, USA: ACM. <https://doi.org/10.1145/2939672.2939785>
- 1166 67. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annu Rev Psychol* 2007;**58**:593-
1167 614. doi:10.1146/annurev.psych.58.110405.085542
1168 pmid:<https://pubmed.ncbi.nlm.nih.gov/16968208/>
- 1169 68. Tingley D, Yamamoto T, Hirose K *et al.* mediation: R Package for Causal Mediation Analysis. *J*
1170 *Stat Softw* 2014;**59**:1–38. doi:10.18637/jss.v059.i05
- 1171 69. Ho DE, Imai K, King G, *et al.* MatchIt: Nonparametric Preprocessing for Parametric Causal
1172 Inference. *J Stat Softw* 2011;**42**:1-28. doi:10.18637/jss.v042.i08