

RESEARCH

Open Access



# Plasma metabolomics provides new insights into the relationship between metabolites and outcomes and left ventricular remodeling of coronary artery disease

Qian Zhu<sup>1,2,3†</sup>, Min Qin<sup>1,2,3†</sup>, Zixian Wang<sup>1,2</sup>, Yonglin Wu<sup>1,2</sup>, Xiaoping Chen<sup>4</sup>, Chen Liu<sup>5</sup>, Qilin Ma<sup>6</sup>, Yibin Liu<sup>1,2,3</sup>, Weihua Lai<sup>1</sup>, Hui Chen<sup>1,3</sup>, Jingjing Cai<sup>7</sup>, Yemao Liu<sup>7</sup>, Fang Lei<sup>7</sup>, Bin Zhang<sup>2,3</sup>, Shuyao Zhang<sup>8</sup>, Guodong He<sup>2,3</sup>, Hanping Li<sup>2</sup>, Mingliang Zhang<sup>9</sup>, Hui Zheng<sup>9</sup>, Jiyan Chen<sup>2</sup>, Min Huang<sup>10</sup> and Shilong Zhong<sup>1,2,3\*</sup>

## Abstract

**Background:** Coronary artery disease (CAD) is a metabolically perturbed pathological condition. However, the knowledge of metabolic signatures on outcomes of CAD and their potential causal effects and impacts on left ventricular remodeling remains limited. We aim to assess the contribution of plasma metabolites to the risk of death and major adverse cardiovascular events (MACE) as well as left ventricular remodeling.

**Results:** In a prospective study with 1606 Chinese patients with CAD, we have identified and validated several independent metabolic signatures through widely-targeted metabolomics. The predictive model respectively integrating four metabolic signatures (dulcitol,  $\beta$ -pseudouridine, 3,3',5'-Triiodo-L-thyronine, and kynurenine) for death (AUC of 83.7% vs. 76.6%, positive IDI of 0.096) and metabolic signatures (kynurenine, lysoPC 20:2, 5-methyluridine, and L-tryptophan) for MACE (AUC of 67.4% vs. 59.8%, IDI of 0.068) yielded better predictive value than trimethylamine N-oxide plus clinical model, which were successfully applied to predict patients with high risks of death ( $P = 0.0014$ ) and MACE ( $P = 0.0008$ ) in the multicenter validation cohort. Mendelian randomisation analysis showed that 11 genetically inferred metabolic signatures were significantly associated with risks of death or MACE, such as 4-acetamidobutyric acid, phenylacetyl-L-glutamine, tryptophan metabolites (kynurenine, kynurenic acid), and modified nucleosides ( $\beta$ -pseudouridine, 2-(dimethylamino) guanosine). Mediation analyses show that the association of these metabolites with the outcomes could be partly explained by their roles in promoting left ventricular dysfunction.

**Conclusions:** This study provided new insights into the relationship between plasma metabolites and clinical outcomes and its intermediate pathological process left ventricular dysfunction in CAD. The predictive model integrating metabolites can help to improve the risk stratification for death and MACE in CAD. The metabolic signatures appear to

<sup>†</sup>Qian Zhu and Min Qin contributed equally to the study and are considered co-first authors

\*Correspondence: [gdph\\_zhongsl@gd.gov.cn](mailto:gdph_zhongsl@gd.gov.cn)

<sup>2</sup> Guangdong Provincial Key Laboratory of Coronary Heart Disease Prevention, Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510080, Guangdong, China

Full list of author information is available at the end of the article

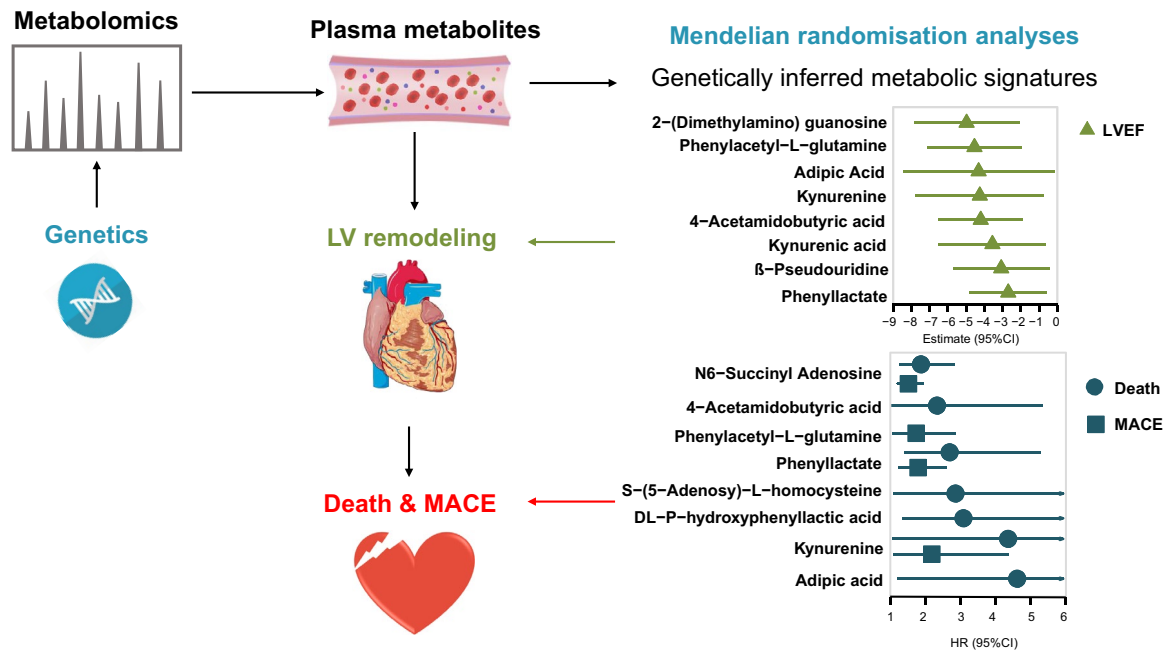


increase death or MACE risks partly by promoting adverse left ventricular dysfunction, supporting potential therapeutic targets of CAD for further investigation.

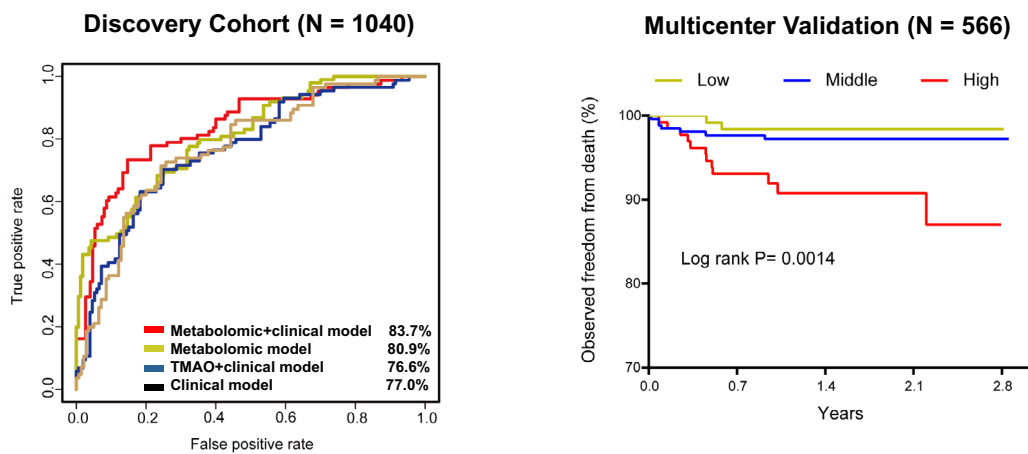
**Keywords:** Metabolomics, Coronary artery disease, Death, Major adverse cardiovascular events, Metabolic signature, Left ventricular remodeling, Mendelian randomisation

**Graphical Abstract**

**Metabolomics Identified Metabolic Signatures with Novel Linkages in LV Remodeling and Outcomes**



**Metabolomic Model for Risk Stratification of Death in CAD**



## Background

Coronary artery disease (CAD) imposes a major burden on modern society with annual morbidity and mortality comparable to cancer [1–3]. Despite the advances in pharmaceutical and operative treatments, the mortality of CAD remains unacceptably high. It is well-known that significant left ventricular (LV) remodeling is a severe and common issue in CAD [4], which contributes to the development of heart failure and the noticeably elevated risk of mortality and cardiovascular events [5, 6]. However, available molecular biomarkers in reflecting progressive worsening cardiac function and the long-term outcomes during CAD progression are still limited [7–9]. Incorporating novel molecular biomarkers to explain detailed molecular processes and pathophysiological mechanisms involving adverse outcomes, may improve early risk stratification and indicate novel targets in preventive therapies for CAD patients.

CAD is a metabolically perturbed pathological condition [10, 11]. Plasma metabolites reflect a functional output for genetic makeup and environmental exposure to disease phenotypes [12, 13]. Metabolomics profiling can assist in shedding light on underlying molecular mechanisms involving the pathophysiology of disease states and support personalized risk prediction for disease development and prognosis [13]. In the setting of prevalent CAD, emerging metabolomic profiling has begun to illustrate specific molecular signatures relating to CAD characterization [10]. Besides, previous studies have also identified a few metabolic markers for cardiovascular events or all-cause mortality [14–18], mainly in Westerners or free of cardiovascular disease. Moreover, researchers have indicated that different study populations may lead to different molecular signatures and conflicting findings [19–22]. While most metabolomic studies have little focus on further explanation of the causal effects and potential pathophysiologic mechanisms of metabolic markers. The mechanism promoting death and major adverse events (MACE) occurrence in CAD patients are heterogeneous, the potential pathological mechanism of adverse LV remodeling during CAD progression has become one of the hot issues [23]. However, few studies have examined the association of metabolites with clinical outcomes and cardiac remodeling in CAD, and the potential causal roles that the spectrum of metabolites play during the disease progression, before clinical endpoints onset, is thus unclear.

Given the fact that metabolism has been closely implicated in CAD pathogenesis and development, there is a need to identify novel metabolic signatures and integrate the genetic regulation of circulating metabolites to improve understanding in causality, and thus provide potential pathological insights and therapeutic targets

to improve CAD survival and prognosis. Therefore, we performed a study on 1606 Chinese patients with CAD through metabolomic profiling in plasma to evaluate the contribution of metabolites to the risks of death or MACE as well as LV remodeling and then built a good prognostic model based on metabolic signatures. Moreover, we further conducted Mendelian randomisation (MR) analysis by integrating genomic data to infer the potential causal effects of metabolites, and mediation analysis to explore possible mediation effects through promoting LV remodeling.

## Results

### Patient characteristics

The baseline characteristics are listed in Table 1. We recruited 1606 patients with CAD, including 1040 patients for the discovery cohort (63.03 years, 79.62% male), and 566 multicentre patients for the multicenter validation cohort (62 years, 74.16% male). Patients with high death and MACE risks were commonly old, suffered from diabetes, high aspartate aminotransferase, SYN-TAX scores, N-terminal-pro brain natriuretic peptide (proBNP) and LVMI, and low estimated glomerular filtration rate and LVEF, see Additional file 1: Table S1.

It is noteworthy that adverse LV remodeling manifested as low LVEF and high LVMI was strongly related to increased death or MACE risks in CAD (Additional file 2: Fig. S1), compared with other clinical characteristics including coronary lesion score. The association between characteristics with LVEF and LVMI is included in Additional file 1: Table S2.

### Metabolomic associations with the clinical outcomes

In the discovery phase, among the 202 metabolites, we identified 35 plasma metabolites that were significantly associated with the risks of death, and 24 metabolites remained significant after adjustment for confounding factors ( $FDR < 0.05$ ), see Fig. 1A and Additional file 1: Table S3. The most significant metabolites including 4-acetamidobutyric acid (HR, 1.60; 95% CI 1.38–1.87;  $FDR = 3.27E-08$ ),  $\beta$ -pseudouridine (HR, 1.77; 95% CI 1.41–2.23;  $FDR = 9.64E-06$ ), dulcitol (HR, 1.44; 95% CI 1.25–1.67;  $FDR = 7.10E-06$ ), (2-(dimethylamino) guanosine (HR, 1.73; 95% CI 1.29–2.31;  $FDR = 6.73E-04$ ), S-(5-adenosyl)-L-homocysteine (HR, 2.03; 95% CI 1.49–2.77;  $FDR = 3.92E-05$ ) and kynurenine (HR, 2.07; 95% CI 1.41–3.03;  $FDR = 6.65E-04$ ). In the multicenter validation cohort, 15 metabolites positively associated with the death risk were reproduced, including 4-acetamidobutyric acid,  $\beta$ -pseudouridine, 2-(dimethylamino) guanosine, S-(5-Adenosyl)-L-homocysteine, kynurenine, cyclic AMP, adipic acid, 3-methylcrotonyl glycine, 5'-deoxy-5'-(methylthio) adenosine, DL-P-hydroxyphenyllactic acid,

**Table 1** Baseline characteristics in 1606 CAD patients

Characteristics	Discovery cohort (n = 1040)	Multicenter validation cohort (n = 566)
Demographic data		
Age	63.03 ± 10.04	62.29 ± 10.18
Sex (male)	828 (79.62)	419 (74.16)
BMI, kg/m <sup>2</sup>	24.28 ± 4.79	24.06 ± 3.38
SBP, mm Hg	130.66 ± 18.89	133.04 ± 20.29
DBP, mm Hg	76.19 ± 11.03	76.46 ± 12.05
Current smoking	294 (28.52)	160 (28.73)
Family of CVD	29 (2.79)	–
Comorbidities		
Arrhythmia	92 (8.86)	51 (9.17)
DM	286 (27.55)	164 (29.39)
HyperT	627 (60.35)	340 (60.93)
Dyslipidemia	729 (72.54)	400 (74.07)
Biomedical measurements		
ALT, U/L	27.41 ± 13.18	27.65 ± 24.56
AST, U/L	26.64 ± 10.62	32.12 ± 55.48
eGFR, mL/min/1.73 m <sup>2</sup>	94.32 ± 73.69	91.37 ± 110.84
GLUC, mmol/L	6.74 ± 2.73	6.21 ± 3.82
CHOL, mmol/L	4.28 ± 1.12	4.29 ± 1.77
LDLC, mmol/L	2.58 ± 0.93	2.7 ± 1.00
HDLC, mmol/L	0.97 ± 0.26	0.99 ± 0.25
TRIG, mmol/L	1.62 ± 1.14	1.85 ± 1.85
CKMB, U/L	7.48 ± 5.92	19.37 ± 52.83
proBNP, pg/mL	774.51 ± 1597.35	1299.46 ± 4922.09
Medications		
BB	929 (89.5)	477 (84.57)
ACEI	660 (63.58)	286 (50.71)
CCB	295 (28.42)	165 (30.05)
PPI	506 (48.75)	380 (67.26)
SYNTAX score	16.43 ± 10.74	16.45 ± 13.09
LVEF, %	60.1 ± 11.54	59.43 ± 11.63
LVMI, g/m <sup>2</sup>	122.13 ± 36.05	116.54 ± 35.08

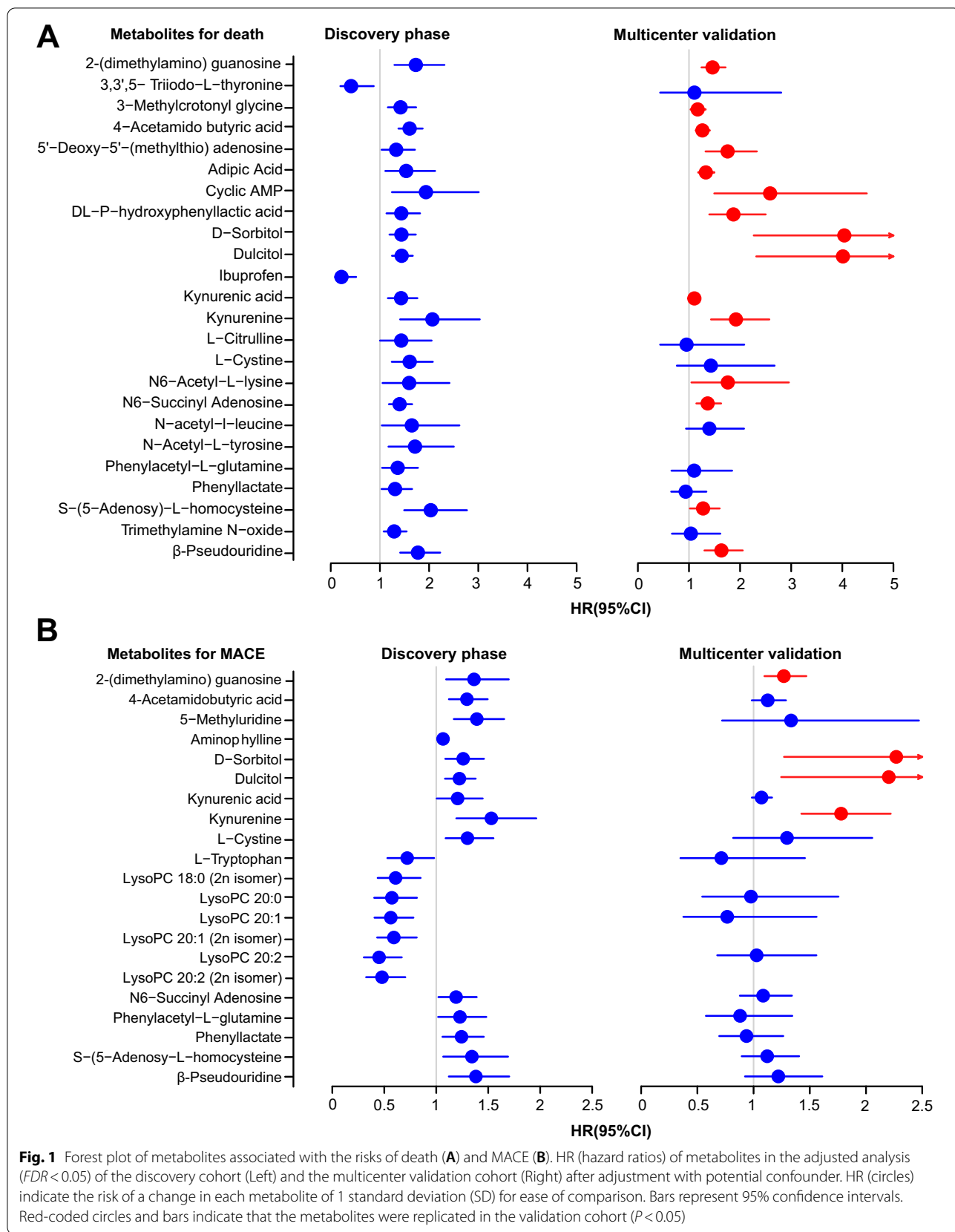
Data are number (%) or mean ± SD when appropriate

SD standard deviation, BMI body mass index, SBP systolic blood pressure, CVD cardiovascular disease, DM diabetes, HyperT hypertension, ALT alanine aminotransferase, AST aspartate aminotransferase, eGFR estimated glomerular filtration rate, GLUC glucose, CHOL cholesterol, LDLC low-density lipoprotein cholesterol, HDLC high-density lipoprotein cholesterol, TRIG triglyceride, CKMB creatine kinase MB, proBNP N-terminal pro brain natriuretic peptide, BB β-blockers, ACEI angiotensin converting enzyme inhibitors, CCB calcium channel blockers, PPI proton pump inhibitors, SYNTAX score Synergy between PCI with TAXUS and Cardiac Surgery score, LVEF left ventricular ejection fraction, LVMI left ventricular mass index

D-sorbitol, Dulcitol, kynurenic acid, N6-succinyl adenosine, and N6-acetyl-L-Lysine.

Twenty-one metabolites were significantly associated with the risks of MACE after adjustment for confounders ( $FDR < 0.05$ ) in the discovery phase, see Fig. 1B and Additional file 1: Table S4. Notably, 4-acetamidobutyric acid (HR, 1.42; 95% CI 1.25–1.60;  $FDR = 8.87E-06$ ) was also the most significant metabolites. During multicenter validation, 2-(dimethylamino) guanosine (HR, 1.27; 95% CI

1.10–1.47;  $FDR = 1.19E-02$ ), dulcitol (HR, 2.20; 95% CI 1.25–3.90;  $FDR = 2.82E-02$ ), D-sorbitol (HR, 2.27; 95% CI 1.27–4.05;  $FDR = 3.17E-02$ ) and kynurenine (HR, 1.78; 95% CI 1.43–2.22;  $FDR = 5.88E-06$ ) were well-replicated to be positively associated with MACE risks. 4-Acetamidobutyric acid, β-pseudouridine, kynurenic acid, and L-tryptophan were significant only before adjustment. In particular, kynurenine was validated to be robustly associated with both death and MACE risks.



**Table 2** Independent metabolic signature selection using LASSO

Terms	Coefficient ( $\beta$ )	HR	Frequency
LASSO based signature selection for death			
Dulcitol	0.22	1.25	200
4-Acetamidobutyric acid	0.29	1.34	200
N6-succinyl adenosine	0.05	1.05	195
L-Cystine	0.08	1.08	191
$\beta$ -Pseudouridine	0.05	1.05	173
2-(Dimethylamino) guanosine	0.02	1.02	137
Kynurenine	0.04	1.04	43
3,3',5-Triiodo-L-thyronine	-0.12	0.89	43
D-Sorbitol	0.04	1.04	21
DL-P-hydroxyphenyllactic acid	0.02	1.03	21
Phenyllactate (PLA)			
Cyclic AMP	0.02	1.02	5
S-(5-Adenosy)-L-homocysteine	0.02	1.02	2
LASSO based signature selection for MACE			
4-Acetamidobutyric acid	0.06	1.06	200
L-Cystine	0.06	1.06	200
L-Tryptophan	-0.24	0.79	200
Dulcitol	0.10	1.10	200
5-Methyluridine	0.28	1.33	200
Kynurenine	0.22	1.25	200
Phenyllactate (PLA)	0.10	1.11	200
LysoPC 20:2	-0.51	0.60	200
D-Sorbitol	0.02	1.02	199
LysoPC 20:1	-0.04	0.96	193
N6-succinyl adenosine	-0.01	0.99	2

The regression coefficients were calculated by averaging the coefficients obtained from tenfold cross-validation lasso Cox regression with 200 repeats, adjusted for 17 main clinical confounders. The confounders included age, sex, AST, eGFR, DM, HyperT, CHOL, HDLC, PPI, ACEI, BB, CCB, current smoking, family history of CVD, SYNTAX, SBP, and GLUC. The variables that appear zero times were removed and the variables left were further selected to develop a predictive model, abbreviations are as in Table 1

### Independent metabolic signatures and optimized prediction model for clinical outcomes

To identify a minimal set of metabolites and develop a prognostic model, we firstly selected independent metabolic signatures based on the 24 metabolites associated with the risks of death from the discovery cohort, using lasso Cox regression adjusting for potential confounders (200 repeats, Table 2). Then, irrelevant metabolites were filtered out and 13 endogenous metabolites were retained, including 4-acetamidobutyric acid, dulcitol, N6-succinyl adenosine, L-cystine,  $\beta$ -pseudouridine, 2-(dimethylamino) guanosine, kynurenine, 3,3',5-triiodo-L-thyronine, D-sorbitol, DL-P-hydroxyphenyllactic acid, phenyllactate, cyclic AMP, and S-(5-Adenosy)-L-homocysteine. Besides, among the 21 metabolites associated with MACE risk, 11 endogenous metabolites were

identified as independent metabolic signatures, including 4-acetamidobutyric acid, L-cystine, L-tryptophan, dulcitol, 5-methyluridine, kynurenine, phenyllactate, lysoPC 20:2, D-sorbitol, lysoPC 20:1, N6-succinyl adenosine.

Then, we developed prognostic models for predicting risks of clinical endpoints in the discovery cohort, using a multivariate Cox regression model based on minimal AIC. For the risk of death, a clinical model was built based on 17 clinical risk factors, and eight factors were retained in the final model. The trimethylamine oxide (TMAO) model consisted of TMAO and five clinical variables. The metabolomic model was built only based on the metabolites screened after lasso Cox regression and finally included five metabolites (3,3',5-Triiodo-L-thyronine,  $\beta$ -pseudouridine, Cyclic AMP, dulcitol, and kynurenine), see Additional file 1: Table S5. The metabolomic plus clinical model combining variables of the metabolomic and clinical model, in the final included four metabolites (3,3',5-Triiodo-L-thyronine,  $\beta$ -pseudouridine, dulcitol, and kynurenine) and five clinical variables. The metabolomic model yielded higher predictive efficiency than the clinical model (AUC, 80.9% vs. 77.0%). Moreover, the model combining metabolomics and the clinical model for death yielded an improved predictive efficiency than the TMAO and the clinical model (Table 3), with increased AUC (83.7% vs. 76.6% vs. 77.0%, Fig. 2A), positive IDI of 0.096, continuous NRI of 0.230 and 0.121. In addition, a metabolomic model for MACE contains six metabolites (Additional file 1: Table S5) that performed well than the clinical model (AUC, 66.0 vs. 58.4%). Furthermore, adding four metabolomic variables (kynurenine, lysoPC 20:2, 5-methyluridine, and L-tryptophan) to the clinical model was also with better predictive value than that of the TMAO and the clinical model (Table 3), with increased AUC (67.4% vs. 59.8% vs. 58.4%, Fig. 2B), positive IDI of 0.068 and 0.072, and continuous NRI of 0.144 and 0.106.

Subsequently, the optimized prediction model was used to estimate survival probabilities from death and MACE of each patient in the multicenter validation cohort. Survival curves showed that the model combining multiple metabolites with clinical factors could successfully differentiate patients with a low, middle, and high risk of death (log-rank test,  $P=0.0014$ ; Fig. 2C) and MACE ( $P=0.0008$ , Fig. 2D).

### Metabolomic associations with LV remodeling

Network analysis of metabolites associated with risks of death (Fig. 3A, Additional file 1: Table S6) or MACE (Additional file 1: Table S7, Additional file 2: Fig. S2) not only showed the importance of hub metabolites of

**Table 3** Model performance measures for mortality and MACE risks in the discovery phase

Predictive model	AUC	IDI (95% CI)	Continuous NRI (95% CI)
Prediction of death			
Metabolomic + clinical <sup>a</sup>	83.7		
Metabolomic <sup>b</sup>	80.9	0.072 (− 0.067 to 0.238)	0.013 (− 0.259 to 0.263)
TMAO + clinical <sup>c</sup>	76.6	0.096 (0.031–0.235)	0.230 (− 0.032 to 0.446)
Clinical <sup>d</sup>	77.0	0.096 (0.012–0.231)	0.121 (− 0.127 to 0.369)
Prediction of MACE			
Metabolomic + clinical <sup>e</sup>	67.4		
Metabolomic <sup>f</sup>	66.0	0.066 (0.005–0.124)	0.097 (− 0.049 to 0.238)
TMAO + clinical <sup>g</sup>	59.8	0.068 (0.029–0.118)	0.144 (0.005–0.324)
Clinical <sup>h</sup>	58.4	0.072 (0.034–0.128)	0.106 (− 0.001–0.321)

The metabolic variables screened from LASSO, TMAO, and the 17 traditional clinical factors including age, sex, AST, eGFR, DM, HyperT, CHOL, HDLC, PPI, ACEI, BB, CCB, current smoking, family history of CVD, SYNTAX, SBP, and GLUC were input into multivariate Cox proportional hazards regression analysis to fit model, using a forward and backward stepwise process based on AIC (Akaike information criterion). The model with the smallest AIC value was considered the best and variables with  $P < 0.1$  were retained. IDI (integrated discrimination improvement) and continuous NRI (net reclassification improvement) were calculated by comparing the Metabolomic + clinical model with TMAO + clinical and Clinical model, and the Metabolomic model with Clinical model, 95% CIs were calculated by 1000 bootstrap resampling

<sup>a</sup> Metabolomic + clinical model = Dulcitol,  $\beta$ -Pseudouridine, 3,3',5-Triiodo-L-thyronine, Kynurenine, age, current smoking, GLUC, AST, SBP

<sup>b</sup> Metabolomic model = Dulcitol, Kynurenine, Cyclic AMP, 3,3',5-Triiodo-L-thyronine,  $\beta$ -Pseudouridine

<sup>c</sup> TMAO model = TMAO, age, AST, current smoking, SBP, GLUC

<sup>d</sup> Clinical model = age, AST, HDLC, CCB, current smoking, SYNTAX, SBP, GLUC

<sup>e</sup> Metabolomic + clinical model = lysoPC 20:2, 5-methyluridine, kynurenine, L-tryptophan, AST, DM, PPI, SYNTAX

<sup>f</sup> Metabolomic model = lysoPC 20:2, 5-methyluridine, kynurenine, L-tryptophan, D-sorbitol, phenyllactate

<sup>g</sup> TMAO + clinical model = TMAO, AST, DM, PPI, CCB, SYNTAX

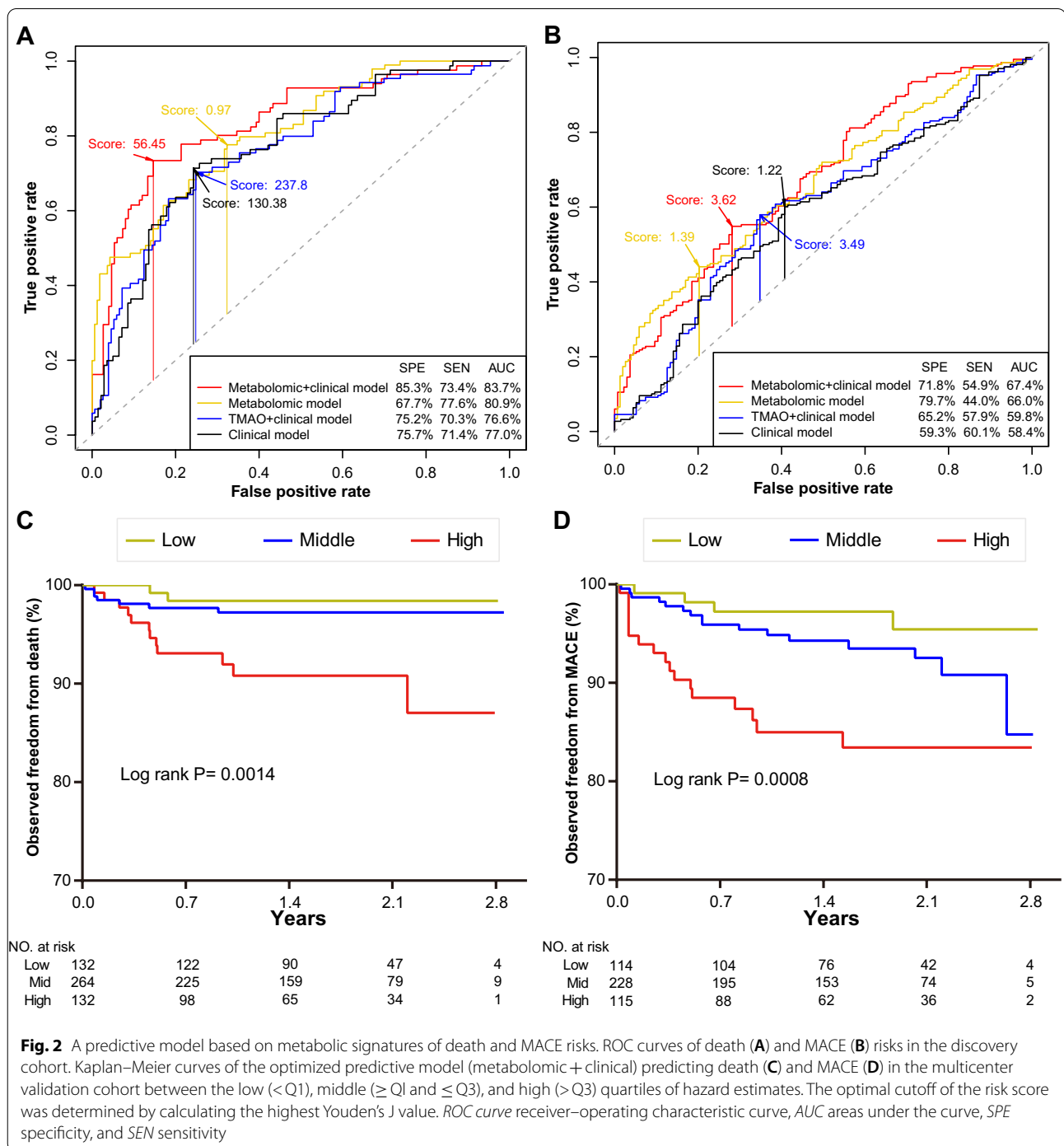
<sup>h</sup> Clinical model = AST, DM, PPI, CCB, SYNTAX. AUC = area under the curve, other abbreviations are as in Table 1

2-dimethylguanosine and kynurenine, but also important roles in reduced LVEF and increased proBNP. The relationships between metabolites with LVEF (Fig. 3B) and LVMI (Fig. 3C) in the discovery cohort were revealed, 32 metabolites were associated with LVEF after adjustment ( $FDR < 0.05$ ), and 20 metabolites were validated during multicenter validation ( $P < 0.05$ , Additional file 1: Table S8). Besides, 16 metabolites were associated with LVMI and eight metabolites were validated ( $P < 0.05$ , Additional file 1: Table S9). Interestingly, the hub metabolites, 2-(dimethylamino) guanosine was the most significant metabolite associated with reduced LVEF and increased LVMI. Another modified nucleoside,  $\beta$ -pseudouridine also showed a similar finding. In particular, 15 metabolites were not only associated with increased death risks but also reduced LVEF, 10 metabolites were associated with increased MACE risks and reduced LVEF (Fig. 4A). Moreover, five metabolites were shared by increased LVMI and death risks, three metabolites were shared by increased LVMI and MACE risks. For example, 2-(dimethylamino) guanosine (estimate [SE], − 4.75 [0.69];  $FDR = 4.72E-10$ ),  $\beta$ -pseudouridine (estimate [SE], − 2.80 [0.56];  $FDR = 9.56E-6$ ), kynurenine (estimate [SE], − 3.15 [0.68];  $FDR = 2.82E-05$ ), N6-succinyl adenosine (estimate [SE], − 2.58 [0.64],  $FDR = 1.95E-04$ ), 4-acetamidobutyric acid (estimate

[SE], − 1.79 [0.46];  $FDR = 3.55E-04$ ), dulcitol (estimate [SE], − 1.67 [0.46];  $FDR = 7.49E-04$ ), kynurenic acid (estimate [SE], − 2.16 [0.61];  $FDR = 9.15E-04$ ), phenylacetyl-L-glutamine (estimate [SE], − 1.86 [0.58];  $FDR = 2.62E-03$ ), adipic acid (estimate [SE], − 2.30 [0.74];  $FDR = 3.35E-03$ ), DL-P-hydroxyphenyllactic acid (estimate [SE], − 1.41 [0.58];  $FDR = 1.86E-02$ ), and cyclic AMP (estimate [SE], − 1.36 [0.68];  $FDR = 4.82E-02$ ) were validated to be associated with reduced LVEF. Besides, 2-(dimethylamino) guanosine (estimate [SE], 7.89 [2.46];  $P = 1.38E-03$ ), cyclic AMP (estimate [SE], 5.27 [2.35];  $P = 2.53E-02$ ), kynurenine (estimate [SE], 4.88 [2.35];  $P = 3.83E-02$ ) and phenylacetyl-L-glutamine (estimate [SE], 4.20 [2.05];  $P = 4.12E-02$ ) was confirmed to be related to increased LVMI.

#### Causality and mediation effects inference between metabolites, LV remodeling and clinical outcomes by MR analyses and mediation model

As secondary analyses, we performed MR analyses using genetic variants as the instrumental variables, to rule out the influences from confounders, thus providing inference of causality between the important metabolic signatures and clinical outcomes as well as LV remodeling (Fig. 4B and Additional file 1: Table S10). The MR analyses

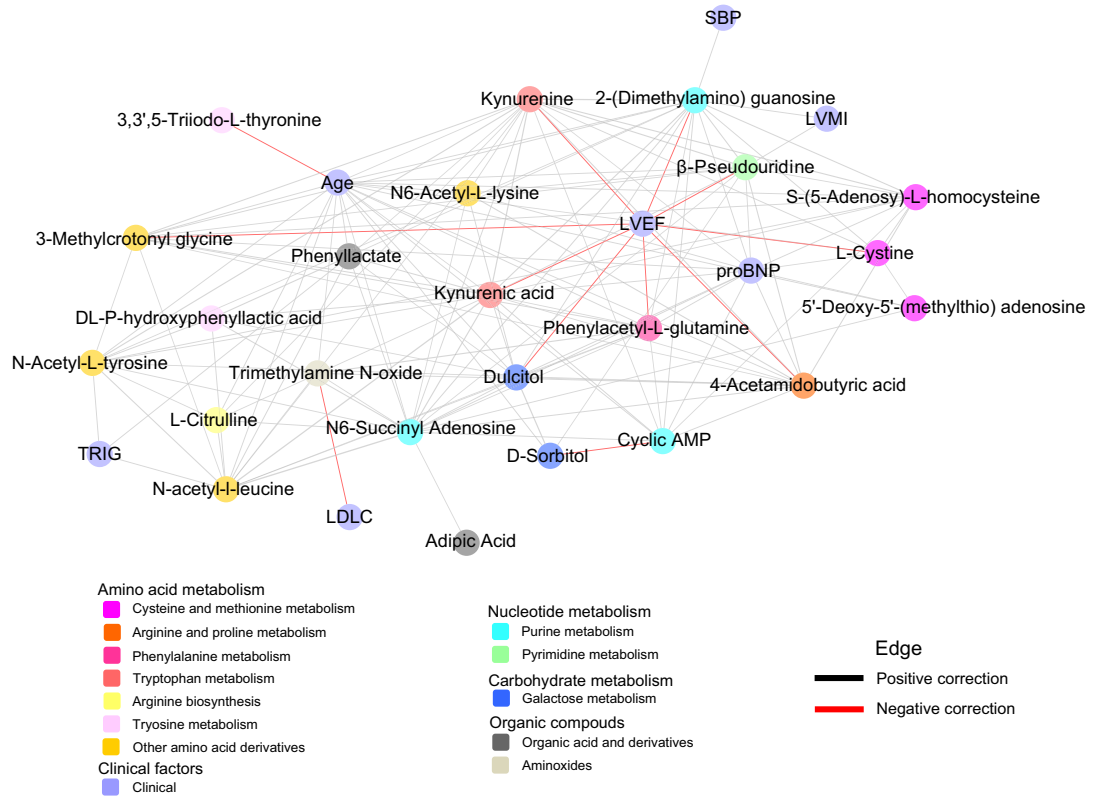


identified that N6-Succinyl adenosine (HR, 1.88 and 1.51), phenyllactate (HR, 2.71 and 1.79), and kynurenine (HR, 4.40 and 2.19) were associated with increased death and MACE risks ( $P < 0.05$ ). DL-P-hydroxyphenyllactic acid with HR of 3.11, adipic acid (HR = 4.67), S-(5-Adenosyl)-L-homocysteine (HR = 2.88), trimethylamine N-oxide (HR = 1.76), 3,3',5'-Triiodo-L-thyronine (HR = 0.11), and

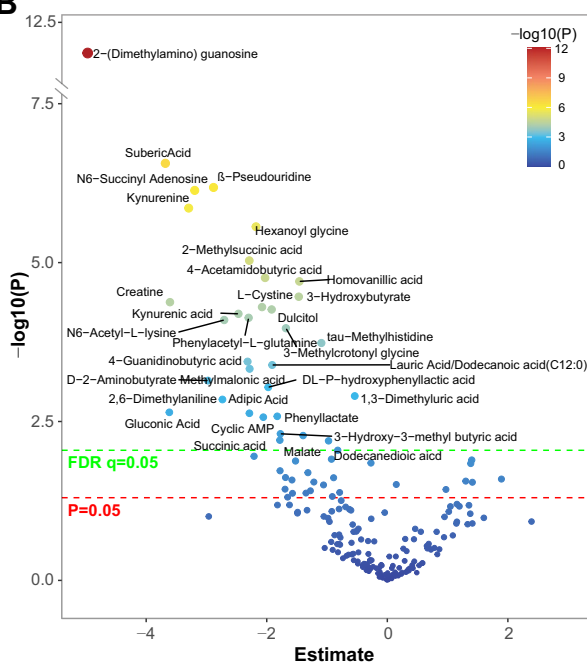
4-acetamidobutyric acid (HR = 2.34), were linked to the risks of death. D-Sorbitol with HR of 1.76 and phenylacetyl-L-glutamine with HR of 1.74 were linked to increased risk of MACE. Moreover, we still observed 14 metabolites, including 4-acetamidobutyric acid (estimate [SE], - 4.19 [1.18]), phenylacetyl-L-glutamine (estimate



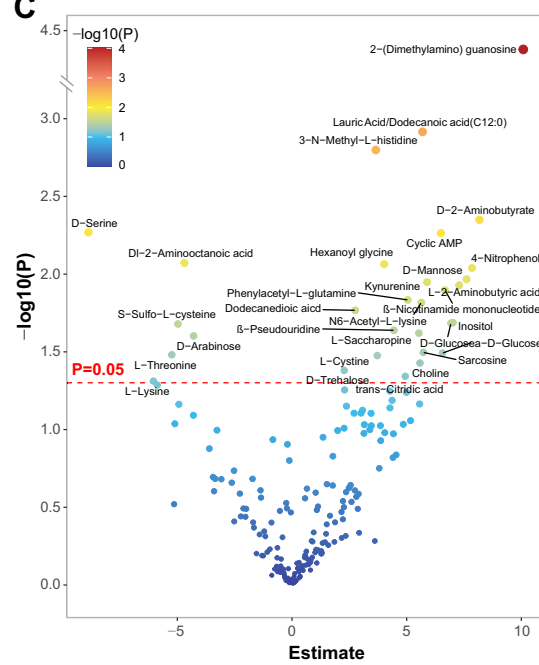
**A**



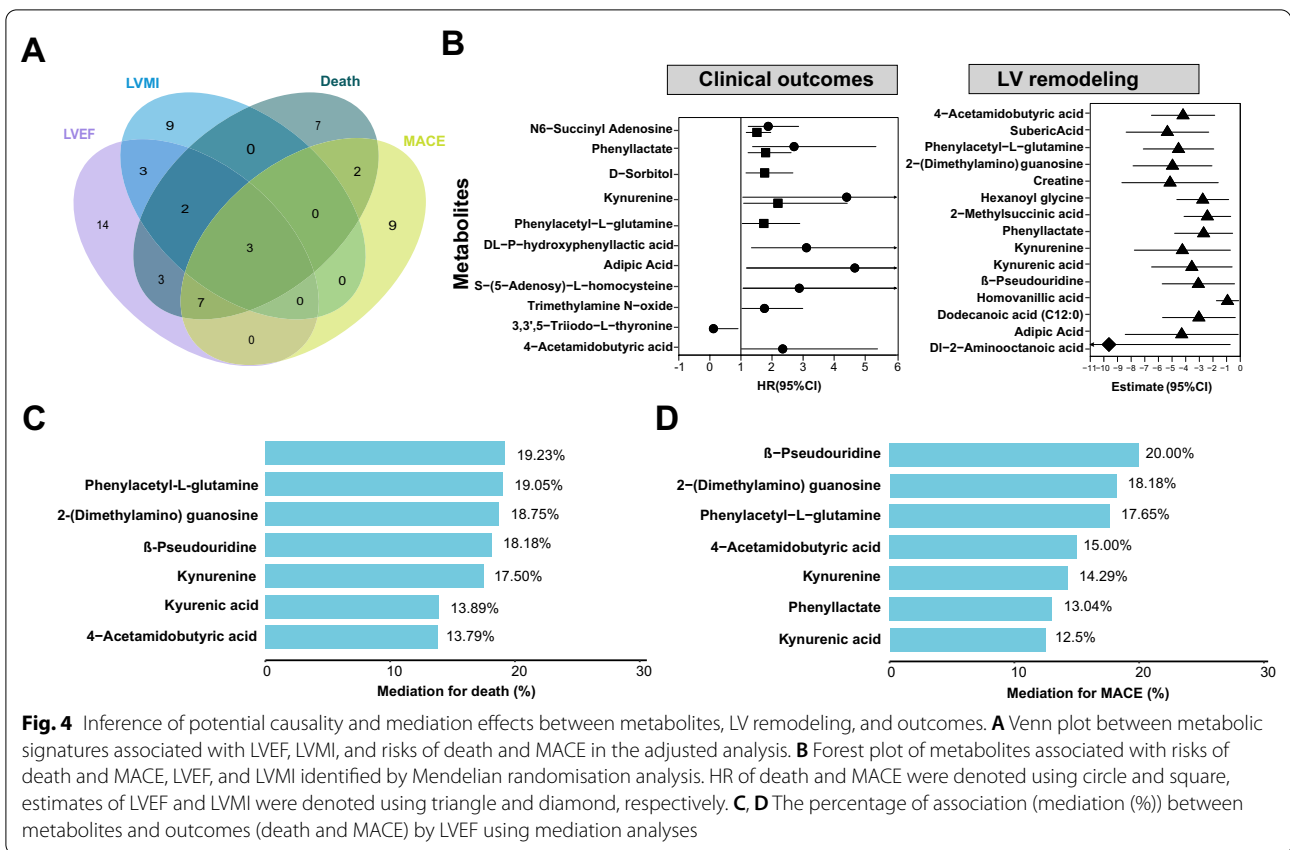
**B**



**C**



**Fig. 3** Correlation network of metabolic signatures and clinical factors, and metabolomic association with LV remodeling. **A** Correlation network of the metabolic signatures for death risk in the discovery cohort and traditional clinical factors by Spearman correlations with distant nodes of  $|\rho| > 0.1$  for clinical factors and  $|\rho| > 0.2$  for metabolites,  $P < 0.01$ , the rho and  $P$ -value were provided in Additional file 1: Table S6. Volcano plot presentation of univariate association of metabolites with LVEF (**B**) and LVM1 (**C**) in a linear regression model. Metabolites with  $P < 0.05$  are labeled (above the red dotted line),  $P$ -value cutoff equivalent to  $FDR < 0.05$  are labeled (above the green dotted line), abbreviations are as Table 1



**Fig. 4** Inference of potential causality and mediation effects between metabolites, LV remodeling, and outcomes. **A** Venn plot between metabolic signatures associated with LVEF, LVMI, and risks of death and MACE in the adjusted analysis. **B** Forest plot of metabolites associated with risks of death and MACE, LVEF, and LVMI identified by Mendelian randomisation analysis. HR of death and MACE were denoted using circle and square, estimates of LVEF and LVMI were denoted using triangle and diamond, respectively. **C, D** The percentage of association (mediation (%)) between metabolites and outcomes (death and MACE) by LVEF using mediation analyses

[SE], -4.53 [1.32]), phenylactate (estimate [SE], -2.68 [1.09]), kynurenine (estimate [SE], -4.24 [1.80]), adipic acid (estimate [SE], -4.29 [2.11]), kynurenic acid (estimate [SE], -3.55 [1.51]), 2-(dimethylamino) guanosine (estimate [SE], -4.97 [1.48]), β-pseudouridine (estimate [SE], -3.07 [1.35]) were associated with decreased LVEF, and DI-2-aminoctanoic acid (estimate [SE], -9.63 [4.53]) were associated with decreased LVMI.

Moreover, mediation analysis of LVEF was performed to further explain the associations between the metabolites and the clinical endpoints. Here, only those metabolites that were observed to have causal effects on LVEF in MR analysis were focused on. Mediation analysis proposed that the seven metabolites increase death and MACE risks could partly through impair LVEF (*P* value of Total effect and Med-eff < 0.05, Fig. 4C and Additional file 1: Table S11). The percentage of the total effect mediated by decreased LVEF on both death and MACE of β-pseudouridine was estimated at 18–20%, 2-(dimethylamino) guanosine and phenylacetyl-L-glutamine was at 18–19%, phenylactate was at 19% (death) and 13% (MACE), kynurenine was at 18% on death and 14% on MACE, 4-acetamidobutyric acid and kynurenic acid were at 13–15%.

## Discussion

In this prospective cohort of Chinese CAD patients, we described a comprehensive metabolomic study to identify endogenous metabolites that were associated with the major clinical outcomes of CAD, i.e. death and MACE. Based on these findings, we built prognostic models and found that combining metabolic signatures and clinical factors outperformed other models that used solely clinical risk factors, or clinical factors with established biomarker TMAO. Subsequently, we performed MR analysis to assess the potential causal effects of the metabolites on clinical outcomes. We observed that 11 metabolites were causally associated with the risk of death or MACE, including kynurenine, N6-Succinyl adenosine, phenylactate, DL-P-hydroxyphenylactic acid, 3,3',5-Triiodo-L-thyronine (T3), adipic acid, S-(5-Adenosy)-L-homocysteine, TMAO, 4-acetamidobutyric acid, D-sorbitol, and phenylacetyl-L-glutamine. In addition, we proposed for the first time that 7 metabolic signatures inferred as causal on the reduced LVEF by MR analysis increased the risks of death and MACE possibly partly through impairing LV function, including

kynurenine, phenylacetyl-L-glutamine, 4-acetamidobutyric acid, phenyllactate, kynurenic acid, and two modified nucleosides. To our knowledge, our study provided the most comprehensive metabolomic and causal inference for the major clinical outcomes of CAD to date, suggesting potential pathological mechanisms, and offering biomarkers and potential therapeutic targets for the secondary prevention of CAD.

Deciphering the details of the circulating metabolome could help to predict the risk of CAD and subsequent cardiovascular events [24]. To date, many studies have already identified some biomarkers for CAD risk, such as sphingolipids [25] and TMAO [26]. Moreover, prior studies identified several metabolic markers for all-cause mortality in CAD patients of European ancestry [14, 15]. However, the metabolomic biomarkers have not been well studied for clinical outcomes of CAD in Asian population especially Chinese at present [27]. In our study on Chinese CAD patients, we replicated several of them, including kynurenine, T3, S-(5-adenosyl)-L-homocysteine, TMAO, and phenylacetyl-L-glutamine. Interestingly, the impacts of these metabolites on the death or MACE risk were further confirmed by our MR analysis. Their potential biological mechanisms in CAD were discussed as follows. Increased catabolism of tryptophan to kynurenine was related to cardiovascular events and mortality [28, 29], and activation of tryptophan–kynurenine pathway was correlated with LV dysfunction [30]. Consistent with this, our mediation analysis showed that 13–18% of adverse effects of kynurenine and its metabolites kynurenic acid on death and MACE risks could be explained by reduced LVEF. Kynurenine and several of its downstream can bind to transcription factor aryl hydrocarbon receptor and mediate T cell apoptosis and vascular inflammation, which in turn regulates inflammation-induced CVD [31]. Therefore, dysregulation of tryptophan metabolism represents a functional mechanism and generates biomarkers for early prognosis and therapeutic intervention. Low T3 level was associated with an increased risk of death, indicating an important role of thyroid function in CAD, which was consistent with a previous meta-analysis that low serum T3 level was related to increased risk of all-cause mortality and MACE [32]. S-(5-adenosyl)-L-homocysteine was reported as a strong predictor of mortality in cirrhosis [33], and as a more sensitive biomarker than homocysteine for cardiovascular events [34, 35].

TMAO and phenylacetyl-L-glutamine were two well-known gut microbiota-derived metabolites. TMAO is an established biomarker for cardiovascular events and mortality [36–39]. TMAO impacts multiple aspects of ‘patient vulnerability’, including atherosclerotic plaque development, platelet hyperresponsiveness [37, 40], and

variation in macrophage and endothelial cell phenotype [41]. More recently, phenylacetyl-L-glutamine began to attract attention as a novel marker for CAD [42] and other cardiovascular diseases [43]. It was derived from gut microbiota, and via adrenergic receptors, it could enhance platelet responsiveness and athero-thrombosis, resulting in cardiovascular events. Here we reported for the first time that phenylacetyl-L-glutamine increased the risks of death and MACE in CAD could be explained at 18–19% by reduced LVEF, indicating a novel mechanistic link to cardiovascular events. The phenylacetyl-L-glutamine is regarded as a predictor of the event of heart failure [44], the accumulation of phenylacetylglutamine in heart failure may be due to increased amino acid degradation. Taken together, these findings support the notion that gut microbiota-derived metabolites are closely associated with cardiovascular health, and provide clues for developing novel therapeutic strategies in the cardiovascular field.

In addition to the known metabolic markers, we discovered additional markers that were not reported in previous CAD studies yet were revealed through our comprehensive study design of recruiting patients from multiple clinical centers. These novel markers included modified nucleosides, aromatic lactic acids, 4-acetamidobutyric acid, etc. The finding that increased plasma levels of modified nucleosides in CAD patients at high risk of major clinical outcomes is intriguing. Two modified nucleosides, 2-(dimethylamino) guanosine and  $\beta$ -pseudouridine, were found in our study as robust metabolic signatures for high risks of death and MACE, despite no causal relationship being observed in the MR analysis. Importantly, they displayed the strongest correlation with LV remodeling and were causally associated with reduced LVEF in MR. Furthermore, mediation analysis also inferred that their associations with clinical outcomes could be explained about 18–20% through impaired LVEF. It was consistent with recent studies that increased levels of 2-(dimethylamino) guanosine and  $\beta$ -pseudouridine levels were related to incident heart failure and LV remodeling [45, 46]. Furthermore, 2-(dimethylamino) guanosine has been linked to increased risk of coronary artery calcium [47] and all-cause mortality in diabetes [48], and elevated circulating levels of  $\beta$ -pseudouridine were associated with atrial fibrillation [49], heart failure [50] and CAD [51]. Here we report for the first time that  $\beta$ -pseudouridine was also a metabolic marker for the risks of death and MACE in CAD. Moreover, increased circulating modified nucleosides were also associated with the development of pulmonary hypertension, reflecting elevated stress and increased proliferation of pulmonary vascular cell [52]. Modified nucleosides consist of various RNA species, reflecting

the upregulation of translation and hypercatabolism in general [53]. Taken together, our findings suggested that CAD patients may experience disorders of modified nucleoside, causing excessive damaging stress in cardiomyocytes, which in turn promoted cardiac remodeling and major clinical outcomes.

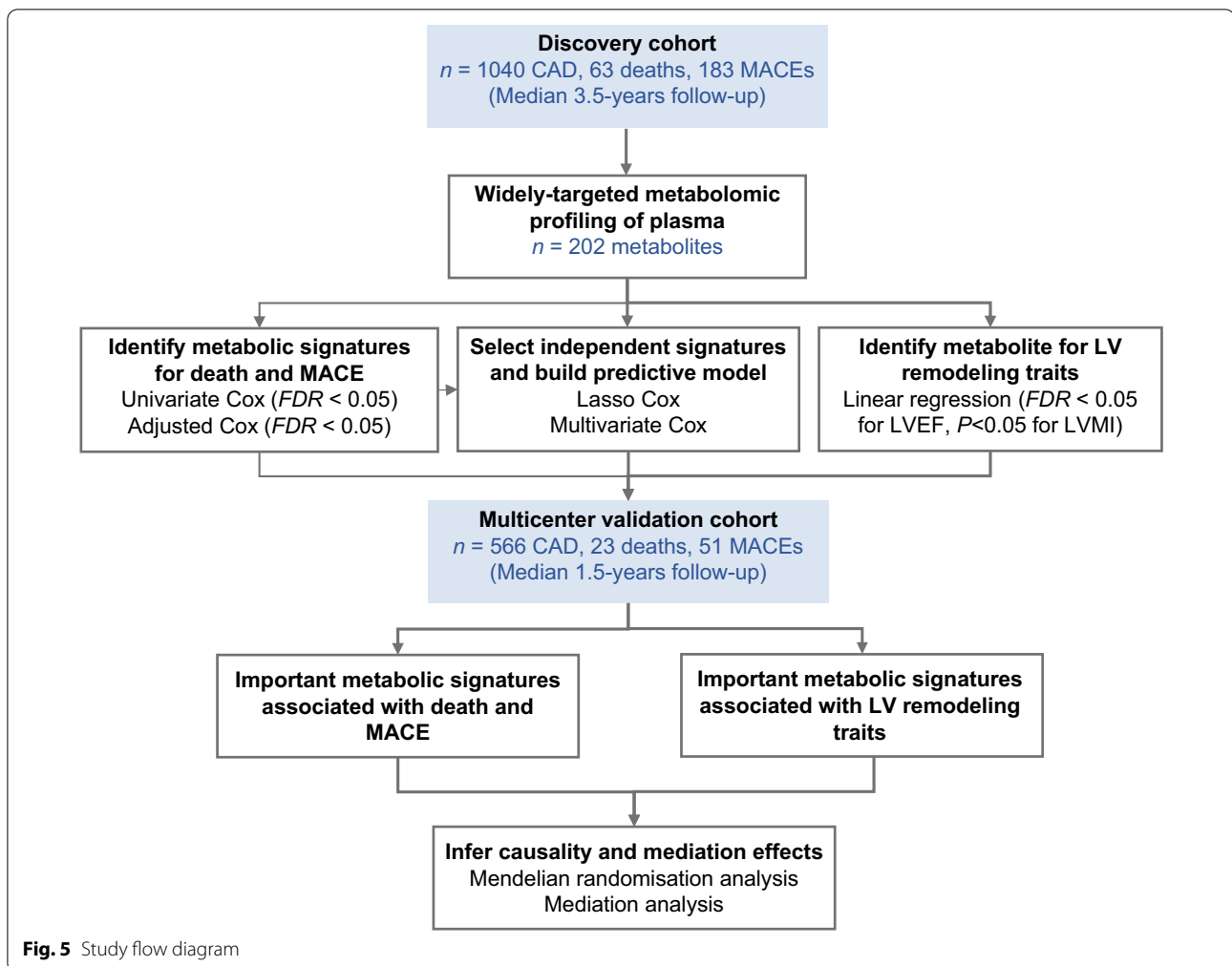
4-Acetamidobutyric acid appeared in our study as the metabolite most strongly associated with both risks of death and MACE, and about 14–15% of these effects could be mediated by promoting LV malfunction. Recent studies also found that 4-acetamidobutyric acid was associated with incident heart failure and LVMI [46], rheumatoid arthritis [54], and severity of liver and kidney disease in cirrhosis [33]. 4-Acetamidobutyric acid is a product of polyamine, arginine and proline metabolism. Polyamines play important roles in cell proliferation, apoptosis, and tumor growth, thus serving as a target for cancer prevention and treatment [55]. Polyamine stress response could be activated by acute ischemia, leading to polyamine accumulation and cardiac cell death [56, 57]. The primary pathophysiological mechanisms of polyamines during ischemia/reperfusion injury and cardiac failure were in regulating cardiomyocyte death. Moreover, inhibition of polyamine biosynthesis could protect cardiac cells from norepinephrine-mediated apoptosis [58]. Therefore, we speculated that 4-acetamidobutyric acid is a key polyamine underlying the progression of heart failure and CAD, which warrants further investigation.

In addition, it is interesting to note that two elevated aromatic lactic acids, phenyllactate, and DL-P-hydroxyphenyllactic acid, were inferred being causally associated with the adverse outcomes in CAD. They have been reported to be produced by the human gut microbiome [59] and were positively associated with hepatic steatosis [60], cirrhosis [33], and hepatocellular carcinoma development [61]. Recent studies reported that 3-(4-hydroxyphenyl) lactate significantly increased the risk of diabetes [62], and in rodents, hydroxyphenyllactate may decrease the production of reactive oxygen species in mitochondria and neutrophils [59]. While their biological linkage with CAD and heart failure needs further study. *N*6-succinyl adenosine, a purine metabolite, was reported to accumulate in the fluid of patients with adenylosuccinase deficiency, causing severe neurological impairment [63]. *N*6-succinyladenosine was reported to promote inflammasome activity and IL-1 $\beta$  production in subjects with high expression of the inflammasome module, which was related to all-cause mortality [64]. Moreover, *N*6-succinyl adenosine was dramatically increased following myocardial infarction [65] in rats and significantly elevated in patients with chronic thromboembolic

pulmonary hypertension [66], providing clues to the association with CAD. Adipic acid is a product of lipid oxidation and could predict the development of islet autoantibodies [67]. A study reported that accumulated adipic acid as toxic metabolites from medium-chain acyl-CoA dehydrogenase deficiency-induced significant DNA damage in vitro [68].

Understanding the effects of these metabolic markers on the clinical events to be causal or merely associated could shed light on potential novel intervention targets for CAD. Therefore, unique in this study, we have explored whether there are potential causal effects of metabolic markers on the prognosis and LV remodeling of CAD using MR analysis. MR analysis is a popular causality inference tool that uses the inherent genotype as the instrumental variable which follows Mendelian laws of inheritance [65]. Genotypes as instrumental variables are not affected by confounding factors such as diet and other environmental factors. Hence, MR analysis could help us to rule out the influences from confounders and have causal timing, thus providing unbiased estimates of causality between exposure and outcome. However, although MR analysis and mediation analysis provided important causal inference for our findings, those newly identified metabolic markers, are still worth-well to further causality confirmation and mechanism study.

In addition, a recent study revealed that heterogeneous metabolic deviation profiles were evident even in a homogenous subgroup of acute coronary syndrome patients, which emphasizes that personalized risk stratification and preventative measures are essential in CAD [69]. Our study confirmed that combining metabolic signatures and clinical risk factors yields a better prognostic prediction for the clinical endpoints in CAD patients, compared to using metabolic signatures or clinical risk factors alone. Furthermore, it outperformed the model of the established marker TMAO plus clinical risk factors, even though the plasma level of TMAO was previously suggested as a predictor of death and MACE in CAD [17, 36]. It was also reported that the model including TMAO, for predicting clinical endpoints resulted in limited C-statistics and no significant improvement compared with the traditional risk factors [36]. The progression of CAD is multifactorial with complex physiopathology, which concurrently alters multiple metabolic pathways. Therefore, a joint metabolic signature has better strength to capture the complex changes, leading to improved tools for early risk stratification for the secondary prevention of CAD. Future investigations will help evaluate and improve the clinical utility of our model for different patient populations.



### Limitations

First, the study was based on Chinese patients, and the sample size of the validation cohort was relatively small, therefore we only validated the results with strong signals and could miss the weak signals that were weak. Studies expanding to other ethnicities and large-scale cohorts were also needed. Secondly, due to the metabolomic platform and method updating, the conditions for metabolite detection between the two cohorts were different, leading to 600 metabolites being detected in the multicenter validation cohort, while only 160 metabolites were identical to the discovery cohort. If more metabolites were added, then more novel metabolic signatures could be found. Another limitation is that the effect of CAD severity on the clinical outcomes was controlled by adjusting the SYNTAX scores in the discovery cohort, but the multicenter cohorts did not eliminate this bias since the data was unavailable. Finally, even though this study is prospective research based on observational data and explored the potential causal relationship between

the key metabolites with LV remodeling traits and the outcomes, further mechanistic experiments are needed to verify the biological linkage and whether the biomarkers we identified precede LV remodeling incidence or concomitant result. Future studies using improved metabolomic detection methods can leverage the analysis framework we described in this study for discovering more novel metabolic signatures and better prognostic models.

### Conclusions

This study provided new insights into the relationship between plasma metabolites and clinical outcomes and LV remodeling in CAD. The prediction model based on the metabolic signatures and clinical risk factors can significantly improve the risk stratification for death and MACE, serving as a potential prognostic tool. Several metabolic markers such as kynurenines, phenylacetyl-L-glutamine, 4-acetamidobutyric acid, and modified nucleosides may causally increase the risk of death and

MACE, and these impacts appear to be partly mediated by impaired cardiac function, which still merits further study. These findings suggest important risk markers and potential therapeutic targets for the secondary prevention of CAD.

## Methods

### Study cohort

The study flow is depicted in Fig. 5. A total of 1606 Chinese patients with CAD were enrolled, including a discovery cohort and a multicenter validation cohort. Patients of the discovery cohort were sequentially enrolled from Guangdong Provincial People's Hospital between January 2010 and December 2013 and followed up for the primary endpoint (all-cause death) and secondary endpoint (MACE) from June 2010 through April 2017 for a median of 3.5 years, with 63 deaths and 183 MACEs. Patients in the multicenter validation cohort were enrolled from Guangdong Provincial People's Hospital ( $n=354$ ), Xiangya Hospital of Central South University ( $n=178$ ), and First Affiliated Hospital of Sun Yat-sen University ( $n=34$ ) from September 2017 to October 2018 and followed up until August 2020 for a median of 1.5 years, with 23 deaths and 51 MACEs. Most patients of this study were also included in our previous published lipidomic study [70].

Patients who were diagnosed with CAD (50% stenosis on coronary angiography) or received the percutaneous coronary intervention (PCI) were included. The exclusion criteria for patients in the discovery cohort included the following: (1) age < 18 years or > 80 years, (2) renal insufficiency (defined as serum creatinine concentration > 2 times the upper limit of normal [230  $\mu\text{mol/L}$ ], history of renal transplantation or dialysis), (3) hepatic insufficiency (defined as serum transaminase concentration > 2 times the upper limit of normal [80 U/L], or a diagnosis of cirrhosis), (4) being pregnant or lactating, (5) advanced cancer or haemodialysis, (6) history of thyroid problems and taking antithyroid drugs or thyroid hormone medication, (7) incomplete information about cardiovascular events during follow-up. The exclusion criteria for patients in the multicenter validation cohort was the seventh criterion above. Baseline informations, including demographics, medical history, biochemical measurements, and medication, were obtained from the hospital information database.

The primary outcome was death and the second was MACE (including death, nonfatal myocardial infarctions, coronary revascularization, and cerebral infarction). Coronary angiography was performed, and the Synergy between PCI with TAXUS and Cardiac Surgery (SYNTAX) score was calculated to assess the severity of CAD.

The echocardiography was conducted to determine left ventricular ejection fraction (LVEF) and left ventricular mass index (LVMI). For detailed information see Additional file 2: Methods.

### Metabolite quantification

Widely targeted metabolomic profiling was conducted in the discovery cohort (UPLC, Shim-pack UFLC SHIMADZU CBM30A; MS, Applied Biosystems 4500 QTRAP) and the multicenter validation cohort (UPLC, Shim-pack UFLC SHIMADZU CBM30A; MS, Applied Biosystems 6500+ QTRAP) by Wuhan Metware Biotechnology, detailed information sees Additional file 2: Methods and Fig. S3. The ionisation modes and ion pairs of all metabolites see Additional file 1: Table S12. Due to the metabolomic platform and method updating, the conditions for metabolite detection between the two cohorts were different, mainly including changed MS system, ESI source operation parameters, flow rate, and injection volume. So, this led to 600 metabolites being detected in the multicenter validation cohort, while only 202 metabolites were detected in the discovery cohort, with 160 identical metabolites.

### Statistical analysis

The analytical method of this study was similar to our previous published lipidomic research [70]. Firstly, data were presented using the number (percent) for categorical variables and mean  $\pm$  standard deviation (SD) for continuous variables. For metabolomic analysis, raw signals with a coefficient of variation (CV) > 50% in the quality control (QC) samples were removed, and missing values below the limit of detection were assigned to the minimum detection level for the metabolites. The CV of all metabolites see Additional file 1: Table S13. The Quality Control–Robust Loess Signal Correction algorithm was used for correction and integration to reduce the bias from a batch effect [71]. QC-RLSC is an effective way to normalize the metabolite features of the QC samples within an analytical block. Each block of the metabolomic data was scaled using Pareto scaling [72]. Detailed information to reflect our data improvement after batch correction see Additional file 2: Figs. S4 and S5.

Univariate and adjusted Cox regression analysis was performed to identify the clinical characteristics and metabolites that were associated with clinical endpoints and to estimate the hazard ratio (HR) and 95% confidence intervals (CI). Univariate and adjusted Linear regression was used to reveal the clinical characteristics and metabolites that were related to LVEF and LVMI with results presented as estimates  $\pm$  standard errors (SE). In the adjusted analysis, covariates included 17 potential

confounders, including age, sex, aspartate aminotransferase, estimated glomerular filtration rate, diabetes, hypertension, systolic blood pressure, glucose, cholesterol, high-density lipoprotein cholesterol, proton pump inhibitors, angiotensin-converting enzyme inhibitors,  $\beta$ -blockers, calcium channel blockers, current smoking, family history of cardiovascular disease (CVD) and SYN-TAX. A two-tailed  $P$  of 0.05 was considered for statistical significance, and Benjamini–Hochberg method was used to control the false discovery rate ( $FDR$ ) for correcting the number of metabolites in multiple hypothesis testing.

Time-to-event free survival was shown using Kaplan–Meier curves, and the  $P$ -values were analyzed using the log-rank test. The Spearman correlations in the metabolites associated with death and MACE risks in the adjusted analysis, together with traditional clinical factors (age, pro-brain natriuretic peptide, blood lipids, LVEF, LVMI, SYNTAX, glucose, blood pressure), correlations above the significant level of  $P < 0.01$  and  $|\rho| > 0.1$  for clinical factors and  $|\rho| > 0.2$  for metabolites, were visualized using Cytoscape (version 3.7) to construct a correlation network. The metabolites with  $FDR < 0.05$  in the adjusted Cox regression analysis were used for least absolute shrinkage and selection operator (lasso) Cox regression analysis (“glmnet” package) to reduce the dimensions of variables and select the powerful metabolic features to build a multivariable prognostic model. This procedure was performed within a tenfold cross-validation framework (200 repeats), and the variables with repeats zero were removed and the metabolites left were selected to establish the prognostic model using multivariate Cox regression, based on the Akaike Information Criterion (AIC) with a forward and backward stepwise process.

To evaluate the prediction efficiency of the multivariate model, hazard estimates of individuals were calculated using the following formula:  $h(X) = \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i)$ , where  $\beta$  is the regression coefficient, and  $X_i$  is the selected marker. Time-dependent receiver-operating characteristic (ROC) analysis (by ‘timeROC’ package in R) was used to assess the prognostic abilities of the hazard models. The model performance was assessed by calculating the area under the curve (AUC), continuous net reclassification indices (continuous NRI by ‘survIDINRI’ package in R), and integrated discrimination improvement (IDI by ‘survIDINRI’ package in R). The sensitivity and specificity of each model were calculated, and the corresponding optimal cut off values (Youden’s index) were determined according to the formula: sensitivity + specificity – 1 [73]. The multicenter validation of the predictive model was performed in another multicenter cohort based on the individually

hazard estimates calculation, and the hazard stratification amongst the low ( $< Q1$ ), middle ( $\geq Q1$  and  $\leq Q3$ ), and high ( $> Q3$ ) hazard estimate groups were shown using Kaplan–Meier curves.

Statistical analyses were performed using GraphPad Prism 7 and R (version 4.1.0, <http://www.R-project.org/>).

### Mendelian randomisation analysis

In secondary analyses, we performed a one-sample MR analysis to infer the potential causal relationships between the metabolites and the outcomes of CAD as well as indicators of LV remodeling (LVEF and LVMI). A schematic diagram sees Additional file 2: Fig. S6.

Association results between genotypes and the metabolites were obtained from our previous metabolome-based genome-wide association study [74], and the results used in this study were provided in Additional file 1: Table S14. The association analysis between genotypes and LV remodeling was subjected to a linear model based on additive mode, adjusting for age, sex, aspartate aminotransferase, estimated glomerular filtration rate, anti-hypertensive drugs medication, hypertension, diabetes, and the first ten principal components. For each association analysis, we adopted a commonly used  $P < 1 \times 10^{-5}$  as a threshold [75, 76] to select the single nucleotide polymorphisms (SNPs) in the discovery cohort [74] for the following analyses. Linkage disequilibrium (LD) analysis was conducted to retain the SNPs with the lowest  $P$ -value and the independent SNPs with them ( $LD r^2 < 0.001$  in a 10,000 kb window or two SNPs beyond 10,000 kb) as the independent instruments. A two-stage least squares (2SLS) method was used in this MR analysis process and we calculated the 2SLS regression in R (version 3.6.3).

Briefly, in the first stage, we regressed by exposure ( $X$ ) to the instrumental variable (IV) to derive fitted values for exposure to IV ( $G-X$ ). In the second stage, we regressed the outcome ( $Y$ ) based on the fitted values from the first stage regression ( $X-Y$ ). The causal estimated effect size is the regression coefficient for this second stage that reflects the change in outcome due to a unit change in exposure. In the case of multiple IVs, the 2SLS estimator can be viewed as a weighted average of the ratio estimates calculated from a single instrumental variable, with weights determined by the relative strength of the instrumental variable in the first stage regression. The outcome was regressed over the predicted values of exposure by using Cox regression to estimate HR and 95% CI for risks of death and MACE, using Linear regression to calculate estimates and SE for LVEF and LVMI,  $P$ -value  $< 0.05$  was used to define causality.

## Mediation analysis

Mediation analysis decomposes the total exposure–outcome effect into a direct effect and an indirect effect through a mediator variable [77, 78]. The mediation analysis models [79] were constructed to examine whether the association of metabolic signatures and risks of death and MACE could be mediated through their role in promoting LV remodelling and estimate mediation effects size. A schematic diagram sees Additional file 2: Fig. S6. Metabolites were continuous predictor variables (X); LVEF were continuous mediators (M); death and MACE were dichotomous outcome variables (Y). In this study, we performed the mediation analyses in the following four steps: (1) regressing outcomes (death and MACE, respectively) on predictors (model 1  $Y = cX$ ), where  $c$  is total effect; (2) regressing mediators (LVEF) on predictors (model 2  $M = \beta_1 X$ ), where  $\beta_1$  is indirect effect 1; (3) regressing outcomes on mediators (model 3  $Y = \beta_2 M$ ), where  $\beta_2$  is indirect effect 2; (4) regressing outcomes (death and MACE, respectively) on predictors controlling for mediators (LVEF). The regression equation is Model  $Y = \beta_2 M + c'X$ , where  $\beta_2$  is indirect effect 2, and  $c'$  is direct effect; and calculating mediation effect as  $(\beta_1 \times \beta_2)/c$ . Mediation analyses were conducted using Lavaan package of R 4.0.5.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13578-022-00863-x>.

**Additional file 1: Table S1.** Baseline characteristics on death and MACE risks. **Table S2.** Baseline characteristics on LVEF and LVMI. **Table S3.** Relationship between metabolites and death risk. **Table S4.** Relationship between metabolites and MACE risk. **Table S5.** Multivariate Cox proportional hazards model for clinical outcomes. **Table S6.** Spearman correlation analysis between clinical factors and those metabolites associated with death risk. **Table S7.** Spearman correlation analysis between clinical factors and those metabolites associated with death risk. **Table S8.** Relationship between metabolites and LVEF. **Table S9.** Relationship between metabolites and LVMI. **Table S10.** Estimation of the potential causal relationships between the metabolites and outcomes and LV remodeling traits by MR analyses. **Table S11.** Mediation effect of metabolites–LVEF–outcomes association. **Table S12.** The ionisation modes and ion pairs of the metabolites. **Table S13.** The coefficient of deviation of the metabolites between QC samples. **Table S14.** Associations between SNP and metabolites in the metabolome-based genome-wide association study.

**Additional file 2.** Additional Methods. **Fig. S1.** The Kaplan–Meier curves of LVEF (A–B) and LVMI (C–D) for risks of death and MACE in the discovery cohort. **Fig. S2.** Correlation network of metabolic signatures for MACE risk and clinical factors. **Fig. S3.** Representative total ion flow diagrams between different QC samples under positive ion mode (A) and negative ion mode (B). **Fig. S4.** The visualization image analysis of the quality improvement procedures for typical features (Kynurenine) in metabolomics data. **Fig. S5.** Comparison of the cumulative frequency of RSD% of all features in QC samples before and after batch correction by QC-RLSC. **Fig. S6.** Diagram of Mendelian randomization analysis and mediation analysis.

## Acknowledgements

We would like to thank the participants in this study.

## Author contributions

SZ was the principal investigator of this study and designed the study. QZ and MQ performed the data analysis. QZ drafted the manuscript. ZW and YW assisted in statistical analysis and critically revised the manuscript. XC, CL, YL and HC were responsible for patient recruitment and clinical data collection. QM, WH, JC, YL, FL BZ, SZ, GH, HL, MZ, HZ, JC and MH assisted in data curation and revised the manuscript. All authors reviewed the final manuscript. All authors read and approved the final manuscript.

## Funding

This study was funded by the National Nature Science Foundation of China (No. 81872934, 81673514), the Key-Area Research and Development Program of Guangdong Province, China (No. 2019B020229003), the Science and Technology Planning Project of Guangdong Province, China (No. 2017B030314041).

## Availability of data and materials

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

## Declarations

### Ethics approval and consent to participate

This study was approved by the Medical Ethical Review Committee of Guangdong General Hospital (No. GDREC2010137H, GDREC2017071H) and conducted according to the Declaration of Helsinki. Informed consent (No. 20100910, 20170211) was obtained from all individual participants of the study, participants were identified by number.

### Consent for publication

Not applicable.

### Competing interests

The authors declared no conflict of interest.

### Author details

<sup>1</sup>Department of Pharmacy, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510080, Guangdong, China. <sup>2</sup>Guangdong Provincial Key Laboratory of Coronary Heart Disease Prevention, Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510080, Guangdong, China. <sup>3</sup>School of Medicine, South China University of Technology, Guangzhou 510080, Guangdong, China. <sup>4</sup>Department of Clinical Pharmacology, Xiangya Hospital, Central South University, Changsha 410008, Hunan, China. <sup>5</sup>Department of Cardiology, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou 510080, Guangdong, China. <sup>6</sup>Department of Cardiology, Xiangya Hospital, Central South University, Changsha 410008, Hunan, China. <sup>7</sup>Institute of Model Animal, Wuhan University, Wuhan 430072, Hubei, China. <sup>8</sup>Department of Pharmacy, Guangzhou Red Cross Hospital, Jinan University, Guangzhou 510220, Guangdong, China. <sup>9</sup>Wuhan Metware Biotechnology Co., Ltd., Wuhan 430000, Hubei, China. <sup>10</sup>Institute of Clinical Pharmacology, School of Pharmaceutical Sciences, Sun Yat-Sen University, Guangzhou 510006, Guangdong, China.

Received: 26 January 2022 Accepted: 28 July 2022

Published online: 14 October 2022

## References

1. Timmis A, Townsend N, Gale CP, Torbica A, Lettino M, Petersen SE, et al. European society of cardiology: cardiovascular disease statistics 2019. *Eur Heart J.* 2020;41(1):12–85.



2. Zhou M, Wang H, Zeng X, Yin P, Zhu J, Chen W, et al. Mortality, morbidity, and risk factors in China and its provinces, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2019;394(10204):1145–58.
3. Musunuru K, Kathiresan S. Genetics of common, complex coronary artery disease. *Cell*. 2019;177(1):132–45.
4. Sharpe N. Cardiac remodeling in coronary artery disease. *Am J Cardiol*. 2004;93(9):17–20.
5. Redfield MM, Jacobsen SJ, Burnett JC, Mahoney DW, Bailey KR, Rodeheffer RJJ. Burden of systolic and diastolic ventricular dysfunction in the community: appreciating the scope of the heart failure epidemic. *JAMA*. 2003;289(2):194–202.
6. Gaasch WH, Zile MR. Left ventricular structural remodeling in health and disease: with special emphasis on volume, mass, and geometry. *J Am Coll Cardiol*. 2011;58(17):1733–40.
7. Lindholm D, Lindback J, Armstrong PW, Budaj A, Cannon CP, Granger CB, et al. Biomarker-based risk model to predict cardiovascular mortality in patients with stable coronary disease. *J Am Coll Cardiol*. 2017;70(7):813–26.
8. Li XS, Obeid S, Klingenberg R, Gencer B, Mach F, Räber L, et al. Gut microbiota-dependent trimethylamine N-oxide in acute coronary syndromes: a prognostic marker for incident cardiovascular events beyond traditional risk factors. *Eur Heart J*. 2017;38(11):814–24.
9. Fatima T, Hashmi S, Iqbal A, Siddiqui AJ, Sami SA, Basir N, et al. Untargeted metabolomic analysis of coronary artery disease patients with diastolic dysfunction show disturbed oxidative pathway. *Metabolomics*. 2019;15(7):1–12.
10. Zhang L, Wei TT, Li Y, Li J, Fan Y, Huang FQ, et al. Functional metabolomics characterizes a key role for *N*-acetylneuraminic acid in coronary artery diseases. *Circulation*. 2018;137(13):1374–90.
11. Ussher JR, Elmariah S, Gerszten RE, Dyck JR. The emerging role of metabolomics in the diagnosis and prognosis of cardiovascular disease. *J Am Coll Cardiol*. 2016;68(25):2850–70.
12. Tremblay BL, Guénard F, Lamarche B, Pérusse L, Vohl MC. Familial resemblances in human plasma metabolites are attributable to both genetic and common environmental effects. *Nutr Res (New York, NY)*. 2019;61:22–30.
13. Johnson CH, Ivanisevic J, Siuzdak G. Metabolomics: beyond biomarkers and towards mechanisms. *Nat Rev Mol Cell Biol*. 2016;17(7):451.
14. Deelen J, Kettunen J, Fischer K, van der Spek A, Trompet S, Kastenmuller G, et al. A metabolomic profile of all-cause mortality risk identified in an observational study of 44,168 individuals. *Nat Commun*. 2019;10(1):3346.
15. Yu B, Heiss G, Alexander D, Grams ME, Boerwinkle E. Associations between the serum metabolome and all-cause mortality among African Americans in the atherosclerosis risk in communities (ARIC) study. *Am J Epidemiol*. 2016;183(7):650–6.
16. Balasubramanian R, Paynter NP, Giulianini F, Manson JE, Zhao Y, Chen JC, et al. Metabolomic profiles associated with all-cause mortality in the Women's Health Initiative. *Int J Epidemiol*. 2020;49(1):289–300.
17. Tang WH, Wang Z, Levison BS, Koeth RA, Britt EB, Fu X, et al. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *N Engl J Med*. 2013;368(17):1575–84.
18. Wurtz P, Havulinna AS, Soininen P, Tynkkynen T, Prieto-Merino D, Tillin T, et al. Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. *Circulation*. 2015;131(9):774–85.
19. Vignoli A, Tenori L, Giusti B, Takis PG, Valente S, Carrabba N, et al. NMR-based metabolomics identifies patients at high risk of death within two years after acute myocardial infarction in the AMI-Florence II cohort. *BMC Med*. 2019;17(1):3.
20. Hu JR, Coresh J, Inker LA, Levey AS, Zheng Z, Rebholz CM, et al. Serum metabolites are associated with all-cause mortality in chronic kidney disease. *Kidney Int*. 2018;94(2):381–9.
21. Yin J, Liao SX, He Y, Wang S, Xia GH, Liu FT, et al. Dysbiosis of gut microbiota with reduced trimethylamine-N-oxide level in patients with large-artery atherosclerotic stroke or transient ischemic attack. *J Am Heart Assoc*. 2015;4(11):e002699.
22. Mueller DM, Allenspach M, Othman A, Saely CH, Muendlein A, Vonbank A, et al. Plasma levels of trimethylamine-N-oxide are confounded by impaired kidney function and poor metabolic control. *Atherosclerosis*. 2015;243(2):638–44.
23. Zhao W, Zhao J, Rong J. Pharmacological modulation of cardiac remodeling after myocardial infarction. *Oxid Med Cell Longevity*. 2020;2020:8815349.
24. Karagiannidis E, Sofidis G, Papazoglou AS, Deda O, Panteris E, Moysidis DV, et al. Correlation of the severity of coronary artery disease with patients' metabolic profile- rationale, design and baseline patient characteristics of the CorLipid trial. *BMC Cardiovasc Disord*. 2021;21(1):79.
25. Poss AM, Maschek JA, Cox JE, Hauner BJ, Hopkins PN, Hunt SC, et al. Machine learning reveals serum sphingolipids as cholesterol-independent biomarkers of coronary artery disease. *J Clin Investig*. 2020;130(3):1363–76.
26. Yu N, Wang R, Liu B, Zhang L. Bibliometric and visual analysis on metabolomics in coronary artery disease research. *Front Cardiovasc Med*. 2022;9:804463.
27. Leu HB, Yin WH, Tseng WK, Wu YW, Lin TH, Yeh HI, et al. Identification of new biosignatures for clinical outcomes in stable coronary artery disease—the study protocol and initial observations of a prospective follow-up study in Taiwan. *BMC Cardiovasc Disord*. 2017;17(1):42.
28. Yu E, Ruizcanela M, Guaschferré M, Zheng Y, Toledo E, Clish CB, et al. Increases in plasma tryptophan are inversely associated with incident cardiovascular disease in the Prevención con Dieta Mediterránea (PREDIMED) study. *J Nutr*. 2017;147(3):314–22.
29. Zuo H, Ueland PM, Ulvik A, Eussen SJPM, Vollset SE, Nygård O, et al. Plasma biomarkers of inflammation, the kynurenine pathway, and risks of all-cause, cancer, and cardiovascular disease mortality: The Hordaland Health Study. *Am J Epidemiol*. 2016;183(4):kvv42.
30. Verheyen N, Meinitzer A, Grubler MR, Ablasser K, Kolesnik E, Fahrleitner-Pammer A, et al. Low-grade inflammation and tryptophan-kynurenine pathway activation are associated with adverse cardiac remodeling in primary hyperparathyroidism: the EPATH trial. *Clin Chem Lab Med*. 2017;55(7):1034–42.
31. Sallée M, Dou L, Cerini C, Poitevin S, Brunet P, Burtey S. The aryl hydrocarbon receptor-activating effect of uremic toxins from tryptophan metabolism: a new concept to understand cardiovascular complications of chronic kidney disease. *Toxins*. 2014;6(3):934–49.
32. Wang B, Liu S, Li L, Yao Q, Song R, Shao X, et al. Non-thyroidal illness syndrome in patients with cardiovascular diseases: a systematic review and meta-analysis. *Int J Cardiol*. 2017;226:1–10.
33. Mindikoglu AL, Opekun AR, Putluri N, Devaraj S, Sheikh-Hamad D, Vierling JM, et al. Unique metabolomic signature associated with hepato-renal dysfunction and mortality in cirrhosis. *Transl Res*. 2018;195:25–47.
34. Xiao Y, Zhang Y, Wang M, Li X, Su D, Qiu J, et al. Plasma S-adenosylhomocysteine is associated with the risk of cardiovascular events in patients undergoing coronary angiography: a cohort study. *Am J Clin Nutr*. 2013;98(5):1162–9.
35. Zawada AM, Rogacev KS, Hummel B, Berg JT, Friedrich A, Roth HJ, et al. S-adenosylhomocysteine is associated with subclinical atherosclerosis and renal function in a cardiovascular low-risk population. *Atherosclerosis*. 2014;234(1):17–22.
36. Li XS, Obeid S, Klingenberg R, Gencer B, Mach F, Räber L, et al. Gut microbiota-dependent trimethylamine N-oxide in acute coronary syndromes: a prognostic marker for incident cardiovascular events beyond traditional risk factors. *Eur Heart J*. 2017;38(11):814–24.
37. Koeth RA, Wang Z, Levison BS, Buffa JA, Elin O, Sheehy BT, et al. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat Med*. 2013;19(5):576–85.
38. Erdmann CC. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *N Engl J Med*. 2013;368(17):1575.
39. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature*. 2011;472(7341):57–63.
40. Zhu W, Gregory JC, Org E, Buffa JA, Gupta N, Wang Z, et al. Gut microbial metabolite TMAO enhances platelet hyperreactivity and thrombotic risk. *Cell*. 2016;165(1):111–24.
41. Seldin MM, Meng Y, Qi H, Zhu W, Wang Z, Hazen SL, et al. Trimethylamine N-oxide promotes vascular inflammation through signaling of mitogen-activated protein kinase and nuclear factor- $\kappa$ B. *J Am Heart Assoc*. 2016;5(2):e002767.
42. Ottosson F, Brunkwall L, Smith E, Orho-Melander M, Nilsson PM, Fernandez C, et al. The gut microbiota-related metabolite

- phenylacetylglutamine associates with increased risk of incident coronary artery disease. *J Hypertens.* 2020;38(12):2427–34.
43. Nemet I, Saha PP, Gupta N, Zhu W, Romano KA, Skye SM, et al. A cardiovascular disease-linked gut microbial metabolite acts via adrenergic receptors. *Cell.* 2020;180(5):862–77.e22.
  44. Tang HY, Wang CH, Ho HY, Lin JF, Lo CJ, Huang CY, et al. Characteristic of metabolic status in heart failure and its impact in outcome perspective. *Metabolites.* 2020;10(11):437.
  45. Razavi AC, Bazzano LA, He J, Li S, Fernandez C, Whelton SP, et al. Pseudouridine and *N*-formylmethionine associate with left ventricular mass index: Metabolome-wide association analysis of cardiac remodeling. *J Mol Cell Cardiol.* 2020;140:22–9.
  46. Tahir UA, Katz DH, Zhao T, Ngo D, Cruz DE, Robbins JM, et al. Metabolomic profiles and heart failure risk in black adults: insights from the Jackson Heart Study. *Circ Heart Fail.* 2021;14(1):e007275.
  47. Chevli PA, Freedman BI, Hsu FC, Xu J, Rudock ME, Ma L, et al. Plasma metabolomic profiling in subclinical atherosclerosis: the Diabetes Heart Study. *Cardiovasc Diabetol.* 2021;20(1):231.
  48. Ottosson F, Smith E, Fernandez C, Melander O. Plasma metabolites associate with all-cause mortality in individuals with Type 2 Diabetes. *Metabolites.* 2020;10(8):315.
  49. Alonso A, Yu B, Sun YV, Chen LY, Loehr LR, O'Neal WT, et al. Serum metabolomics and incidence of atrial fibrillation (from the Atherosclerosis Risk in Communities Study). *Am J Cardiol.* 2019;123(12):1955–61.
  50. Dunn WB, Broadhurst DI, Deepak SM, Buch MH, McDowell G, Spasic I, et al. Serum metabolomics reveals many novel metabolic markers of heart failure, including pseudouridine and 2-oxoglutarate. *Metabolomics.* 2007;3(4):413–26.
  51. Jing Z, Liu L, Shi Y, Du Q, Zhang D, Zuo L, et al. Association of coronary artery disease and metabolic syndrome: usefulness of serum metabolomics approach. *Front Endocrinol (Lausanne).* 2021;12:692893.
  52. Rhodes CJ, Ghataorhe P, Wharton J, Rue-Albrecht KC, Hadinnapola C, Watson G, et al. Plasma metabolomics implicates modified transfer RNAs and altered bioenergetics in the outcomes of pulmonary arterial hypertension. *Circulation.* 2017;135(5):460–75.
  53. Charette M, Gray MW. Pseudouridine in RNA: what, where how, and why. *IUBMB Life.* 2010;49(5):341–51.
  54. Chu SH, Cui J, Sparks JA, Lu B, Tedeschi SK, Speyer CB, et al. Circulating plasma metabolites and risk of rheumatoid arthritis in the Nurses' Health Study. *Rheumatology (Oxford).* 2020;59(11):3369–79.
  55. Gerner EW, Meyskens FL Jr. Polyamines and cancer: old molecules, new understanding. *Nat Rev Cancer.* 2004;4(10):781–92.
  56. Thomas T, Thomas TJ. Polyamines in cell growth and cell death: molecular mechanisms and therapeutic applications. *Cell Mol Life Sci.* 2001;58(2):244–58.
  57. Tantini B, Fiumana E, Cetrullo S, Pignatti C, Bonavita F, Shantz LM, et al. Involvement of polyamines in apoptosis of cardiac myoblasts in a model of simulated ischemia. *J Mol Cell Cardiol.* 2006;40(6):775–82.
  58. Cetrullo S, Tantini B, Facchini A, Pignatti C, Stefanelli C, Caldarera CM, et al. A pro-survival effect of polyamine depletion on norepinephrine-mediated apoptosis in cardiac cells: role of signaling enzymes. *Amino Acids.* 2011;40(4):1127–37.
  59. Beloborodova N, Bairamov I, Olenin A, Shubina V, Teplova V, Fedotcheva N. Effect of phenolic acids of microbial origin on production of reactive oxygen species in mitochondria and neutrophils. *J Biomed Sci.* 2012;19:89.
  60. Caussy C, Hsu C, Lo MT, Liu A, Bettencourt R, Ajmera VH, et al. Link between gut-microbiome derived metabolite and shared gene-effects with hepatic steatosis and fibrosis in NAFLD. *Hepatology.* 2018;68(3):918–32.
  61. Stepien M, Keski-Rahkonen P, Kiss A, Robinot N, Duarte-Salles T, Murphy N, et al. Metabolic perturbations prior to hepatocellular carcinoma diagnosis: findings from a prospective observational cohort study. *Int J Cancer.* 2021;148(3):609–25.
  62. Vangipurapu J, Fernandes Silva L, Kuulasmaa T, Smith U, Laakso M. Microbiota-related metabolites and the risk of type 2 diabetes. *Diabetes Care.* 2020;43(6):1319–25.
  63. Jurecka A, Zikanova M, Kmoch S, Tylki-Szymanska A. Adenylosuccinate lyase deficiency. *J Inher Metab Dis.* 2015;38(2):231–42.
  64. Furman D, Chang J, Lartigue L, Bolen CR, Haddad F, Gaudilliere B, et al. Expression of specific inflammasome gene modules stratifies older individuals into two extreme clinical and immunological states. *Nat Med.* 2017;23(2):174–84.
  65. Nam M, Jung Y, Ryu DH, Hwang GS. A metabolomics-driven approach reveals metabolic responses and mechanisms in the rat heart following myocardial infarction. *Int J Cardiol.* 2017;227:239–46.
  66. Heresi GA, Mey JT, Bartholomew JR, Haddadin IS, Tonelli AR, Dweik RA, et al. Plasma metabolomic profile in chronic thromboembolic pulmonary hypertension. *Pulm Circ.* 2020;10(1):2045894019890553.
  67. Webb-Robertson BM, Bramer LM, Stanfill BA, Reehl SM, Nakayasu ES, Metz TO, et al. Prediction of the development of islet autoantibodies through integration of environmental, genetic, and metabolic markers. *J Diabetes.* 2021;13(2):143–53.
  68. de Moraes MS, Guerreiro G, Sitta A, de Moura CD, Manfredini V, Wajner M, et al. Oxidative damage in mitochondrial fatty acids oxidation disorders patients and the in vitro effect of l-carnitine on DNA damage induced by the accumulated metabolites. *Arch Biochem Biophys.* 2020;679:108206.
  69. Talmor-Barkan Y, Bar N, Shaul AA, Shahaf N, Godneva A, Bussi Y, et al. Metabolomic and microbiome profiling reveals personalized risk factors for coronary artery disease. *Nat Med.* 2022;28(2):295–302.
  70. Qin M, Zhu Q, Lai W, Ma Q, Liu C, Chen X, et al. Insights into the prognosis of lipidomic dysregulation for death risk in patients with coronary artery disease. *Clin Transl Med.* 2020;10(5):e189.
  71. Luan H, Ji F, Chen Y, Cai Z. statTarget: A streamlined tool for signal drift correction and interpretations of quantitative mass spectrometry-based omics data. *Anal Chim Acta.* 2018;1036:66–72.
  72. Bujak R, Dagher-Wojtkowiak E, Kaliszczan R, Markuszewski MJ. PLS-based and regularization-based methods for the selection of relevant variables in non-targeted metabolomics data. *Front Mol Biosci.* 2016;3:35.
  73. Santulli G, Pascale V, Finelli R, Visco V, Giannotti R, Massari A, et al. We are what we eat: impact of food from short supply chain on metabolic syndrome. *J Clin Med.* 2019;8(12):2061.
  74. Wang Z, Zhu Q, Liu Y, Chen S, Zhang Y, Ma Q, et al. Genome-wide association study of metabolites in patients with coronary artery disease identified novel metabolite quantitative trait loci. *Clin Transl Med.* 2021;11(2):e290.
  75. Yang J, Yan B, Zhao B, Fan Y, He X, Yang L, et al. Assessing the causal effects of human serum metabolites on 5 major psychiatric disorders. *Schizophr Bull.* 2020;46(4):804–13.
  76. Choi KW, Chen CY, Stein MB, Klimentidis YC, Wang MJ, Koenen KC, et al. Assessment of bidirectional relationships between physical activity and depression among adults: a 2-sample Mendelian Randomization Study. *JAMA Psychiat.* 2019;76(4):399–408.
  77. MacKinnon MD. Introduction to statistical mediation analysis. New York: Erlbaum; 2008.
  78. Rijnhart JJM, Lamp SJ, Valente MJ, MacKinnon DP, Twisk JWR, Heymans MW. Mediation analysis methods used in observational research: a scoping review and recommendations. *BMC Med Res Methodol.* 2021;21(1):226.
  79. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychol Methods.* 2010;15(4):309–34.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

