UNIVERSITY
OF WARMIA AND MAZURY
IN OLSZTYN

SCHOOL OF MEDICINE
Field of Medical and Health Sciences
Discipline of Medical Sciences

Jakub Morze, M.D.
Doctoral thesis

Application of metabolomics in the assessment of risk and progression of heart failure

Badanie metabolomu w ocenie ryzyka i przebiegu niewydolności serca

Thesis completed in the Department of Cardiology and Internal Medicine,
School of Medicine, University of Warmia and Mazury in Olsztyn

under supervision of
Professor Andrzej Rynkiewicz M.D., Ph.D., D.Sc.
University of Warmia and Mazury in Olsztyn

and

Associate Professor Marta Guasch-Ferré, Ph.D.
University of Copenhagen and Harvard University

Olsztyn, 2023

This work is licensed under a Creative Commons License:
Attribution-NonCommercial-ShareAlike 4.0 International
To find a copy of this license visit the following link:
https://creativecommons.org/licenses/by-nc-sa/4.0/

Table of Contents

Abstracts *9*

English abstract **9**

Polish abstract **10**

1. Introduction *11*

Heart failure **11**

Definition 11

Phenotypes and classification 12

Causes 14

Prevalence 15

Outcomes 16

Mortality 16

Hospitalizations 17

Circulating biomarkers of heart failure **19**

Natriuretic peptides 20

Alternative biomarkers 21

Metabolism and metabolites in heart failure **23**

Metabolism in healthy heart 23

Metabolism in failing heart 26

Other metabolic pathways in heart failure 28

Retrospective studies on metabolomics and heart failure 32

Prospective studies on metabolomics and heart failure 36

2. Hypotheses and study scope *39*

Rationale: 39

Hypothesis 39

Research Aims 39

3. Data sources and methods *41*

UK Biobank Cohort **41**

Overview 41

Data access 42

Ethical approval 42

Eligibility criteria for study subcohorts 42

Baseline covariate assessment 44

Nuclear magnetic resonance metabolite measure profiling **48**

Overview 48

Nomenclature 49

Preparation of Plasma Samples 49

Procedure for NMR Spectroscopy 49

Quantified metabolite measures 50

Quality control 52
Data preprocessing 52
Outcomes 53
Longitudinal health data 53
Hospital inpatient data 53
Death register 54
Terminology for health events 55
Definition of health events 55
Cardiovascular magnetic resonance imaging protocol 58
Statistical analysis 59
Data preprocessing 61
Summary statistics 61
Cox proportional hazards models 62
Subgroup and sensitivity analyses 63
Mediation analyses 64
Additional hypothesis-driven models 64
General linear models 65
Metabolite measure networks 65
Multi-metabolite measure score calculation 66
Predictive performance 67
Statistical software 68
4. Results 69
Single metabolite markers and heart failure risk 70
Metabolome-wide analysis with basic covariate adjustment 70
Metabolome-wide analysis with essential covariate adjustment 72
Subgroup analyses for 90 heart failure related metabolite measures 84
Sensitivity analyses for 90 heart failure related metabolite measures 95
Mediation analyses for indirect effects of interim events on metabolite measure-HF associations 101
Assessment of relative metabolite measure validity 111
Comparison of particle count and lipid content predictive power 114
Multi-metabolite signatures and incident heart failure 115
Heart failure metabolic susceptibility score 115
Metabolite measures networks and incident heart failure 118
Prediction performance of multi-metabolite measure scores for heart failure 125
Metabolite markers in relation to cardiac function in magnetic resonance imaging 129
Metabolome-wide analysis for automatically annotated MRI parameters of left ventricle 130
Multi-metabolite measure scores and automatically annotated MRI parameters of left ventricle 137
Metabolite markers in relation to survival among heart failure patients 142
Metabolome-wide analysis for all-cause and cardiovascular mortality in heart failure patients 144

Multi-metabolite measure scores for all-cause and cardiovascular mortality in heart failure patients 147
5. Discussion 149
Summary of study findings 149
Potential mechanisms 152
Aromatic amino acids 152
Branched-chain amino acids 153
Glycine 154
Histidine 155
Other amino acids 155
Glycolysis 156
Ketone bodies 157
Fatty acids 159
Inflammation 161
Fluid balance 162
Triglyceride-rich lipoproteins 163
Cholesterol-enriched lipoproteins 165
High-density lipoproteins 166
Clinical application and future perspectives 167
Strengths and limitations 170
Conclusions 172
References 174

Tables

Table 1. Comparison of participants' baseline characteristics between those with and without incident heart failure during follow-up. 69
Table 2. Goodness of fit statistics for ratio measures relevant for heart failure risk, their nominators, and denominators. 112
Table 3. Comparison of relative importance of metabolite measures in small and very large HDL class 114
Table 4. Composition and coefficients of heart failure metabolic susceptibility scores. 115
Table 5. Associations of unweighted and weighted heart failure susceptibility scores with incident heart failure. 117
Table 6. Weights for cluster-specific scores derived from Cox regression model with elastic net regularization. 121
Table 7. Hazard ratios (95% CI) for incident heart failure associated with one-standard deviation increase in metabolite measure cluster-specific scores. 123
Table 8. Associations between cluster-specific scores and incident heart failure adjusted for interim events. 124
Table 9. Comparison of area under curve, net reclassification index, integrated discriminatory improvement parameters for HF prediction models with and without multi-metabolite measure scores. 128
Table 10. Comparison of participants' baseline characteristics between overall cohort and sub-cohort who underwent heart magnetic resonance imaging assessment. 129
Table 11. Associations between multi-metabolite measure scores and cardiac index in magnetic resonance imaging. 137
Table 12. Associations between multi-metabolite measure scores and left ventricle ejection fraction in magnetic resonance imaging. 138
Table 13. Associations between multi-metabolite measure scores and left ventricle end-systolic volume index in magnetic resonance imaging. 139
Table 14. Associations between multi-metabolite measure scores and left ventricle end-diastolic volume index in magnetic resonance imaging. 140
Table 15. Associations between multi-metabolite measure scores and stroke volume index in magnetic resonance imaging. 141
Table 16. Comparison of participants' baseline characteristics between overall subcohort of prevalent HF and individuals who died during follow-up. 142
Table 17. Associations between multi-metabolite measure scores and all-cause mortality in participants with prevalent heart failure. 147
Table 18. Associations between multi-metabolite measure scores and cardiovascular mortality in participants with prevalent heart failure. 148

Figures

Figure 1. Flowchart for selection of UK Biobank participants to the present study.....43
Figure 2. Flow-chart summarizing the main steps in the statistical analysis plan. 60
Figure 3. Volcano plot summarizing associations between 249 metabolite measures with incident heart failure adjusted for age, sex, and body mass index.....70
Figure 4. Multivariable-adjusted hazard ratios (95%CI) for heart failure associated with metabolic measures from apolipoprotein, lipoprotein particle concentrations, lipoprotein particle sizes and total lipids classes. 73
Figure 5. Multivariable-adjusted hazard ratios (95%CI) for heart failure associated with metabolic measures from overall cholesterol, cholesteryl esters, free cholesterol, other lipids, phospholipids, and triglycerides classes. 74
Figure 6. Multivariable-adjusted hazard ratios (95%CI) for heart failure associated with metabolic measures from fatty acids class and their ratios. 76
Figure 7. Multivariable-adjusted hazard ratios (95%CI) for heart failure associated with metabolic measures from VLDL size subclasses.....78
Figure 8. Multivariable-adjusted hazard ratios (95%CI) for heart failure associated with metabolic measures from IDL and LDL size subclasses. 79
Figure 9. Multivariable-adjusted hazard ratios (95%CI) for heart failure associated with metabolic measures from HDL size subclasses. 80
Figure 10. Multivariable-adjusted hazard ratios (95%CI) for heart failure associated with metabolic measures from amino acids, ketone bodies, glycolysis-related, fluid balance and inflammation metabolite measures. 82
Figure 11. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by baseline age of participants. 84
Figure 12. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' sex.....85
Figure 13. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' self-reported ethnicity.....86
Figure 14. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' baseline deprivation index.....87
Figure 15. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' baseline body mass index. 88
Figure 16. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' baseline alcohol intake. 89
Figure 17. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' baseline smoking status. 90
Figure 18. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' fasting time during sample collection. 91
Figure 19. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' baseline estimated glomerular filtration rate. 92
Figure 20. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' baseline systolic blood pressure. 93
Figure 21. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' baseline glycated hemoglobin level. 94
Figure 22. Sensitivity analysis for metabolite measures significantly associated with incident heart failure after exclusion of first three years of follow-up. 96
Figure 23. Sensitivity analysis for metabolite measures significantly associated with incident heart failure using age as underlying time scale for regression models. 97

Figure 24. Sensitivity analysis for metabolite measures significantly associated with incident heart failure after further adjustment for baseline self-reported physical activity level. 98
Figure 25. Sensitivity analysis for metabolite measures significantly associated with incident heart failure after further adjustment for baseline healthy diet score. 99
Figure 26. Sensitivity analysis for metabolite measures significantly associated with incident heart failure after inclusion of non-linear covariate terms to regression models. 100
Figure 27. Multivariable-adjusted hazard ratios (95%CI) for interim coronary artery disease (left pane) and incident heart failure after adjustment for this interim event (right pane).....102
Figure 28. Multivariable-adjusted hazard ratios (95%CI) for interim atrial fibrillation (left pane) and incident heart failure after adjustment for this interim event (right pane). 104
Figure 29. Multivariable-adjusted hazard ratios (95%CI) for interim valvular heart disease (left pane) and incident heart failure after adjustment for this interim event (right pane).....106
Figure 30. Multivariable-adjusted hazard ratios (95%CI) for interim diabetes mellitus (left pane) and incident heart failure after adjustment for this interim event (right pane). 108
Figure 31. Proportion of association between metabolite measures and incident heart failure attributable to occurrence of interim events.....110
Figure 32. Histogram presenting distribution of unweighted and weighted heart failure metabolic susceptibility scores. 116
Figure 33. Kaplan-Meier curve of cumulative incidence of heart failure stratified by quartiles of weighted (left pane) and (right pane) unweighted heart failure metabolic suspectibility scores.....116
Figure 34. Networks of metabolite measures derived by Gaussian Graphical Model.....118
Figure 35. Heatmap of correlations between ten derived cluster-specific scores. 123
Figure 36. Receiver operator characteristic curves for HF prediction models with and without multi-metabolite measure scores.....125
Figure 37. Volcano plot summarizing associations between 168 non-ratio metabolite measures with cardiac index. 130
Figure 38. Volcano plot summarizing associations between 168 non-ratio metabolite measures with left ventricle ejection fraction. 132
Figure 39. Volcano plot summarizing associations between 168 non-ratio metabolite measures with left ventricle end-systolic volume index. 133
Figure 40. Volcano plot summarizing associations between 168 non-ratio metabolite measures with left ventricle end-diastolic volume index.....134
Figure 41. Volcano plot summarizing associations between 168 non-ratio metabolite measures with left ventricle end-systolic volume index. 136
Figure 42. Kaplan-Meier curve of the cumulative all-cause mortality among participants with heart failure at baseline. 143
Figure 43. Kaplan-Meier curve of the cumulative cardiovascular mortality among participants with heart failure at baseline. 144
Figure 44. Volcano plot summarizing associations between non-ratio 168 metabolite measures with all-cause mortality among participants with prevalent HF. 145
Figure 45. Volcano plot summarizing associations between non-ratio 168 metabolite measures with cardiovascular mortality among participants with prevalent HF. 146

Abstracts

English abstract

Heart failure (HF) is characterized by a complex clinical syndrome with a multifactorial etiology and elevated rates of morbidity and mortality. High-throughput metabolomics, the comprehensive analysis of metabolites produced by cellular processes, holds potential for elucidating the underlying mechanisms of HF. Despite the growing popularity of this technique, its application to HF has been limited. In this study, identification of a metabolomic signature linked to the development and progression of HF was pursued, utilizing data from a prospective observational cohort in the UK Biobank. Utilizing Nightingale Health's nuclear magnetic resonance platform, 249 metabolite measures were quantified. Emphasis was placed on examining the prospective associations between these metabolite measures and incident HF during the follow-up period. Relationships between metabolite levels and left ventricular function, as determined through magnetic resonance imaging, were also scrutinized, as were survival outcomes in participants with prevalent HF. An inverse association with HF risk was observed in elevated levels of certain metabolite measures, such as omega-3 fatty acids, small HDL particles, albumin, and specific amino acids like alanine and histidine. On the other hand, metabolite measures such as ketone bodies, very large HDL particles, and glycoprotein acetyls were found to have a positive association with HF risk. This study offers observational evidence that may support identification of new metabolic biomarkers in HF.

Polish abstract

Niewydolność serca (NS) jest złożonym zespołem klinicznym o wieloczynnikowej etiologii, charakteryzującym się wysoką chorobowością i śmiertelnością. Metabolomika pozwala na integralną ocenę dużej liczby metabolitów stwarzając potencjał do pogłębienia wiedzy o patogenezie NS. Pomimo rosnącej popularności tej techniki, jej zastosowanie w badaniu NS pozostaje ograniczone. W niniejszej pracy podjęto próbę identyfikacji profili metabolomicznych związanych z rozwojem i przebiegiem NS, w oparciu o dane z prospektywnego, obserwacyjnego badania kohortowego UK Biobank. Wykorzystując platformę metabolomiczną Nightingale Health, scharakteryzowano 249 różnych metabolitów w osoczu krwi uczestników badania. Następnie przeanalizowano prospektywne zależności między stężeniami metabolitów a ryzykiem rozwoju NS. Określono również związek pomiędzy metabolitami a funkcją lewej komory serca ocenioną w rezonansie magnetycznym, jak również wpływ na śmiertelność wśród pacjentów z rozpoznaną już NS. W badaniu zaobserwowano odwrotną zależność między ryzykiem NS a podwyższonymi stężeniami metabolitów, takich jak kwasy tłuszczowe omega-3, małe cząsteczki HDL, albumina oraz specyficzne aminokwasy jak alanina i histydyna. Z kolei metabolity takie jak ciała ketonowe, bardzo duże cząsteczki HDL i acetylowane glikoproteiny wykazały pozytywną zależność z ryzykiem NS. Niniejsze badanie dostarcza dowodów obserwacyjnych, które mogą przyczynić się do identyfikacji nowych biomarkerów metabolicznych w NS.

1. Introduction

Heart failure

Definition

Heart failure (HF) is a clinical syndrome characterized by various presentations, with a range of underlying etiologies and complex pathophysiological mechanisms, rather than constituting a single, specific disease. Consequently, defining HF poses a significant challenge compared to diseases that have definitive pathology measures to confirm their diagnosis such as myocardial infarction or diabetes mellitus (DM). That is why definition of HF exhibits substantial heterogeneity across clinical textbooks, society guidelines, and standard medical practice (Heidenreich et al., 2022; McDonagh et al., 2021; McDonagh et al., 2023). Recently, a joint effort of Heart Failure Society of America, Heart Failure Association (HFA) of the European Society of Cardiology (ESC), Japanese Heart Failure Society led to creation of the universal definition of HF (Bozkurt et al., 2021).

According to this definition, HF can be defined as a clinical syndrome characterized by: (1) the current or previous presence of typical symptoms or signs as a consequence of structural or functional cardiac abnormality, accompanied by (2) increased concentration of natriuretic peptides or (3) objective confirmation of pulmonary or systemic congestion in imaging or hemodynamic assessment conducted at rest or during provocation.

Typical symptoms of HF include breathlessness, orthopnea, and paroxysmal nocturnal dyspnea, which are all indicative of impaired gas exchange and fluid retention in the lungs. Exercise tolerance is notably reduced, often accompanied by fatigue and a general feeling of tiredness. Peripheral edema, particularly ankle swelling, is common and sometimes extends to other body parts. Elevated jugular venous pressure, a third heart sound, and cardiomegaly are among the most specific signs of HF, often accompanied by hepatojugular reflux and Cheyne–Stokes respiration in advanced cases. Less specific signs include peripheral edema, pulmonary rales, and cardiac murmurs, as well as weight fluctuations and tachycardia. Additional manifestations can include hepatomegaly, ascites, and cold extremities (McDonagh et al., 2021).

Considering the low specificity of signs and symptoms in HF, elevated natriuretic peptides, as well as evidence for cardiac structural or functional abnormality, substantially improve the

diagnostic accuracy of the above-mentioned definition. In an ambulatory setting, a brain natriuretic peptide (BNP) concentration of ≥35 pg/ml or an N-terminal prohormone of BNP (NT-proBNP) concentration of ≥125 pg/ml is suggestive of an HF diagnosis. In hospitalized or decompensated patients, a BNP concentration of ≥100 pg/ml or an NT-proBNP concentration of ≥300 pg/ml is similarly indicative of HF (McDonagh et al., 2021).

Structural and functional cardiac abnormalities causing signs and symptoms can include decreased left ventricle (LV) ejection fraction (LVEF), ventricular hypertrophy, decreased global longitudinal strain, increased ratio between early mitral inflow velocity and mitral annular early diastolic velocity, ventricular hypertrophy, atrial enlargement, valvular stenosis or regurgitation (Thomas, 2015). For the objective verification of pulmonary congestion, modalities such as chest radiography, pulmonary ultrasonography, and thoracic computed tomography are instrumental. In the context of systemic congestion, evaluation may be performed using inferior vena cava, jugular vein, or renal ultrasonography (Boorsma et al., 2020).

Phenotypes and classification

Due to its broad definition and varied clinical manifestations, HF can be classified in multiple ways. Among other methods, this includes stratification based on factors such as the onset timing of symptoms, LV systolic function, disease stage or advancement, symptom severity, and underlying cause.

By the timing of symptoms onset, HF is generally categorized into two main categories: chronic HF (CHF) and acute HF (AHF). CHF pertains to individuals with an existing HF diagnosis or those experiencing a gradual emergence of symptoms. When CHF worsens, either abruptly or progressively, the condition is often termed as HF decompensation. On the other hand, AHF is characterized by a rapid or gradual onset of HF symptoms or signs that are severe enough to compel the patient to seek immediate medical intervention. AHF can either represent the new onset of HF or more commonly, result from an acute decompensation of pre-existing CHF (Heidenreich et al., 2022; McDonagh et al., 2021).

Currently, the most clinically important division of HF is based on LVEF. LVEF serves as a crucial parameter in the categorization of HF patients due to its implications for both prognosis and treatment responsiveness. This distinction becomes especially important because most randomized controlled trials (RCTs) focused on HF outcomes primarily enrolled patients with

impaired systolic function (those with an LVEF of ≤35% or ≤40%) (Bozkurt et al., 2021). HF characterized by a reduced LVEF (HFrEF), defined as an LVEF below 40%, is well-understood with established effective treatments. Conversely, HF with preserved LVEF (HFpEF), which can be defined in varying terms such as LVEF greater or equal 40%, 45%, or even 50%, has eluded effective therapeutic strategies, with no clinical trials showing conclusive benefits. The 2016 ESC guidelines introduced a new category known as HF with mid-range LVEF (HFmrEF), defined by an LVEF between 40-49%, and defined HFpEF as LVEF greater or equal 50% (Ponikowski et al., 2016). Subsequent studies have found that HFmrEF exhibits characteristics that are intermediate between HFrEF and HFpEF, but also show distinct similarities to HFrEF. This has led to proposals for renaming HFmrEF to HF with mildly-reduced EF in current guidelines (McDonagh et al., 2021). Most recently, evidence for the consistent benefit of using sodium-glucose co-transporter 2 (SGLT2) inhibitors across a range of LVEF values justifies further subclassification of this third group (McDonagh et al., 2023).

The American College of Cardiology (ACC) and the American Heart Association (AHA) categorize HF into four progressive stages – A through D – applicable to all HF patients with no distinction to LVEF (Heidenreich et al., 2022). Stage A encompasses individuals at high risk for developing HF, often due to risk factors like hypertension, DM, or a family history of cardiomyopathy, but who currently exhibit neither symptoms nor structural heart abnormalities. Stage B consists of patients without HF symptoms but who show evidence of structural heart disease, such as LV hypertrophy, or abnormal cardiac function like diastolic dysfunction, often identified through advanced imaging or elevated biomarker levels. Stage C includes patients who have structural or functional cardiac abnormalities accompanied by current or prior symptoms of HF. Lastly, Stage D refers to advanced HF cases characterized by severe symptoms at rest, recurrent hospitalizations, and the need for specialized interventions like transplantation or mechanical circulatory support. Each stage represents a point along a continuum, allowing for targeted interventions based on symptomology, structural abnormalities, and disease progression. Despite the intuitive nature of ACC/AHA staging, especially for the identification of risk factors and opportunities for preventive intervention, it has not been adopted as a guideline for recruitment in ongoing RCTs.

Stratification of HF patients according to the New York Heart Association (NYHA) functional classification serves both as a cornerstone in clinical assessment and as a criterion for trial eligibility (Bozkurt et al., 2021; McDonagh et al., 2021). This system is divided into four

classes. Class I represents patients who experience no limitations in physical activity and do not suffer from undue breathlessness, fatigue, or palpitations during ordinary activities. Class II includes those who are comfortable at rest but face slight limitations during regular physical activities, experiencing symptoms such as breathlessness, fatigue, or palpitations. In Class III, patients are also comfortable at rest, but experience marked limitations with even less than ordinary physical activity, manifesting similar symptoms. Finally, Class IV consists of patients who are unable to engage in any form of physical activity without experiencing discomfort; these individuals may also have symptoms while at rest. Despite the traditional role of NYHA in HF management, recent research suggests that NYHA inadequately stratifies patients across the spectrum of functional impairment (Caraballo et al., 2019).

HF can also be phenotyped according to its underlying etiology. For example, recent analyses of the HOMAGE and UK Biobank cohorts have shown considerable differences in clinical characteristics and prognostic factors at the time of diagnosis between participants with and without a prior history of myocardial infarction (Rastogi et al., 2022). However, the broad application of this classification is limited by a lack of consensus on confirming a specific etiology in HF patients and the overlap between predisposing events.

Causes

HF can be considered the end-stage downstream manifestation of various cardiovascular diseases (CVDs). The most common predisposing factors for HF include coronary artery disease (CAD), hypertension, valvular heart disease (VHD), arrhythmias (particularly atrial fibrillation [AF]), cardiomyopathies (including toxic-induced), infective conditions, and congenital heart diseases (McDonagh et al., 2021).

According to the ESC Heart Failure Long-Term Registry, CAD was the identified etiology in 43% of cases, hypertension in 8%, idiopathic dilated cardiomyopathy in 30%, and VHD in 8% (Chioncel et al., 2017). In the Global Congestive Heart Failure Registry, the respective etiologies were observed in 38%, 20%, 16%, and 9% of cases (Joseph et al., 2020). The Global Burden of Disease study estimated that CAD and hypertension were the cause of 41% and 31% of HF cases in men, and 35% and 35% in women, respectively (Yan et al., 2023).

CAD as an etiology can be twice as frequent in HFrEF, while cardiomyopathy can be three times as frequent. On the other hand, hypertension can be four times more likely to be a cause of HFpEF (Chioncel et al., 2017).

Prevalence

Providing an accurate estimation of the HF prevalence poses a significant challenge. This is not only due to differences across definitions of HF used by various study groups, but also despite the number of efforts (Bachtiger et al., 2022), lack of easy and reliable screening tool that could be adopted on a mass scale. Therefore, compared with cross-sectional and cohort studies, which adopt a uniform panel of assessments for each participant allowing to capture previously undiagnosed cases, HF seems to be underreported in population-based national registries (Cuthbert et al., 2019). While hospital-based observational studies are traditionally recognized to be prone to selection bias, population-based studies are often non-representative for the local population making their prevalence estimates less generalizable (Fry et al., 2017).

Nevertheless, irrespective of the source of data on HF prevalence, it is commonly recognized as a global epidemic (Roger, 2018). According to data from the Global Burden of Disease, an estimated 56.2 million individuals globally were living with HF in 2019, translating to a worldwide prevalence rate of 0.7% (Yan et al., 2023). Compared to the 1990 estimate of 27.2 million individuals, this represents a more than twofold increase in the prevalence of HF in under three decades. Results from the National Health and Nutrition Survey (NHANES) indicated that approximately 6.7 million Americans aged 20 or older (or ~2.6% of the US adult population) were living with HF between 2017 and 2020. This figure represents an increase from roughly 6.0 million, as reported by NHANES for the period 2015 to 2018 (Tsao et al., 2023). The HFA Atlas survey based on national statistics data conducted in 2018-2019 across 42 ESC countries revealed a median prevalence of HF of 17.2 cases per 1000 people, which translates to 1.7%. The prevalence ranged from less than or equal to 1.2% in Greece and Spain to more than 3.0% in Lithuania and Germany (Seferović et al., 2021). Existing data on the prevalence of HF are disproportionately represented by studies from Europe and the U.S., despite emerging evidence of a rising burden in Asia (Shimokawa et al., 2015).

HF prevalence sharply increases with age. In the elderly population (individuals aged 65 and over) the prevalence exhibits a nearly exponential growth. Analysis of primary care data from 56,320 adults in the Nijmeegs Monitoring Project, showed HF prevalence of 0.04% in the 18-44 age group, 0.3% for ages 45-54, 0.8% for 55-64, 2.2% for 65-74, escalating to 8.0% for ages 75-84, and peaking at 20.9% for those aged 85 and older (Bosch et al., 2019). While men display a slightly higher overall prevalence of HF than women, the more significant distinctions in prevalence between the sexes become evident when categorized by LVEF (Yan et al., 2023).

Men have a higher predisposition toward HFrEF, while women have a higher predisposition toward HFpEF. Data from the Swedish Heart Failure Registry indicated that of 42,131 patients with diagnosed HF women represented 55% of HFpEF, 39% of HFmrEF, and 29% of HFrEF patients (Stolfo et al., 2019). Globally, approximately half of all individuals with HF, based on LVEF are categorized as having HFrEF. In the ESC Heart Failure Long-Term Registry, among 9,134 HF patients, 59.8% were classified as having HFrEF, 24.2% had HFmrEF, and 16.0% had HFpEF (Chioncel et al., 2017). Global Congestive Heart Failure Registry enrolling 23,047 adult HF patients across 40 countries showed that 53.6%, 20.6%, and 24.2% of patients had respectively HFrEF, HFmrEF, and HFpEF (Joseph et al., 2020). There are also significant time trends that show an increase in the prevalence of HFpEF, while the prevalence of HFrEF appears to be plateauing or modestly decreasing. Over a 16-year period, the SwedeHF Registry data revealed a decline in the prevalence of HFrEF from 60% to 49%, an increase in HFpEF from 21% to 30%, while the contribution of HFmrEF at around 20% (Canepa et al., 2022).

The burden of HF patients is expected to grow substantially in the coming years, primarily due to global population growth and increased longevity resulting from improvements in the treatment and diagnosis of CVDs. The prevalence of HFrEF is expected to decline, largely due to advancements in the prevention and treatment of CAD. Conversely, the prevalence of HFpEF is projected to rise, in part because of increasing rates of obesity and hypertension, making it a more frequently occurring form of HF.

Outcomes

The trajectory of HF is often marked by periods of stability and acute exacerbation. Patients with CHF may be well-managed as outpatients but can experience episodes of acute decompensation requiring hospitalization. Accumulating number of acute events is inevitably associated with decreased cardiac function leading to death of a patient. Therefore, mortality, morbidity (as indicated by hospital admissions), and patient-reported quality of life are the most important outcomes considered in the management of HF.

Mortality

HF is commonly recognized as a condition with a poor prognosis. All-cause mortality in patients with HF has gradually decreased over the years due to advancements in treatment, particularly since the first large, randomized trials on HF were conducted in the late 1980s. However, despite increased efforts towards better diagnosis and treatment, all-cause mortality

rates remain very high. A historical report from the Framingham Heart Study (FHS), analyzing HF patients diagnosed between 1950 and 1999, showed significant improvements in age-adjusted mortality rates over the five decades. The 30-day mortality rate in women fell from 18% to 10%, while one-year mortality rates varied but generally decreased for both genders, e.g., from 30% to 28% in men. The five-year mortality rate also declined, dropping from 70% to 59% in men and from 57% to 45% in women (Levy et al., 2002). The population-based cohort study of 55,959 patients with a new diagnosis of HF, made between 2000 and 2017, found substantial improvements in survival rates: one-year survival rates increased from 74.2% in 2000 to 80.8% in 2016, five-year survival rates rose by from 41.0% to 48.2%, and ten-year survival rates grew from 19.8% to 26.2% (Taylor et al., 2019). A large systematic-review and meta-analysis including 60 studies with data for 1.5 million patients with HF estimated survival rates to be 95.7% at 1 month, 86.5% at 1 year, 72.6% at 2 years, 56.7% at 5 years, and 34.9% at 10 years (Jones et al., 2019).

Cardiovascular deaths account for the largest number of deaths in HF participants. The second commonest cause of death for people with HF was respiratory disease followed by cancer (Taylor et al., 2019). In the FHS, CVD were the cause of 66.1% of deaths in HF patients diagnosed between 1971-2004 (Lee et al., 2011). Analysis of 88,416 individuals with HF in the UK Clinical Practice Research Datalink and Hospital Episodes Statistics indicated that CVD deaths contributed to 50% of all HF deaths (Lawson et al., 2019). The same estimate was reported by previously mentioned systematic review of 60 studies, however there was a considerable inconsistency between studies (from 8% to 64%). Paradoxically, while CVD-related deaths in HF patients are generally decreasing over time, non-CVD-related deaths are experiencing a steady increase. A population-based study in the UK observed that one-year CVD mortality rates in HF patients declined from 18% in 2002 to 13% in 2013, while non-cardiovascular mortality rates increased from 13% to 17%, over the same period. The decline in cardiovascular mortality was observed across all age groups, albeit less steeply in the older population, where the increase in non-cardiovascular mortality was largely concentrated (Conrad et al., 2019).

Hospitalizations

Hospitalizations for HF are a significant public health concern, particularly among the elderly. In high-income countries, HF is the leading primary diagnosis for hospitalization among individuals aged over 65, accounting for 1–2% of all hospital admissions (Blecker et al., 2013).

In the US between 2008 and 2018 there were approximately 11.5 million hospitalizations with HF, with the annual number of hospitalizations increasing from ~1 million in 2009 to ~1.2 million in 2018 (Clark et al., 2022). Although, the overall number of hospitalizations rises due to higher HF prevalence, the rates of hospitalization gradually drop due to advancement in HF management (Ziaeian et al., 2017). In the European Union, the age-standardized hospitalization rate for HF in 2019 was 277 admissions per 100,000 individuals, showing marked regional variability. The lowest admission rate was observed in Portugal (87 admissions per 100,000), while the highest was found in Poland (616 admissions per 100,000) (OECD European Union, 2022).

Hospitalization serves as a pivotal turning point in the clinical trajectory of HF, substantially increasing the likelihood of subsequent readmissions and increasing mortality rates in comparison to patients who have not been hospitalized. Initial hospital visit during which a HF diagnosis was made was associated with 46% and 55% higher risk of all-cause and CVD mortality in a retrospective UK cohort (Lawson et al., 2019). Another population-based cohort set in the UK observed 2.4 years shorter survival in patients who were admitted to the hospital at the time of diagnosis. Survival in this group of patients was 81.2%, 51.8%, 28.8%, and 15.5% at respectively 1, 5, 10, 15 years since the diagnosis compared with 68.8%, 36.7%, 17.8% and 8.1% observed for patients in whom the diagnosis was made in the outpatient setting (Taylor et al., 2019).

While hospitalization appears to have a lasting impact on the individual risk profiles of patients, and a considerable number of events do occur during the hospital stay, the overwhelming majority of these events actually occur post-discharge in patients with HF. The immediate 2-3 month period following hospital release is especially fraught, known as the 'vulnerable phase,' and carries a particularly high risk of adverse clinical outcomes (Greene et al., 2015). Analyses of data from the ASCEND trial indicated CHF patients with prior hospitalization within the last 12 months have higher adverse outcomes compared to those without. For instance, the 180-day all-cause mortality rate was 15.5% in patients with prior hospitalization versus 11.9% in those without, and the 30-day all-cause rehospitalization rate was 4.0% versus 3.1%, respectively. However, the difference for 180-day seemed to be explained by other patient characteristics (Blumer et al., 2021).

Hospital admission does not only affect risk of death, but substantially increases morbidity. In the first year after diagnosis of HF, 40% of patients had at least one and 20% had multiple

hospitalizations (Lawson et al., 2019). According to the Medicare data from 2008 to 2010, among participants aged 65 years and more who were admitted to the hospital due to HF, 67.8% were re-admitted during the first year. The highest risk for re-admission was observed for first week since the hospitalization and decreased by half between on the 38th day since the first admission. Compared with those without HF hospitalization, beneficiaries admitted to the hospital due to HF had respectively sixteen and four times higher risks of re-admission at 30 day and 1 year (Dharmarajan et al., 2015). This is also why majority of HF hospitalizations are represented by re-admission rather than primary HF admission. In REPORT-HF registry, in total 43% patients were admitted with new-onset HF, and 57% patients had decompensation of previous HF diagnosis (Tromp et al., 2022). Mortality rates notably rise following each successive hospitalization for HF. In an analysis of UK Clinical Practice Research Datalink among patients with one recurrent HF hospitalization the annualized rate for all-cause mortality was 49 death per 100 person-years and rose to respectively to 73 and 89 deaths per 100 person-years for second and fourth re-admission. As a comparison this rate among patients with no re-admission was 29 deaths per 100 person-years (Lahoz et al., 2020).

Although HF is a leading cause for hospitalization, most admissions among HF patients are surprisingly for non-cardiovascular reasons. In the US between 2002 and 2013 admissions with HF as the primary reason represented 13% of hospitalizations within a year since the diagnosis. Overall, hospitalizations associated with any cardiovascular reason (e.g., HF or another CVD) represented fewer than half of all admissions. Other causes included infections, neoplasms, chronic respiratory diseases, kidney diseases, digestive diseases, and injuries (Conrad et al., 2019).

Circulating biomarkers of heart failure

A biomarker, short for 'biological marker', is described as a measurable trait that serves as an indicator of either normal biological activities, disease-related processes, or the body's response to therapeutic treatments (Biomarkers Definitions Working Group, 2001). The assessment of substances in the bloodstream is the most prevalent method for measuring biomarkers and has gained extensive application in clinical studies. In the context of HF, the use of biomarkers serves multiple objectives, including the identification of high-risk individuals and patients with subclinical disease, prediction of symptom onset, support for HF diagnosis, projection of disease trajectories, guidance for therapeutic management, and serving as surrogate endpoints (Ibrahim & Januzzi, 2018). Despite variations in the definitions of HF, the role of circulating

biomarkers in HF has been steadily increasing over the past two decades. This is especially true for natriuretic peptides like BNP and NT-proBNP, whose utility spans all the aforementioned roles.

Natriuretic peptides

Natriuretic peptides have demonstrated significant utility in risk stratification and screening. In the general population, higher level of BNP are independently linked to all-cause mortality. The risk of death associated with a specific BNP level is comparable between those with HF and those without it (York et al., 2018). Individual participant data from 40 prospective cohort studies, which included data for 95,617 individuals without prior history of CVD, indicated that the top tertile of NT-proBNP concentration was associated with a threefold higher risk of incident HF (Willeit et al., 2016). Two RCTs including individuals with risk factors for HF development, STOP-HF (patients with at least one CVD factor like DM, dyslipidemia, and hypertension) and PONTIAC (patients with type 2 DM) confirmed that NT-proBNP-based screening lowers risk of HF hospitalization, newly diagnosed HF, LV systolic dysfunction and diastolic dysfunction (Huelsmann et al., 2013; Ledwidge et al., 2013). Moreover, in a secondary analysis of two other trials on patients with type 2 DM, only NT-proBNP showed meaningful improvement in HF prediction (Ohkuma et al., 2017; Sharma et al., 2020). Individuals with structural heart disease who are at high risk of progressing to clinical HF may have elevated levels of BNP or NT-proBNP, even if they are asymptomatic. Retrospective analysis of the Olmsted County Heart Function Study showed that NT-proBNP was associated with progression of pre-clinical to clinical HF (Young et al., 2021)

Natriuretic peptides are a cornerstone of HF diagnosis as they are included in the current universal definition of HF (Bozkurt et al., 2021). Major guidelines offer the highest level of recommendation, supported by the highest level of evidence, for the use of natriuretic peptide assessment in both ambulatory settings and among all acute patients presenting with symptoms suggestive of new-onset or worsening HF (Heidenreich et al., 2022; McDonagh et al., 2021). Both BNP and NT-proBNP are established markers for both short-term and long-term prognosis following a diagnosis of HF, including factors such as mortality and hospital admissions (Sharma et al., 2020; York et al., 2018; Zile et al., 2020).

While there is broad consensus on the value of natriuretic peptides for diagnosis and prognostic assessment in HF, the utility of BNP and NT-proBNP as guides for medical treatment remains a subject of debate. Pooled results of small RCTs conducted before 2010 indicated that

natriuretic peptide-guided therapy, when compared to the standard of care, demonstrated benefits in incorporating biomarker levels into treatment decision-making (Felker et al., 2009). These promising results were one of the motivations for the design of GUIDE-IT trial, which aimed to randomize 1100 high-risk patients with HFrEF either to standard therapy or to titrate HF therapy to target an NT-proBNP concentration of less than 1000 pg/mL. NT-proBNP-guided therapy did not improve either the primary or secondary study endpoint (Felker et al., 2017). However, this does not negate the potential utility of biomarker-guided therapy, as the intensification of medical treatments was similar in both arms, and the achieved NT-proBNP levels remained predictive of trial outcomes (Daubert et al., 2019). Additionally, recent trials involving new drug classes such as SGLT-2 inhibitors and angiotensin receptor-neprilysin inhibitors (ARNIs) have indicated a strong correlation between natriuretic peptides and favorable therapeutic outcomes (Cunningham et al., 2020; Januzzi et al., 2022).

Alternative biomarkers

Nevertheless, natriuretic peptides constitute a single class of biomarkers that do not capture the entire spectrum of HF pathophysiology. Over the past 30 years, substantial efforts have been made to identify alternative markers for HF. Recently, accumulating evidence for some of these markers has led to their inclusion in consensus papers and guidelines from relevant societies. These alternative biomarkers can be divided into five general categories: biomarkers of cardiac injury or damage, biomarkers of systemic or localized inflammation and oxidative stress, biomarkers of extracellular matrix remodeling (ECM) and fibrosis, biomarkers of other organ damage, and biomarkers of neurohormonal activation.

Notable examples of biomarkers cardiac injury in HF include high-sensitivity cardiac troponins (hs-cT) and heart-type fatty acid (FA)-binding protein (hFABP). In an AHF, an increase in hs-cT might be a consequence of acute myocardial injury, while in CHF it might reflect increased LV filling pressures and subendocardial ischemia from impaired myocardial oxygen supply-demand (Januzzi et al., 2012; Obokata et al., 2018). A previous meta-analysis of 11 studies, with 9289 patients, concluded that hs-cT was strongly and independently predicting all-cause and cardiovascular mortality, as well as hospitalization in patients with CHF (Aimo et al., 2018). hFABP is located in the sarcoplasm and therefore is actively secreted upon cardiomyocyte injury, contrasting with the passive, slower release mechanisms seen for hs-cT (Otaki et al., 2017).

Interest in the role of inflammation in HF pathogenesis predates the focus on natriuretic peptides. Building on this early interest, studies have explored systemic pro-inflammatory biomarkers such as C-reactive protein, procalcitonin, tumor necrosis factor-alpha, interleukin (IL)-1β, IL-6, IL-18, IL-33, myeloperoxidase, adiponectin, visfatin, and leptin (Reina-Couto et al., 2021; S. Zhao et al., 2021). Particular attention was recently attributed to growth differentiation factor 15, a member of the transforming growth factor (TGF)-β cytokine superfamily, regulating the inflammatory response and tissue repair (Wollert et al., 2017).

In the context of ECM and fibrosis biomarkers, focus has been placed on galectin-3, the soluble isoform of suppression of tumorigenicity 2 (sST2), and various molecules associated with ECM turnover. Galectin-3 is a β-galactoside-binding lectin involved in the regulation of myocardial and microvascular inflammation, the migration of mononuclear cells, and the extracellular build-up of collagen matrix. By promoting collagen overproduction and fibroblast proliferation, it contributes to detrimental changes in cardiac structure and function (Bošnjak et al., 2015). sST2 is a cardiomyocyte protein induced by mechanical stress that opposes the anti-hypertrophic effects of IL-33. Substantial evidence suggests that sST2 levels can serve as markers for cardiac stress and structural changes in HF (Reina-Couto et al., 2021).

When examining the neurohormonal axis in HF, key biomarkers include norepinephrine, adrenomedullin (ADM), copeptin, endothelin-1 and neprilysin (Ibrahim & Januzzi, 2018; Meijers et al., 2021). Activation of the sympathetic nervous system leads to the release of norepinephrine, which in turn results in peripheral vasoconstriction, increased myocardial contraction and accelerated heart rate (Ibrahim & Januzzi, 2018). The primary function of ADM is vasodilation, which leads to reduced blood pressure and enhanced vascular flow, while a secondary role involves maintaining endothelial integrity, the absence of which can cause vascular leakage and systemic and pulmonary edema (Tanaka et al., 2016). Due to its short half-life, a precursor mid regional propeptide is measured as an alternative. Copeptin is the C-terminal segment of the arginine vasopressin precursor peptide, and its concentration has been reported to increase in HF (Zhong et al., 2017). Endothelin-1 is synthesized by the endothelium in reaction to inflammation, neurohormonal activation, and vascular shear stress, and it subsequently acts to induce vasoconstriction, initiate proinflammatory pathways, stimulate cellular proliferation, and facilitate the generation of free radicals (Ibrahim & Januzzi, 2018). Lastly, neprilysin is responsible for breaking down natriuretic peptides as well as other

substances like angiotensin I, angiotensin II, bradykinin, endothelin-1, and substance P. It has recently garnered attention due to the introduction of ARNI (Meijers et al., 2021).

Other biomarkers pertinent to HF but primarily associated with different organs include renal indicators like cystatin C, kidney injury molecule-1, and neutrophil gelatinase-associated lipocalin, markers of micronutrient deficiencies such as iron, and metrics related to thyroid function like thyroid hormones (Meijers et al., 2021).

Metabolism and metabolites in heart failure

Leading medical societies in cardiovascular medicine have recently recognized the significance of discovering novel biomarkers as alternatives to natriuretic peptides. This heightened interest is particularly driven by the identification of individuals with a high baseline risk of HF who could potentially benefit from early therapeutic interventions. Additionally, monitoring the effectiveness of treatment becomes crucial to personalize therapy using a biomarker-guided approach.

Most of the research endeavors in the field of biomarker research for HF have predominantly concentrated on proteins. Nevertheless, small molecules, namely metabolites, have remained largely unexplored as a promising reservoir of potential HF biomarkers. Metabolites exhibit remarkable sensitivity to fluctuations in the cellular milieu and demonstrate swift responsiveness to both internal and external cues. Consequently, shifts in metabolite levels often showcase heightened sensitivity and more rapid reactions compared to corresponding alterations in protein levels. The significant metabolic disruptions that transpire within cardiac cells during the progression of HF hold the promise of serving as valuable indicators of the disease.

Metabolism in healthy heart

Due to the unceasing energy expenditure during cardiac contraction, the heart displays exceptionally high metabolic activity. The adult heart produces vast amounts of adenosine triphosphate (ATP) essential for maintaining its contractile function, primarily deriving this energy from two main sources: mitochondrial oxidative phosphorylation and glycolysis. Typically, around 95% of the myocardial ATP demands are met through mitochondrial oxidative phosphorylation, while glycolysis fulfills the remaining 5%. The predominant share

of ATP production within the mitochondria, approximately 40% to 60%, arises from the oxidation of FAs. The residual portion is sourced from the oxidation of pyruvate (derived from glucose and lactate), ketone bodies, and amino acids (Lopaschuk et al., 2021).

FAs reach the heart either as bound to albumin in the blood or released from triacylglycerols (TGs) in chylomicrons and very low-density lipoprotein (VLDL). Inside cardiomyocytes, they transform into fatty acyl coenzyme A (CoA) via esterification. Carnitine palmitoyl transferase 1 (CPT-1) transfers the FAs to carnitine, forming long-chain acylcarnitines. They enter mitochondria, where FAs transfer back to CoA, creating fatty acyl-CoA. Acetyl-CoA which is a product of further beta-oxidation, enters the tricarboxylic acid (TCA) cycle, along with nicotinamide adenine dinucleotide (NADH) and flavin adenine dinucleotide (FADH2) for ATP generation. The TCA cycle also contributes NADH and FADH2 to the electron transport chain, aiding ATP synthesis with oxygen. Regulation of FA utilization includes: FAs supply, FAs uptake, CPT-1 inhibition by malonyl CoA, FAD/FADH2 and NAD+/NADH ratios affecting enzyme activity, mitochondrial acetyl-CoA/CoA ratio influencing 3-ketoacyl-CoA thiolase, posttranslational enzyme modifications, and transcriptional regulation (Lopaschuk et al., 2021). Despite high ATP yield, FAs demand the most oxygen, making them the least efficient myocardial energy substrates.

Glucose oxidation contributes to 20-40% of energy in a healthy heart. It enters cardiomyocytes primarily through glucose transporter type 4 (GLUT4) and to a lesser extent through GLUT1. In the heart, GLUT4 is the predominant transporter responsible for myocardial glucose uptake, and its translocation is modulated not only by insulin but also by myocardial contraction. Once inside the cell, glucose is metabolized through various pathways, including glycolysis, which converts it to pyruvate or lactate. The intracellular glucose undergoes phosphorylation by hexokinase to form glucose 6-phosphate (G6P), which can be directed into multiple metabolic pathways. These include glycolysis for the generation of pyruvate, glycogen synthesis, and other accessory branches like the hexosamine biosynthetic pathway and the pentose phosphate pathway. Pyruvate, the end product of glycolysis, has two primary fates. It can either be converted to lactate or imported into the mitochondria via the mitochondrial pyruvate transporter complex (MPC). Once in the mitochondria, pyruvate is decarboxylated by the pyruvate dehydrogenase complex (PDH) to form acetyl-CoA, which then enters the TCA cycle for further metabolism. PDH activity is finely regulated through phosphorylation and dephosphorylation mechanisms, influenced by the ratios of acetyl-CoA/CoA and

NADH/NAD+. Lactate also serves as an important energy substrate for the heart, particularly when circulating lactate levels are elevated. It is taken up via monocarboxylate transporters (MCT) and converted back to pyruvate through lactate dehydrogenase (LDH). This pyruvate can then be metabolized similarly to the pyruvate derived from glycolysis (Lopaschuk et al., 2021).

Ketone bodies, predominantly 3-hydroxybutyrate (3-OHB), are increasingly recognized as a significant energy substrate for cardiac metabolism. Under normal conditions, ketone use contributes to 10-20% of total heart energy demand. These ketone bodies are primarily synthesized in the liver from acetyl-CoA, which is mainly sourced from FAs oxidation (Lopaschuk & Dyck, 2023). Interestingly, the potential for ketogenesis within the heart itself has also been proposed. The cellular uptake of 3-OHB is facilitated also by MCT. Once inside the cell, 3-OHB is transported into the mitochondria, where it undergoes a series of enzymatic catabolic reactions to yield acetyl-CoA, a crucial substrate for the TCA cycle. The initial step in this catabolic pathway involves the enzyme mitochondrial D-3-OHB dehydrogenase (BDH1), which catalyzes the oxidation of 3-OHB to acetoacetate. Subsequently, acetoacetate is activated by succinyl-CoA:3-ketoacid-CoA transferase (SCOT1) to form acetoacetyl-CoA. Finally, acetoacetyl-CoA undergoes a thiolysis reaction catalyzed by acetyl-CoA acetyltransferase 1 (ACAT1), resulting in the production of acetyl-CoA, which is then funneled into the TCA cycle. When considering the oxygen consumption required for ATP production, ketones offer a more efficient metabolic pathway compared to FAs, although they are less energy-efficient than glucose (Lopaschuk et al., 2021).

Oxidation of branched-chain amino acids (BCAAs – isoleucine, leucine, and valine) contributes in normal conditions to <1% of ATP production. The initial phase of BCAA metabolism in cardiac mitochondria is orchestrated by mitochondrial branched-chain amino-transaminase, which catalyzes the transamination of BCAAs to their corresponding branched-chain α-keto acids (BCKAs). This reversible reaction involves the transfer of the α-amino group to α-ketoglutarate, subsequently generating glutamate. Following this, BCKAs undergo oxidative decarboxylation facilitated by branched-chain α-keto acid dehydrogenase (BCKDH). The activity of BCKDH is finely tuned through phosphorylation by BCKDH kinase, which inhibits its activity, and dephosphorylation by protein phosphatase C2m, which activates it. The end products of this metabolic cascade are either acetyl-CoA, which feeds into the TCA cycle, or succinyl-CoA, which serves anaplerotic functions. Beyond their metabolic

contributions, BCAAs also modulate critical signaling pathways in the heart, such as insulin and mammalian target of rapamycin (mTOR) signaling (Lopaschuk et al., 2021).

The heart is often referred to as a "metabolic omnivore" due to its remarkable metabolic flexibility, allowing it to seamlessly transition between various energy substrates to sustain the production of ATP (Bertero & Maack, 2018). Increased workflow induces a shift from FA to glucose oxidation. On the other hand, during prolonged fasting or DM, there is a preference for ketone use. The flux of substrates through oxidative pathways is mainly determined by the catalytic activity of rate-limiting enzymes, the number of enzyme molecules available inside the cell, as well as epigenetic and post-translational modifications (Bertero & Maack, 2018; Ritterhoff & Tian, 2023).

Metabolism in failing heart

Two metabolic hallmarks of HF encompass: (1) imbalance between increased energy demand and decreased energy supply, and (2) ‘metabolic rewiring’ that triggers a transition in the utilization of energy substrates (Bertero & Maack, 2018; Ritterhoff & Tian, 2023). During the progression of HF, the energy status of the heart can decline to as low as ~60% of the energy status in a healthy heart. This shift in the utilization of energy substrates includes a lower relative contribution from FA oxidation, coupled with an increased relative utilization of glucose, ketone bodies, lactate, and BCAAs.

The decline in FA oxidation rates HF advances is a subject of considerable research focus. In both animal models and human studies, this metabolic shift is often observed and is believed to be mediated by complex intracellular signaling pathways. One key player is the peroxisome proliferator-activated receptor (PPAR) α, a transcription factor integral for upregulating enzymes involved in FA oxidation. When PPARα signaling is suppressed, as seen in severe stages of HF, FA oxidation decreases substantially. Another potentially influential factor is the hypoxia-inducible factor 1α (HIF1α)–PPARγ signaling axis. This pathway may further downregulate FAs oxidative enzymes and impair FA transport into mitochondria. However, the decline in FA oxidation is not accompanied by a proportional decrease in FA uptake. This creates a significant metabolic mismatch, leading to the intracellular accumulation of lipids. These accumulated lipids are not only stored as TGs but also funneled into non-oxidative pathways. This diversion generates toxic lipid species like ceramides and diacylglycerols, which are implicated in mitochondrial dysfunction and apoptosis. Such lipotoxicity potentially speeds up the progression of HF. Additionally, the accumulation of FAs disrupts insulin

signaling pathways, contributing to insulin resistance. This resistance can further exacerbate HF by impairing glucose uptake, another key source of myocardial energy, especially during stress conditions. Studies have shown that mechanical unloading of the failing heart by LV assist devices can reduce levels of toxic lipid species and restore insulin signaling, suggesting that the metabolic dysregulation is reversible under specific conditions (Bertero & Maack, 2018).

Both clinical and animal models consistently demonstrate increased glucose uptake in the failing heart. However, this uptick in glucose uptake often does not correlate with increased glucose oxidation. Specifically, in idiopathic dilated cardiomyopathy, canine models, and rodent models involving a high-salt diet or pressure overload, glucose uptake surges alongside an upregulation in GLUT1 expression, suggesting increased capacity for glucose assimilation (Lopaschuk et al., 2021). Despite this, the oxidation of pyruvate, the end-product of glycolysis, usually declines, indicating a disconnect between uptake and efficient metabolic utilization. This metabolic incongruity in the failing heart (increased glycolysis but diminished glucose oxidation) has been observed across various scenarios including obesity-associated HF and DM. The inefficiency stems partly from decreased expression of metabolic transporters and enzymes such as MCT1, PDH, and MPCs. Concurrent with these alterations is an increase in anaplerotic reactions—alternate metabolic pathways that replenish TCA cycle intermediates. One such mechanism is the carboxylation of pyruvate to malate via malic enzyme, which is energetically less efficient than pyruvate oxidation and increases the load on the TCA cycle. This metabolic adaptation is significant because it points to a state of compromised redox homeostasis and oxidative stress, which further impairs cardiac function (Lopaschuk et al., 2021).

Evidence suggests that ketone body oxidation is heightened in conditions like HFrEF and diabetic cardiomyopathy, with increased levels of blood ketone bodies reflecting augmented oxidation. This phenomenon appears to be a compensatory mechanism aimed at sustaining energy production in the face of reduced utilization of other substrates. Notably, genes encoding enzymes linked to ketone body oxidation show heightened expression in both patients and animal models with HFrEF or DM. Additionally, changes in enzyme activity, like the increased activity state of the rate-limiting enzyme SCOT1, further indicate enhanced ketone oxidation in failing hearts. This metabolic shift towards ketone body utilization might stem from its potential advantages, such as a higher ATP production: oxygen consumption ratio

compared to FAs. This has led to the notion of 3-OHB as a potentially advantageous "super fuel." However, whether this adaptation signifies a beneficial response or a maladaptive one remains uncertain. While there is evidence of increased ketone body oxidation in HFrEF, the situation is more intricate in HFpEF. Unlike HFrEF, HFpEF models have demonstrated mixed results regarding ketone oxidation. Hyperketonemia seen in conditions like insulin resistance and type 2 DM might also contribute to the metabolic landscape of HFpEF. Nevertheless, the exact role of ketone bodies in HFpEF development requires further investigation (Bertero & Maack, 2018; Lopaschuk et al., 2021; Ritterhoff & Tian, 2023).

In patients with HF, there is a noticeable increase in plasma BCAA levels, which could serve as an early predictor of future CVD. The rise in BCAA levels might partly result from impaired BCAA oxidation in HF. The regulation of BCAA catabolism is controlled by Krüppel-like factor 15 (KPF-15)-mediated transcriptional mechanisms. This process is downregulated during the development of cardiac hypertrophy and failure, as part of metabolic remodeling. While controversy surrounds the reason for the accumulation of BCAAs and BCKAs in the failing heart, genetic impairment of BCAA oxidation in animal models suggests that faulty catabolism contributes to HF development. Interestingly, BCAA oxidation only minimally contributes to ATP production in the heart, implying that its impact on HF pathogenesis is likely through signaling mechanisms (Bertero & Maack, 2018; Lopaschuk et al., 2021; Ritterhoff & Tian, 2023).

Other metabolic pathways in heart failure

Except for cardiac metabolism, also other metabolic pathway seems to be important in the context of HF pathophysiology. Nitric oxide (NO) is synthesized by NO synthases (NOS) from L-arginine, and exists in three primary isoforms: endothelial, neuronal, and inducible. Endothelial NOS is pivotal in regulating vascular tone through the activation of cyclic guanosine monophosphate and protein kinase G, leading to smooth muscle relaxation and vasodilation. In HF, however, there is a reduced vascular NO bioavailability, resulting in compromised vasodilation. This adds to the cluster of risk factors like hypertension and atherosclerosis that exacerbate HF. The availability of L-arginine, a substrate for NOS, and other modulators such as tetrahydrobiopterin influence NO production. Additionally, endogenous inhibitors like asymmetric/symmetric dimethylarginine (ADMA/SDMA) and N-monomethyl L-arginine (L-NMMA) further regulate this pathway. ADMA and L-NMMA, whose levels are controlled by the enzyme DDAH1, can inhibit all forms of NOS and thereby

attenuate NO production. Both ADMA/SDMA, NMMA were previously linked with a higher risk of incident HF (Goni et al., 2022).

Another pathway considered in HF pathogenesis is choline metabolism by gut microbiota. Dysregulation of gut microbiota has been linked to elevated levels of circulating trimethylamine N-oxide (TMAO). In heart cells, TMAO triggers the TGF-β1/Smad3 and p65 NF-κB signaling cascades, leading to reduced energy metabolism and compromised mitochondrial function. TMAO further impairs the heart's ability to contract effectively and hampers the regulation of intracellular calcium. In the cells lining blood vessels, TMAO activates NLRP3 inflammasomes, resulting in the release of pro-inflammatory cytokines IL-1β and IL-18. Additionally, TMAO exacerbates endothelial dysfunction through the activation of the PKC/NF-κB/VCAM-1 signaling pathway (Goni et al., 2022).

Metabolomics

As summarized in previous chapters, the dysregulation across specific metabolic pathways during HF can result in elevated circulating levels of various classes of metabolites, offering potential candidates for HF biomarkers. Traditionally, in the realm of biochemical research, specific molecules were selected based on hypotheses surrounding disease pathophysiology. This reductionist, deductive approach was time-consuming, necessitating the generation of initial hypotheses and the measurement of a single specific metabolite. This process had to be repeated for each metabolite of interest. Furthermore, this methodology overlooked the potential for unveiling unknown pathways and previously unrecognized interconnections between compounds and pathways.

With over 37,000 metabolites identified in human blood according to the Human Metabolome Database, traditional hypothesis-based experiments often fall short in unlocking the full exploratory potential of metabolomic studies (Wishart et al., 2022). Suffix "omics" refers to a suite of advanced, comprehensive analytical approaches in molecular biology used to study a wide range of biological traits such as genes, mRNAs, proteins, or metabolites. In the context of metabolites – metabolomics is a comprehensive analysis aiming to quantify all the metabolites of a biological using high-throughput methods. A complete set of metabolites in a given biological compartment (biospecimen) in given physiological conditions will be defined as a metabolome.

Two main approaches in metabolomic studies include non-targeted and targeted metabolomics. Non-targeted metabolomics is a broad analytic strategy designed to detect, identify, and relatively quantify as many metabolites as possible in a biological sample. The objective is to isolate unique metabolite features linked with clinical features. However, key challenges in nontargeted metabolomics include the need for established or in-house metabolite databases, the absence of reliable fragmentation chemistries for all metabolites of interest, and the chemical/biochemical modification of metabolites that result in spectral shifts. Despite these hurdles, the extensive coverage afforded by non-targeted approaches offers the potential to discover new biomarkers. Conversely, targeted metabolomics focuses on the precise measurement of a selected set of metabolites, employing customized analytical methods for absolute quantification. Such quantification is commonly attained using stable isotope-labeled internal standards or by generating calibration curves via external standards. This stands in distinction to nontargeted metabolomics, where the method of choice for representing metabolites is usually relative quantification, often based on comparing the peak areas of the metabolites under investigation (McGarrah et al., 2018).

High-throughput techniques used for metabolomics can be divided into mass spectrometry (MS)-based and nuclear magnetic resonance (NMR). In MS, the ionization source transforms analytes into ions, the intensity of which is measured and plotted according to their mass-to-charge ratios (m/z). These spectra of m/z ratios, along with fragmentation data, facilitate the identification of metabolites. Usually, a front-end separation step such as gas (GC) or liquid chromatography (LC) precedes the mass spectrometer, enhancing selectivity and mitigating ion suppression by limiting the number of compounds that enter the mass spectrometer simultaneously. GC is commonly employed for the profiling of organic/amino acids, as well as both non-esterified and esterified FAs. LC can separate a wide array of metabolites, making it ideal for comprehensive, high-throughput metabolomic studies of bodily fluids and tissues. The common stationary phases used include reverse phase for nonpolar metabolites like glycerolipids, phospholipids, FAs, acylcarnitines, and medium/long chain acyl-CoAs; hydrophilic interaction chromatography for polar metabolites such as sugar monomers, nucleotides/nucleosides, phosphate compounds, and organic acids; and porous graphitized carbon for very polar molecules like glycans, oligosaccharides, and sugar phosphates. An advantage of LC is no necessity to volatize the sample before analysis but, on the other hand, is less reproducible (McGarrah et al., 2018).

The principle of NMR exploits the inherent magnetic properties of specific atomic nuclei, attributed to their nuclear spin. When placed in an external magnetic field, these nuclei align along the field. By applying radio frequency pulses, the system's magnetic alignment is disturbed. As the nuclei relax back to their equilibrium state, they emit electromagnetic radiation. The emitted electromagnetic radiation during the relaxation phase of these nuclei is subsequently captured. The frequency of this released energy, known as the chemical shift, is influenced by both the magnetic field's strength and the specific compound characteristics. Signal splitting in NMR spectra occurs due to spin coupling, as neighboring nuclei impact the resonance frequency of a particular nucleus. ^{1}H-NMR stands out as the most widely used NMR method, credited to its rapid acquisition times, the high sensitivity of ^{1}H-nuclei, and the prevalence of ^{1}H in biological entities. Compared to MS, NMR has the advantage of being non-destructive and noninvasive, enabling in vivo studies in humans and animals as well as in situ investigations in perfused tissues. However, its sensitivity lags significantly behind MS (µmol/L vs. pmol/L), and it offers less comprehensive metabolome coverage (covering hundreds vs. thousands of metabolites). Financially, the initial and operational expenses for NMR systems are considerably higher than for GC-MS and are comparable to or exceed those for top-tier LC-MS/MS setups. Another drawback of NMR is the larger sample size requirement. While MS can work with µL volumes of plasma or mere milligrams of tissue, NMR generally demands more substantial sample sizes, posing challenges, particularly in rodent studies where the available tissue and plasma are limited (McGarrah et al., 2018).

Metabolomics in molecular epidemiology of chronic diseases

Over the past decade, metabolomics has successfully transitioned into observational studies, enabling the exploration of the molecular underpinnings of disease, investigating the relationships between environmental factors and disease outcomes, and facilitating the discovery of biomarkers. Early observational studies that employed metabolomics in the context of chronic diseases often had limitations such as small sample sizes, retrospective design, the use of in-house analytical platforms (i.e., limiting comparability and reproducibility of results), and a lack of external validation. Over time, metabolite data became available for prospective cohort studies, initially in the form of nested case-control and case-cohort designs, and later extending even to cohorts with repeated measurements. Simultaneously, several commercial companies began offering externally validated extensive metabolomic assays, thereby enhancing the standardization and reproducibility of results.

Metabolomics has been extensively researched in relation to various outcomes, including all-cause mortality, CAD, stroke or type 2 DM (Balasubramanian et al., 2020; Morze et al., 2022; Ruiz-Canela et al., 2017). For instance, a systematic review current as of 2015 identified 19 prospective studies reporting associations between metabolites and type 2 DM (Guasch-Ferré et al., 2016). An updated version of this review in 2022 revealed 61 prospective studies on the same topic, representing more than a threefold increase in 7 years (Morze et al., 2022). However, while research on other common non-communicable disease endpoints has been expanding across various cohorts, the focus on HF appears to be notably lacking.

Retrospective studies on metabolomics and heart failure

The literature review pertaining to retrospective studies exploring associations between specific metabolites and HF confronts several methodological challenges. Firstly, it proves challenging to categorically determine whether a given study can be classified as a metabolic analysis. This is predominantly due to the utilization of high-throughput platforms that analyze a restricted set of metabolites, either based on a priori decisions or due to their linkage with a singular metabolic pathway. Secondly, some investigations employ a comparative analysis across diverse types or etiologies of HF or assess the ramifications of specific interventions on the metabolomic profile of HF patients in contrast to healthy controls. To navigate these complexities, inclusion criteria were established to consider studies that evaluate metabolites from more than two distinct pathways, or a panel assessing greater than 20 metabolites (Ruiz-Canela et al., 2017). Studies were incorporated into the literature review only if they facilitated a comparative analysis of metabolites between HF subjects and control groups.

Several retrospective case-control studies have used different chromatography methods combined with MS to compare the metabolomic profiles of HF patients and healthy controls. The first to author's best knowledge study was conducted by Dunn and colleagues. They collected serum samples from 52 patients diagnosed with systolic HF, characterized by LVEF <40% along with overt signs and symptoms of failure. These were compared to samples from 57 control subjects. GC-MS was employed for analytical profiling, resulting in 272 statistically robust metabolite peaks, including 38 significant after multiple-testing correction. Notably, two of such metabolites were pseudouridine, a modified nucleotide found in t- and rRNA that serves as a marker for cell turnover, and 2-oxoglutarate, an intermediate in the TCA cycle. In addition to these, the study also identified three new compounds – 2-hydroxy-2-methylpropanoic acid,

erythritol, and 2,4,6-trihydroxypyrimidine – as excellent discriminators between the patients and the control group (Dunn et al., 2007).

The majority of subsequent studies included small sample sizes, utilized in-house metabolomic platforms, and in most cases lacked external validation. The findings consistently showed higher levels of acylcarnitines, particularly those with short and long fatty acyl chains (Chen, Liu, et al., 2020; Chen et al., 2023; Liu et al., 2020; Ruiz et al., 2017; Zordoky et al., 2015), and lower levels of various phosphatidylcholines (PC) and lysophosphatidylcholines in HF patients compared to controls (Lindahl et al., 2016; Marcinkiewicz-Siemion et al., 2018; Mueller-Hennessen, Sigl, et al., 2017; Zhao et al., 2020; Zordoky et al., 2015). There were also signs that HF could be significantly linked to the concentration of ceramides (especially those with a single monounsaturated bond) and sphingomyelins (SM) (Deng et al., 2014; Mueller-Hennessen, Sigl, et al., 2017; Yang et al., 2021), several amino acids (increased levels of phenylalanine and glutamate, and decreased levels of histidine) (Liu et al., 2020; Ruiz et al., 2017; Saleem et al., 2020; Zordoky et al., 2015), ketone and energy-related metabolites – β-OHB, citrate, and pyruvate (Chang et al., 2020; Ruiz et al., 2017; Yang et al., 2021), as well as gut-microbiota related metabolites (Cui et al., 2018).

There were a few notable large, comprehensive case-control studies. Cheng and colleagues reported a case-control study of 515 participants, which included 114 normal control subjects and 401 HF patients. It was conducted in two phases: a discovery phase with 234 participants (51 normal control subjects and 183 HF patients classified into stages A, B, and C) and a validation phase involving an independent set of 63 normal control subjects and 218 stage C HF patients. In the discovery phase, an untargeted metabolomics analysis was performed on 119 plasma samples, specifically from 51 normal control subjects and 68 stage C HF patients, using ultra-performance LC coupled with time-of-flight MS. Subsequently, a targeted analysis of 184 metabolites was conducted using LC-MS. Metabolites such as amino acids, biogenic amines, and phospholipids were observed to differentiate between HF stages A, B, and C. Plasma metabolites related to arginine metabolism, like glutamine and citrulline, were reduced in stage C HF patients, while glutamate, ornithine, spermine, spermidine, tyrosine, and phenylalanine levels were elevated. Receiver-operating characteristic curves revealed that the areas under the curves for BNP and multi-metabolite score (including four metabolites histidine, phenylalanine, spermidine, and PC C34:4) were 0.99 and 1, respectively (Cheng et al., 2015).

Hunter and colleagues analyzed data from the CATHGEN repository for 282 patients with HFpEF cases, 279 patients with HFrEF controls, and 191 controls. Plasma samples from these participants were analyzed using tandem flow-injection targeted MS to quantify a panel of 60 metabolites – comprising 45 acylcarnitines and 15 amino acids. An additional three metabolites (ketones, non+esterified FAs, and 3-hydroxybutyrate) were measured by conventional methods. Principal Component Analysis was employed to derive 14 metabolite scores. Comparative analysis revealed that HFrEF cases exhibited an increase in the scores for long-chain acylcarnitines, BCAA and related metabolites, alanine, proline, and free FAs. Conversely, they showed a decrease in scores for histidine, arginine, tiglylcarnitine, 3-hydroxylinoleyl/hexadecadienedioyl carnitine, valine, glutamine, and glutamate. In HFpEF cases, the long-chain acylcarnitine score was increased, while scores for histidine, arginine, tigylcarnitine, and 3-hydroxylinoleyl/hexadecadienedioyl carnitine were decreased (Hunter et al., 2016).

Another study recruited 887 individuals including HFrEF due to either ischemic (n = 257) or non-ischemic cardiomyopathy (n = 269) and healthy controls (n = 327), and a cohort of patients with pulmonary diseases (n = 34). The study was conducted in two phases: initially, a single-center identification study involved 238 participants, followed by a multicenter confirmation study with 649 individuals. Metabolite profiling was performed on human plasma samples using untargeted GC-MS and LC-MS/MS techniques. A differential distribution of 92 metabolites between HFrEF patients and controls was identified. Amino acids like cysteine and glutamate were predominantly upregulated, while a complex pattern of up and downregulation was noted for lipids and FAs. In a principal component analysis of the metabolomic data, HFrEF patients were observed to have lower scores on the second principal component compared to healthy controls. A cardiac lipid panel (CLP) was subsequently established, consisting of three specific metabolomic features: TG 18:1/18:0/18:0, PC 16:0/18:2, and the sum of the 3 isobaric SM species. This CLP was combined with NT-proBNP levels in a logistic regression model, enabling effective discrimination between HFrEF patients and healthy controls (Mueller-Hennessen, et al., 2017).

Wang and colleagues studied 1288 participants who underwent LC-MS metabolite profiling. This group included 129 normal controls and 712 patients with stages A to D of HF. In the validation cohort, there were 447 patients with stage C HF. In patients with stage C HF, the amino acids ornithine, phenylalanine, aspartate, glutamate, tyrosine, isoleucine, tryptophan,

and lysine were elevated compared to normal controls. Conversely, the amino acids histidine, glutamine, citrulline, serine, and arginine were found to be decreased. Three metabolites - histidine, ornithine, and phenylalanine - were notably significant in identifying stage C HF patients and were used to create the multi-metabolite score. That score outperformed BNP in distinguishing between patients at various stages, most notably in those who were elderly, had AF, high body mass index (BMI), or kidney dysfunction. Moreover, the score showed a better correlation with the 6-minute walking distance than did BNP (Wang et al., 2018).

Lastly, a Chinese study enrolled a total of 628 inpatients, consisting of 182 non-HF control subjects, 176 patients with HFpEF, 119 patients with HFmrEF, and 151 patients with HFrEF. Utilizing a targeted metabolomics approach, 102 metabolites were quantified in serum samples via LC-MS/MS. This panel included 23 amino acids, 35 acylcarnitines, and 44 metabolite ratios. Out of these, 63 metabolites were found to significantly differentiate controls from HF patients. Employing ordinal logistic regression models, a diagnostic metabolic score (DMS) was developed based on a panel of 18 metabolites, which clearly distinguished cases and controls. The panel consisted of specific amino acids and their ratios (valine/phenylalanine, phenylalanine, citrate), short-chain acylcarnitines (C2, C3), medium-chain acylcarnitines (C6, C18), and other carnitine derivatives and ratios (C6DC, C4-OH, C3/C16, C2/C0). Additional metabolite ratios such as C5-OH/C8, C5-OH/C0, C3DC/C4-OH, C5DC/C5-OH, C5DC/C16, and Total AC-DC/Total AC, along with the phenylalanine/tyrosine ratio, were also included in the DMS, providing a comprehensive metabolic profile for HF subtyping (Zhao et al., 2021).

A limited number of studies have explored NMR metabolomics in the context of HF. An Italian study involving 185 HF patients and 111 healthy controls found that specific metabolites showed directional changes that effectively discriminated between the two groups: phenylalanine, tyrosine, isoleucine, and creatine increased, while lactate, citrate, and L-dopamine decreased (Tenori et al., 2013). A study involving 46 patients with HF and 15 age-matched controls revealed that 3-OHB, acetone, and succinate were significantly associated with elevated myocardial energy expenditure (Du et al., 2014). Another study that enrolled 9 control subjects, 9 HF patients with mild-to-moderate LVEF impairment, and 15 HF patients with severe LVEF impairment suggested that 2-hydroxybutyrate, glycine, methylmalonate, and myo-inositol were the most effective metabolites for stratifying the considered subgroups (Deidda et al., 2015).

Prospective studies on metabolomics and heart failure

In molecular epidemiology, population-based cohort studies are favored over case-control studies. These designs mitigate selection, recall, and misclassification biases, enabling the study of multiple outcomes and disease mechanisms. Most importantly cohort studies are prospective in nature, allowing for the establishment of a temporal sequence between exposure and outcome. This strengthens the causality, which is particularly crucial when investigating the molecular mechanisms underlying disease etiology. To date, there is a limited number of prospective cohort studies which adopted high-throughput metabolomics in order to identify metabolites significantly associated with risk of HF during follow-up.

The first such study was published by Zheng and colleagues. They analyzed data from 1,744 African Americans aged 45-64 years in the Atherosclerosis Risk in Communities (ARIC) study. The participants were free of HF at the beginning of the study, and they were followed for a median of 20 years (from 1987 to 2008). Using untargeted methods, the researchers measured 204 stable serum metabolites and used Cox proportional hazards (PH) regression to evaluate their associations with incident HF hospitalizations (n = 276). After adjusting for established risk factors, 16 metabolites were found to be significantly associated with HF. Among these, six named metabolites were primarily involved in amino acid metabolism, with one being a dipeptide (prolylhydroxyproline) and another a sugar alcohol (erythritol). Further adjustment for kidney function left two metabolites (X-11308 and X-11787) significantly associated with HF, with hazard ratios (HRs) of 0.75 and 1.23, respectively. These two metabolites were later identified as dihydroxydocosatrienoic acid and an isoform of either hydroxyleucine or hydroxyisoleucine.

Stenemo and colleagues performed MS-based metabolomic profiling in plasma or serum samples obtained from three community-based cohorts free of HF at the baseline. The study involved a total of 3,924 participants and yielded 341 incident HF events during a median follow-up period ranging from 4.6 to 13.9 years. Two discovery cohorts were used: the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) with 920 participants and the Uppsala Longitudinal Study of Adult Men (ULSAM) with 1,121 participants. An independent replication cohort, TwinGene, was also employed, consisting of 1,797 participants. In these cohorts, 74 participants from PIVUS, 183 from ULSAM, and 84 from TwinGene were eventually diagnosed with HF. Cox PH models were used to evaluate the association of each of the 206 identified metabolites with incident HF. The study found that

higher circulating levels of urobilin and lower levels of SM 30:1 were consistently associated with incident HF, both in the discovery and replication cohorts (Stenemo et al., 2019).

In the Jackson Heart Study involving 2,199 black individuals, targeted LC-MS was used to analyze 312 metabolites. Over a mean follow-up period of 9.6 years, 188 participants developed HF. A multivariate model identified 14 metabolites associated with the incidence of HF hospitalization. Of these, three metabolites – homoarginine, 2-aminoadipic acid, and uridine – were associated with a decreased risk. Conversely, eleven metabolites were associated with an increased risk: pseudouridine, 2-deoxyuridine, methylhistidine, 4-acetamidobutanoate, choline, ectoine, N-acetylspermidine, diacetylspermine, dimethylguanidinovaleric acid (DMGV), N-acetyl-l-alanine, and N2-N2 dimethylguanosine. These associations for choline, uridine, and 2-aminoadipic acid were subsequently replicated in a validation study using the FHS cohort. An exploratory analysis using competing risk models further revealed that eight metabolites were differentially associated with HFpEF (Orotic acid, C18-0-LPE_A, ATP, C16-1-LPC-plasmologen, C18-1-LPC, N-methylproline, C18-1-LPC-plasmologen, and C16-0-LPE), and three metabolites were specifically associated with HFrEF. These findings provide valuable insights into the metabolic pathways potentially involved in the development of HF and its subtypes, enriching our understanding of the disease's pathophysiology (Tahir et al., 2021).

Recently, several pathway-focused studies have reported metabolites associated with incident HF in the Prevención con Dieta Mediterránea (PREDIMED) trial (Bulló et al., 2021; Goni et al., 2022; Papandreou et al., 2021; Papandreou et al., 2020; Razquin et al., 2021; Ruiz-Canela et al., 2022; Wittenbecher et al., 2021). A nested case-control included approximately (depending on the report) 320 incident HF cases and 420 controls. Metabolites were quantified using targeted ultra-high performance LC-MS/MS platform. Elevated levels of medium-chain and long-chain acylcarnitines were associated with increased HF risk (Ruiz-Canela et al., 2022). In the TCA-related metabolite class, higher concentrations of malate, aconitate, isocitrate, and D/L-2-hydroxyglutarate were associated with a higher HF risk (Bulló et al., 2021). Among arginine-related metabolites, higher concentrations of arginine and homoarginine were negatively associated with the HF risk, while citrulline, asymmetric/symmetric dimethylarginine, N1-acetylspermidine, and diacetylspermine were positively associated with a higher HF risk (Goni et al., 2022). In the tryptophan–kynurenine degradation pathway, the kynurenine/tryptophan ratio, kynurenic acid, and quinolinic acid

were positively associated with an increased HF risk (Razquin et al., 2021). Higher concertation of choline and its pathway-related dimethylglycine were linked with a higher HF risk (Papandreou et al., 2021). Moreover, a higher concentration of glutamate was associated with a higher risk, while glutamine-to-glutamate ratio with a lower HF risk (Papandreou et al., 2020). Regarding lipid-related markers, palmitate, palmitoleate and stearate-containing lipids were linked to a higher HF risk (Wittenbecher et al., 2021).

Two prospective cohort studies investigated NMR metabolomics with incident HF. Delles and colleagues used data from the PROspective Study of Pravastatin in the Elderly at Risk (PROSPER) examine metabolic markers associated with HF hospitalization. Utilizing Nightingale ^{1}H NMR spectroscopy, 80 metabolic measures were analyzed in 5,341 participants from PROSPER, yielding 182 HF hospitalizations a 2.7-year follow-up. In unadjusted analysis, four measures, namely phenylalanine, creatinine, 3-OHB, and glycoprotein acetyls, were higher among individuals with HF than those without HF during follow-up. Conversely, nine lipoprotein-related measures, most notably those tied to high-density lipoprotein (HDL), lower in individuals with HF. Replicating the work in FINRISK 1997 cohort, involving 7,330 participants and 133 HF events over a 5-year period, 24 out of 85 metabolic measures were significantly associated with HF risk. Six measures, including 3-OHB, ratios of specific FAs and HDL2-cholesterol, were inversely related to HF, while 18 others (glucose, pyruvate, lactate, tyrosine, isoleucine, phenylalanine, alanine, glycoprotein acetyls and HDL triglycerides, VLDL and triglycerides) were positively related to HF risk. Notably, phenylalanine emerged as the strongest predictor of higher HF risk in both cohorts (Delles et al., 2018). In another Finnish study analyzing data on the from the Metabolic Syndrome in Men cohort including 10,197 men, authors examined association of clinical variables and 27 NMR metabolites measured using the same Nightingale platform with incident HF. After a median of 4.4 years, 172 men developed incident HF. Of the metabolites assessed, acetoacetate and pyruvate remained significantly associated with incident HF even after substantial adjustment for covariates (Jauhiainen et al., 2021).

2. Hypotheses and study scope

Rationale:

HF is a complex clinical syndrome with a multifactorial etiology and significant morbidity and mortality rates. Recent advancements in molecular biology have enabled the investigation of metabolic processes at an unprecedented level of detail. Metabolomics, the systematic study of the unique chemical fingerprints that specific cellular processes leave behind, may offer new insights into the underlying mechanisms of HF. As presented in the literature review in the previous section, prospective studies evaluating changes in circulating metabolites in the context of incident HF are limited, particularly regarding NMR metabolomics. The most significant limitations of previous studies include small cohorts with a low number of cases, which do not have sufficient statistical power to detect significant biological signals, as well as the inability to appropriately address residual confounding issues. This clearly identifies a research gap for a large-scale, prospective cohort study with validated metabolomic assays and comprehensive clinical and phenotypical data.

Hypothesis

This study relies on the hypothesis that metabolic alterations in specific pathways (e.g., FA metabolism, glycolysis, and amino acid metabolism) are significantly associated with the onset and progression of HF. It is hypothesized that these metabolic changes can be identified through metabolomic analysis and correlated with longitudinal HF risk. Additionally, these metabolite measures involved in HF contribution can be captured by subclinical changes in LV function at the stage preceding the diagnosis of HF. Lastly, these metabolite measures, as linked with HF pathogenesis, are associated with survival in patients with established HF as an indicator of prognosis.

Research Aims

The primary aim (Research Aim 1) of this study was to identify metabolic pathways independently associated with the development of HF. Specific steps included:

- (**Research Aim 1a**) To evaluate associations between plasma metabolite measures and incident HF during longitudinal follow-up.

- (**Research Aim 1b**) To derive metabolomic signatures associated with HF risk representing joint metabolic risk in the pathogenesis of the condition.

Secondary aims included:

- (**Research Aim 2**) To evaluate the association of plasma metabolite measures and HF metabolomic signatures with LV function, as assessed through magnetic resonance imaging.

- (**Research Aim 3**) To evaluate the association of plasma metabolite measures and HF metabolomic signatures with survival among participants with HF.

3. Data sources and methods

UK Biobank Cohort

UK Biobank is large-scale prospective cohort study in the United Kingdom designed to explore the genetic and non-genetic factors influencing diseases prevalent in middle to old age. With an in-depth evaluation of various exposures and rigorous monitoring of diverse health-related outcomes, the initiative seeks to bolster innovative scientific research through widespread availability of its data and resources.

Overview

UK Biobank aimed to recruit approximately 500,000 participants from a general UK population aged 40-69 without restriction to their health status between 2006 and 2010.

The full rationale and design of UK Biobank was published previously (Sudlow et al., 2015; UK Biobank, 2007). Potential participants were identified through centrally coordinated processes, drawing predominantly from population-based registers such as those held by the National Health Service (NHS). Potential candidates aged 40-69, living within a reasonable travelling distance of an assessment center, were dispatched an invitation package typically 6-8 weeks ahead of a provisional appointment. Recipients were prompted to confirm their appointment within a fortnight of receipt, either by telephoning the provided toll-free service, or by mailing the enclosed reply form or through the study website. For those who confirmed, a written affirmation of their assessment center slot was provided, accompanied by preparatory advice for their visit. In cases of non-participation, feedback on reasons for declining was actively pursued. Approximately 9.2 million invitations were sent letters and 502,682 participants were recruited via 23 centers across England, Wales and Scotland.

During the initial visit, each participant signed an electronic consent, completed a detailed touch-screen questionnaire; underwent a nurse-led computer-assisted interview, several physical and functional measures and collection of blood, urine, and saliva.

Except for the initial visit selected participants have contributed further by participating in additional data collection. These have involved: recording of physical activity over a 7-day period using accelerometers and online surveys addressing a range of subjects such as dietary factor, mental acuity, work history, psychological well-being, digestive health, and pain. From August 2012 to June 2013, a comprehensive reassessment was conducted for 20,000

participants at the UK Biobank Co-ordinating Centre in Stockport, UK. Since 2014, an ongoing imaging study has been underway, aiming to collect data on brain, cardiac, and abdominal MRI, dual-energy X-ray absorptiometry, and carotid ultrasound for 100,000 participants. Furthermore, in 2019, a follow-up imaging assessment was initiated.

Data access

Data from the UK Biobank can be accessed and used by external collaborators after submitting a research proposal for approval and upon payment of an access fee. The data used for this study were obtained under project no. 75001, titled "*Association of Metabolomic Profiles with MRI Traits of Cardiac Structure and Function and with the Risk of Incident Heart Failure in the UK Biobank Cohort*". This project was approved on September 23rd, 2021, for a duration of three years. Thesis author is a principal investigator in this project.

Ethical approval

The UK Biobank study received approval from the NHS North West Multicenter Research Ethics Committee (approval number: 11/NW/0382) and had obtained informed consent from all participants. The ethical approval was granted 2011 and renewed subsequently in 2016 and 2021 as scheduled for every 5-year renewal. Under this approval, separate ethical clearance for the purpose of this study was not required.

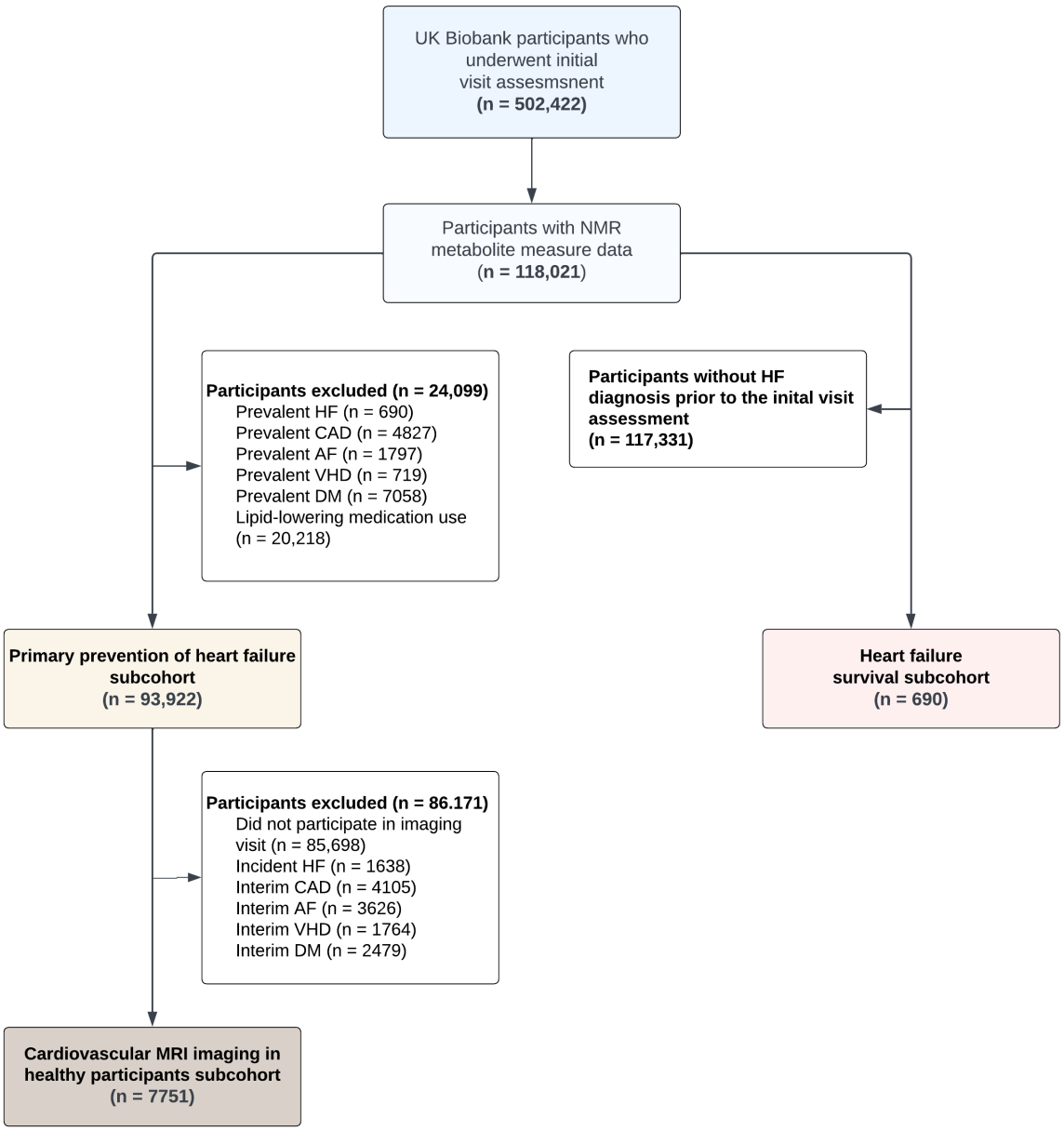
Eligibility criteria for study subcohorts

In line with research aims of this study, three subcohorts were defined. Study participants selection is presented in **Figure 1**. At the point of data access, the total cohort comprised of 502,422 participants. This differs from the above-mentioned number of recruited participants, due to deletion of participants who withdrew their consent for participation in UK Biobank. Missing covariate data were in general not an exclusion criterion as they were imputed at later stage using appropriate methods. Specific eligibility criteria for each of subcohorts were as following:

1. Primary prevention of HF subcohort (**Research aims 1a and 1b**) – Eligible were participants who had no prevalent HF during initial visit assessment and underwent NMR profiling of metabolite measures from baseline plasma sample. Furthermore, participants with prevalent CAD, VHD, AF, DM or users of lipid-lowering medication were excluded. Exclusion of participants with those prevalent events was motivated by

establishing a more homogeneous cohort by excluding participants at increased HF risk. This ensures that any metabolomic measure changes observed can be more confidently linked to the development of HF rather than other existing diseases. Similarly, exclusion of lipid-lowering medications users aimed to omit a confounding from the fact that therapy influences lipoprotein measures which constitute to substantial part of NMR panel.

Figure 1. Flowchart for selection of UK Biobank participants to the present study.



Abbreviations: AF - Atrial Fibrillation, CAD - Coronary Artery Disease, DM - Diabetes Mellitus, HF - Heart Failure, NMR - Nuclear Magnetic Resonance, VHD - Valvular Heart Disease.

2. Cardiovascular MRI imaging in healthy participants subcohort (**Research Aim 2**) – Eligible were participants who at baseline met criteria for primary prevention subcohort. Further participants who developed conditions significantly impairing cardiac function (incident HF, CAD, VHD, AF, DM) between baseline and date of MRI assessment. The goal of this was to emulate conditions where apparently healthy participants underwent assessment of cardiac function in order to explore subclinical changes associated with metabolite measures.

3. HF survival subcohort (**Research Aim 3) –** Eligible were participants who had prevalent HF (diagnosis of HF made before the initial assessment visit) and underwent NMR profiling of metabolite measures from baseline plasma sample, without regard to status of HF-predisposing health events.

Baseline covariate assessment

Below are details about the covariates collected during the initial assessment visit, which were used in this study:

Age – Age of participants (years) was derived from a continuous variable "Age at recruitment" (Data field: 21022), which was based on NHS Primary Care Trust data available to UK Biobank prior to the recruitment.

Sex – Sex of participants was derived from a binary variable "Sex" (Data field: 31) with two levels: (1) Male; (2) Female, similarly, extracted from NHS Primary Care Trust data.

Self-reported ethnicity – As a part of touchscreen questionnaire during initial visit, participants declared their ethnicity stored as "Ethnic background" (Data field: 21000), a character variable with 20 levels reflecting UK Office for National Statistics ethnicity classification. Original variable was re-coded into variable with four levels: (1) White; (2) Asian; (3) Black, (4) Other/Mixed. This was accomplished by grouping: "White", "British", "Irish", "Any other white background" as (1) White; "Asian or Asian British", "Indian", "Pakistani", "Bangladeshi", "Chinese" or "Any other Asian background" as (2) Asian; "Black or Black British", "Caribbean", "African", "Any other Black background" as (3) Black; "Mixed", "White and Black Caribbean", "White and Black African", "Any other mixed background", "Other ethnic group", "Do not know" as (4) Other/Mixed. "Prefer not to answer" was treated as missing data.

Townsend Deprivation Index (TDI) – This is a socioeconomic status measure based on four census variables (unemployment, non-car ownership, non-home ownership, household overcrowding) reflecting degree of deprivation of living area for a given participant. It was derived from a continuous variable "Townsend deprivation index at recruitment" (Data field: 21000), calculated from participants' postcodes available from NHS Primary Care Trust data.

Medication use – During nurse-led computer-assisted interview participants were asked about number and type of prescribed and over-the-counter medication taken on regular basis. Their answers were stored in a list-type character variable "Treatment/medication code" (Data field 20003), containing 6745 free-text labels (substance and brand names) of taken medication. Free-text labels reported by >1 participants were manually recoded into closest level four or five Anatomical Therapeutic Chemical (Codes) Classification System. Then, three new binary variables (a) Use of anti-hypertensive medication, (b) Use of anti-diabetic medication, (c) Use of lipid-lowering medication were created based on self-reported information of taking any medication linked with respective ATC codes (Levels: [1] User, [2] Non-user). Anti-diabetic medications were defined as reporting of A10 class use (*Drugs used in diabetes*). Anti-hypertensive medications were defined as reporting of C02 (*Antihypertensives*), C03(A-E) (*Diuretics* excluding *Other diuretics*), C07 (*Beta blocking agents*), C08 (*Calcium channel blockers*), and C09(A-D) (*Agents acting on the renin-angiotensin system* excluding *Other agents acting on the renin-angiotensin*) class use. Lipid-lowering medications were defined as reporting of C10 class use (*Lipid modifying agents*) excluding use of omega-3/fish oil supplements. Information on dosing and frequency, time on therapy or adherence was not available.

Self-reported history of diseases and operations – As a part of nurse-led computer-assisted interview participants were asked about conditions that participants had been told by their doctor or operative procedures they underwent prior to the initial visit assessment. Additionally, participants were asked about an approximate age or year of diagnosis/procedure. Information on self-reported non-cancer disease and self-reported procedures were derived from a list-type character variables "Non-cancer illness code, self-reported" (Data field 20002) and "Operation code" (Data field 20004). Dates of diagnosis or procedure were derived from numerical variables "Non-cancer illness year/age first occurred" (Data field 87 – Array 0 to 33) "Operation year/age first occurred" (Data field 92 – Array 0 to 31). If participant reported age of diagnosis, it was converted to year of diagnosis by subdividing age at baseline. January

1st was assumed as exact date for each self-reported event. Data extracted in that way were further utilized to define heath events according to methods summarized in **Outcomes subsection**.

Smoking status – Questions on smoking habits were part of a touchscreen status. Current smoking status was derived from a character variable "Smoking status" (Data field 20116) with three levels: (1) Current; (2) Previous; (3) Never. "Prefer not to answer" was treated as missing data.

Alcohol intake – As part of the touchscreen questionnaire participants were asked about average intake of six beverage types: red wine, champagne plus white wine, beer plus cider, spirits, fortified wine, and other alcoholic types. Depending on overall frequency of alcohol intake participants were asked about monthly or weekly intake (number of pints for beer plus cider or glasses for other beverage types). These data were stored as 12 numeric variables (Data fields: 1568, 1578, 1588, 1598, 1608, 4407, 4418, 4429, 4440, 4451, 4462 and 5364). Previously published alcohol unit equivalents were used to calculate Daily ethanol intake (g/day) stored as a new variable for this study.

Healthy diet score – A shortened screener-like food frequency questionnaire was a part of touchscreen questionnaire. To approximate overall diet quality of participants, a Healthy Diet Score (HDS) was calculated, according to previously published criteria. HDS consists of seven domains, each scored up to 1 point (total range 0-7 points), with a higher score indicating a better diet quality. Domains include (1) fruit intake (1 point if ≥3 servings/day), (2) vegetable intake (1 point if ≥3 servings/day) (3) fish intake (1 point if ≥2 servings/week), (4) processed meat intake (1 point if ≤1 serving/week), (5) unprocessed red meat intake (1 point if ≤1.5 servings/week), (6) whole grains intake (1 point if ≥3servings/day), and (7) refined grains intake (1 point if ≤1.5 servings/day). Fruit intake was derived as a sum of numeric variables (pieces/day) "Fresh fruit intake" (Data field 1309) and "Dried fruit intake" (Data field 1319). Vegetable intake was derived as a sum of numeric variables (tablespoons/day) "Cooked vegetable intake" (Data field 1289) and "Salad / raw vegetable intake" (Data field 1299). Fish intake was derived as a sum of character variables "Oily fish intake" (Data field 1329) and "Non-oily fish intake" (Data field 1339) converted to numeric servings/week. Processed meat intake was derived from a character variable "Processed meat intake" (Data field 1349), while unprocessed red meat intake was derived as a sum of "Beef intake" (Data field 1369), "Lamb/mutton intake" (Data field 1379) and "Pork" (Data field 1389), all converted to numeric

servings/week. Whole and refined grains intakes were derived from variables on quality ("Bread type" [Data field 1448] and "Cereal type" [Data field 1468) and quantity ("Bread intake" [Data field 144838] and "Cereal intake" [Data field 1458) converted to serving/day (treating slice of bread or bowls of cereals as one serving). Categorical HDS >5 points was interpreted as categorical High diet quality.

BMI – During initial visit assessment, participants underwent a number of physical measurements including anthropometry. Standing height in meters was measured using SECA 240 cm height measure and weight in kilogram was measured using TANITA BCV418MA Body Composition Analyzer. BMI was calculated by diving weight in kilograms per squared height in meters (kg/m^2). BMI was derived from an available numeric variable "Body mass index" (Data field 21001).

Blood pressure measurements – Participants' systolic (SBP) and diastolic blood pressure (DBP) was measured twice, while setting using either a manual sphygmometer or automatic OMRON 705 IT Electronic Blood Pressure Monitor. SBP and DBP were derived as mean of two taken measurements of SBP and DBP using manual or automatic method. Raw SBP and DBP values were stored as numeric variables "Systolic blood pressure, automated reading" (Data field 93), "Systolic blood pressure, automated reading" (Data field 4080), "Diastolic blood pressure, manual reading" (Data field 94) and "Diastolic blood pressure, automated reading" (Data field 4079).

Fasting time – Prior to biological sample collection participants were asked about time since least meal or non-water drink. This was stored and derived from a numeric (hours) variable "Fasting time" (Data field 74).

Non-NMR biomarkers – During initial assessment visit participants underwent non-fasted collection of blood and urine samples. Stored fractions included among others plasma (EDTA and EDTA in a gel plasma separation tube), buff coat (EDTA), red cells (EDTA) and serum (clot activator in a gel serum separation tube). Plasma and serum sample were maintained at 4ºC before they were dispatched to the central processing laboratory in a temperature-maintained conditions. Collected blood samples were dived into 26 1.4 ml aliquots and stored at -80ºC or in liquid nitrogen. Thirty-four serum biomarkers were selected and quantified for all baseline participants (Fry et al., 2019). For this study, hemoglobin A1c (HbA1c) and creatinine were utilized. HbA1c was derived from a numeric variable "Glycated haemoglobin

(HbA1c)" (Data field 30750). The HbA1c assay was performed using high-performance LC Bio-Rad Variant II Turbo analyzer (Bio-Rad Laboratories, Inc). Measurements were given in mmol/mol, and the analytical range specified by the manufacturer is between 15-184 mmol/mol (Tierney et al., 2018). Creatinine was derived from a numeric variable "Creatinine" (Data field 30700). The creatinine assay was performed using enzymatic technique on Beckman Coulter AU5800 analyzer (Beckman Coulter, Ltd, UK). Measurements were given in μmol/L, and the analytical range specified by the manufacturer is between 0-4420 μmol/L (Fry et al., 2019). Creatinine level together with age, sex and ethnicity to calculate estimated glomerular filtration rate (eGFR) using MDRD formula (Levey et al., 2006).

Nuclear magnetic resonance metabolite measure profiling

Overview

Nightingale Health Ltd. (Helsinki, Finland) was responsible for high-throughput NMR profiling of metabolite measures in UK Biobank participants. As of August 2023, analyzed data are available from initial phase 1 release, the subsequent phase 2 release, and minor updates between the two. With release of future phases, all UK Biobank participants will have quantified metabolite measure data.

During the phase 1 release, data on metabolite measures from a random sample pf around 118,000 participants at their initial recruitment and 5,000 during their repeat assessment (with about 1,500 having both) were made available. These samples underwent measurement between June 2019 and April 2020 at Nightingale Health's labs in Finland. In contrast, the phase 2 release added data from an extra 157,000 participants at the start of their recruitment and 12,000 during their repeat assessments. The time frame for these measurements spanned from April 2020 to June 2022. Consequently, the total count of samples with metabolic biomarker data has risen to 292,000, breaking down into 275,000 from initial recruitments and 17,000 from repeat assessments. Among these, approximately 15,500 samples represent both initial and repeat assessments. In the present study, metabolite measure data came from phase 1, with NMR profiling available initially for 118,021 participants shared with principal investigators who were granted access in September 2021.

Detailed NMR methodology and analytical protocol (Inouye et al., 2010; Soininen et al., 2015; Soininen et al., 2009) and specific protocol in UK Biobank were published previously (Julkunen et al., 2023).

Nomenclature

In the context of the Nightingale platform and others like it, many authors have described them as "metabolomic" or "metabonomic" platforms. This has led to the biomarkers measured by this technology being referred to as "metabolites," a term that may not always be precise. While the Nightingale platform does quantify actual metabolites, such as amino acids or ketone bodies, a significant portion of its output represents combined metrics, such as the total content of a specific lipid fraction (e.g., total phospholipids) or functional measures expressed in units other than concentration (e.g., the average diameter of lipoproteins). Recognizing this distinction, the UK Biobank has adopted the term "metabolic biomarker." However, this might be confused with classical clinical chemistry biomarkers and those derived from other techniques. To avoid this confusion, throughout this study, biomarkers from the Nightingale platform will be referred to as "metabolite measures".

Preparation of Plasma Samples

In the UK Biobank laboratory located in Stockport, UK, EDTA plasma samples from aliquot 3 were systematically arranged and processed using 96-well plates. At least 90 μL of plasma was aliquoted in each well using TECAN freedom EVO 150 robotic liquid handlers. These samples, batched between 5,000 to 20,000 units, were securely transported on dry ice to the Nightingale Health labs in Finland.

Once the samples arrived at Nightingale Health laboratories, they were kept at −80 °C. Prior to any processing, samples were methodically thawed at a steady +4 °C overnight. After the samples were gently mixed, and centrifugation (3 min, 3400 × g, +4 °C) was performed to eliminate any potential precipitate. Aliquots of each sample were transferred into 3-mm outer-diameter NMR tubes and mixed in 1:1 ratio with a phosphate buffer (75 mM Na_2HPO_4 in 80%/20% H_2O/D_2O, pH 7.4, including also 0.08% sodium 3-(trimethylsilyl) propionate-2,2,3,3-d4 and 0.04% sodium azide) automatically with an automated liquid handler. (PerkinElmer, Inc, USA).

Procedure for NMR Spectroscopy

Plasma samples were analyzed with six 500 MHz Bruker AVANCE IIIHD NMR spectrometers. Personnel involved in NMR analyses were blinded towards health status of participants. The pre-processed plasma samples, settled on 96-well plates, were put on a temperature-regulated sample changer, keeping them at +6 °C until measurement. Two distinct

NMR spectra were measured in each plasma sample. The first was a presaturated proton spectrum, which represented resonances arising mainly from proteins and lipids within various lipoprotein particles. Conversely, the second was Carr-Purcell-Meiboom-Gill T_2-relaxation-filtered spectrum, which allowed for detection of small-scale metabolites by minimizing broader macromolecule and lipoprotein lipid signals. The metabolic biomarkers were quantified using Nightingale Health's proprietary software (Nightingale Health biomarker quantification library 2020).

Quantified metabolite measures

Nightingale NMR high-throughput platform includes 168 crude metabolite measures and 81 relative metabolite measures organized in 45 classes. Out of 168 metabolite measures, 106 primary metabolite measure which were non-derived meaning that they were directly measured by spectrometer rather than calculated from other measures. Primary metabolite measures included (unit for all is mmol/L, unless other indicated):

- **Amino acids**: alanine, glutamine, glycine, histidine, isoleucine, leucine, phenylalanine, tyrosine, valine;
- **Glycolysis-related metabolites**: citrate, glucose, lactate, pyruvate;
- **Ketone bodies**: 3-OHB, acetate, acetoacetate, acetone;
- **Fluid balance**: albumin, creatinine;
- **Inflammation**: glycoprotein acetyls;
- **FA**: degree of unsaturation [dimensionless], docosahexaenoic acid (DHA), linoleic acid, omega-3 FAs, omega-6 FAs, monounsaturated FAs (MUFAs), saturated FAs (SFAs);
- **Apolipoproteins**: apolipoprotein A1 (apoA1), apolipoprotein B (apoB);
- **Lipoprotein particle sizes**: average diameter for VLDL, average diameter for low-density lipoprotein (LDL), average diameter for HDL [all three in nm];
- **14 lipoprotein size subclasses based on average diameter of particles** – for each class there were six primary measures: particle count and concentration of cholesteryl esters, free cholesterol, triglycerides, and phospholipids – 84 measures together. Size subclasses included:
    - **Chylomicrons and extremely large VLDL** - particle diameters from 75 nm upwards;
    - **Very large VLDL** - average diameter 64 nm;

    - **Large VLDL** - average diameter 53.6 nm;
    - **Medium VLDL** - average diameter 44.5 nm;
    - **Small VLDL** - average diameter 36.8 nm;
    - **Very small VLDL** - average diameter 31.3 nm;
    - **Intermediate-density lipoprotein (IDL)** - average diameter 28.6 nm;
    - **Large LDL** - average diameter 25.5 nm;
    - **Medium LDL** - average diameter 23 nm;
    - **Small LDL -** average diameter 18.7 nm;
    - **Very large HDL** - average diameter 14.3 nm;
    - **Large VLDL** - average diameter 53.6 nm;
    - **Medium HDL** - average diameter 10.9 nm;
    - **Small HDL** - average diameter 8.7 nm;
- **Other lipids**: PCs, Phosphoglycerides, SMs, Total Cholines;

Further 62 measures were derived using directly quantified using metabolite measures:

- **Amino acids**: total concentration of BCAAs – sum of leucine, isoleucine and valine.
- **FAs**: polyunsaturated FAs (PUFA) – sum of omega-3 and omega-6 FAs.
- **14 lipoprotein size subclasses**: for each size subclasses total cholesterol concentration was calculated by summing class-specific cholesteryl esters and free cholesteryl and total lipid concentration was calculated by summing class-specific total cholesterol, triglycerides and phospholipids – 28 measures together.
- **Lipoprotein-level lipid fractions**: for each class VLDL, LDL, HDL: particle count, cholesterol, cholesteryl esters, free cholesterol, triglycerides, phospholipids, and total lipid concentration by summing respective fractions within each lipoprotein class (six for VLDL, three for LDL, and four for HDL) – 18 measures together.
- **Total lipid fractions**: Particle count, total cholesterol, cholesteryl esters, free cholesterol, triglycerides, phospholipids and total lipid concentration by summing respective fractions in VLDL, IDL, LDL and HDL – 6 measures together. These together with lipoprotein-level lipid fractions measures comprised for: (1) **Lipoprotein particle concentration**, (2) **Cholesterol**, (3) **Cholesteryl esters**, (4) **Free cholesterol**, (5) **Phospholipids** and (6) **Total Lipids** classes.
- **Cholesterol**: Clinical LDL cholesterol (Nightingale nomenclatures) was calculated using Friedewald formula from NMR total cholesterol, HDL cholesterol and

triglycerides (Friedewald et al., 1972). Remnant cholesterol was calculated by subdividing LDL and HDL cholesterol from total cholesterol (sum of VLDL and IDL fractions). Non-HDL cholesterol was calculated by subdividing HDL cholesterol from total cholesterol.

Following 81 ratio measures were calculated:

- **FAs**: relative DHA, linoleic acid, omega-3 FAs, omega-6 FAs, PUFA, MUFAs, SFAs as percentages of total FAs, omega-6 FAs to omega-3 FAs ratio, PUFA to MUFAs ratio.
- **Apolipoproteins**: apoB to apoA1 ratio.
- **14 lipoprotein size subclasses:** for each size subclasses relative cholesterol, cholesteryl esters, free cholesterol, triglycerides and phospholipids as percentages of total lipid concentration per particle subclass – 70 measures together.
- **Other lipids**: triglycerides to phosphoglycerides ratio.

Quality control

Nightingale NMR platform has received CE-marking, fulfilling the laboratory criteria for clinical application (IVD 98/79/EC; Directive on In Vitro Diagnostic Medical Devices). High agreement has been indicated when Nightingale metabolite measures were compared with clinical biochemistry in cohort studies, with correlation coefficients being commonly observed as Pearson's r>0.9. The coefficient of variation (CV) across spectrometers has been measured using internal control samples and blind duplicates provided by Nightingale Health. For most of the metabolic biomarkers, the CVs have been found to be below 5%. The sample quality is verified through integrated quality procedures by Nightingale, and signs of degradation and contamination issues are reported as quality control (QC) flags alongside the result data. QC flags that affect the entire sample are reported as sample-level flags; those affecting specific biomarkers are reported as metabolite measure-level flags and were provided in a separate variable for each biomarker. Common QC flags included "Below limit of quantification" in case of very low lipid concentrations (<0.01 mmol/L) or "High ethanol" in case of sample contamination with skin disinfectant.

Data preprocessing

Metabolite measure data were stored as 249 numeric variables (Data fields 23400 to 23647), sample processing information as 11 character, numeric or date variables (Data field 23649 to

23600), and metabolite measure QC flags as 249 character variables (Data field 23661 to 23948).

Although NMR data released for researcher could be utilized in analysis, a report from an independent group pointed about sourced of unaccounted technical variation that for some measure like alanine explained ~30% of variance (Ritchie et al., 2023). Therefore, a method proposed by authors of this report was adopted to remove sources of technical variation. The method involved multi-step robust linear regression models constructed on residuals from previous step. Steps included regression of log-concentration over sample degradation time, plate row, plate column, time drift within spectrometer, removal of outlier plates and back-transformation to original concentrations.

Outcomes

The primary outcome of this study was incident HF (**Research Aim 1a** and **1b**). Secondary outcomes included MRI LV function measures (**Research Aim 2**), all-cause and cardiovascular mortality among participants with prevalent HF (**Research Aim 3**). Below are described details of how outcomes were derived.

Longitudinal health data

Data collected from participants during study visits were integrated with longitudinally collected health information, encompassing both hospital inpatient records and death registries. This integration enabled the determination of health events that transpired both before and after recruitment. Each data source underwent regular updates. By the time of data retrieval for this study, health events were available up to 31st March 2021. This date was considered as end of follow-up for this study and participants were censored at this date.

Hospital inpatient data

The UK Biobank collates hospital inpatient data from three primary regions in the United Kingdom. For England, the Data Access Request Service, overseen by NHS England, supplies data from the Hospital Episode Statistics Admitted Patient Care (APC) dataset. In Wales, the data comes from the Secure Anonymised Information Linkage Databank hosted by the University of Swansea, under the stewardship of NHS Wales Informatics Service's Information Services Division. This data corresponds to the Patient Episode Database for Wales APC. Scotland contributes inpatient data through Public Health Scotland, a subsidiary of NHS

National Services Scotland. The Scottish data is segmented into two distinct datasets: the General Acute Inpatient and Day Case - Scottish Morbidity Record and the Mental Health Inpatient and Day Case - Scottish Morbidity Record (UK Biobank, 2023a).

Historically, the dataset inception dates are 1997 for England, 1998 for Wales, and a notably earlier date of 1981 for Scotland. These datasets encapsulate coded data relating to hospital admissions, surgical procedures, and other interventions.

Clinically, the datasets employ universally recognized coding systems. The World Health Organization's ICD (International Classification of Diseases and Related Health Problems) is used for disease classification. Meanwhile, interventions and surgical procedures are coded as per the OPCS (Office of Population, Censuses and Surveys: Classification of Interventions and Procedures) system. Pertinently, all current hospital data linked to UK Biobank from England and Wales adhere to the ICD-10 and OPCS-4 conventions. Due to the earlier initiation of data collection in Scotland (1981), any Scottish data gathered before 1997 is aligned with the ICD-9 and OPCS-3 classifications.

Hospital inpatient data can be accessed in two major formats, record-level and summary-level. In record-level dataset, every inpatient episode for a participant is represented either by a single data row or by several consecutive records. In contrast, summary-level dataset stores only first occurrence of a given ICD or OPCS code along with its date. Four-digit ICD-10 along with date of events were derived from a listed character variable "Diagnoses – ICD10" (Data field 41270) and date variable "Date of first in-patient diagnosis – ICD10" (Data field 41280, Array 0 to 258). Four-digit ICD-9 along with date of events were derived from a listed character variables "Diagnoses – ICD9" (Data field 41271) and date variable "Date of first in-patient diagnosis – ICD9" (Data field 41281, Array 0 to 46). Four-digit OPCS-4 along with date of procedure were derived from a listed character variables "Operative procedures – OPCS4" (Data field 41272) and date variable "Date of first operative procedure – OPCS4" (Data field 41282, Array 0 to 125).

Death register

Data on deaths of UK Biobank participants that occurred during follow-up were obtained from NHS England for participants in England & Wales and from the NHS Central Register (NHSCR), part of the National Records of Scotland, for participants in Scotland. Data provided to the UK Biobank from death registry covered the death date and both primary and

contributory causes of death, all encoded using the ICD-10 system (UK Biobank, 2023b). Notably, there were instances where two death records were made available by data sources, often when a secondary death certificate was issued after a post-mortem examination. In such a situation, for this study reported causes of death from both certificates were compared and in case of matching event identification, the one reporting it was selected. Primary and secondary causes of death along with date of death were derived from a listed character variables "Underlying (primary) cause of death: ICD10" (Data field 40001, Instances 0 to 1), "Contributory (secondary) causes of death: ICD10" (Data field 40002, Instances 0 to 1) and a date variable "Date of death" (Data field 40000).

Terminology for health events

In this study three types of health events were considered: prevalent, incident, and interim. By prevalent event should be understood an event which date of diagnosis was made before the date of recruitment to the UK Biobank. Incident event was defined as an event that occurred between date of recruitment and the end of follow-up. Interim events were used only in the context of disease predisposing to the occurrence of primary outcome (incident HF) and were defined as events occurring between the date of recruitment and date of primary outcome diagnosis.

Incident and interim events were determined using hospital inpatient data and death record and were considered as earliest diagnosis or death with a code matching to the definition of a particular health event. For prevalent events, additionally self-reported disease and procedures were incorporated, but in case when self-reported outcome was confirmed by hospital inpatient data the date of hospitalization was preferred over approximated self-reported data from nurse-led interview.

Definition of health events

Below are listed self-reported disease and procedure names, ICD-9, ICD-10, and OPCS-4 codes that were used to define following health events:

HF: 1) Self-reported "Heart failure" or "Cardiomyopathy" during nurse-led interview OR 2) hospitalization of death due to ICD-10: I11.0 (*Hypertensive heart disease with [congestive] heart failure*) I13.0 (*Hypertensive heart and renal disease with [congestive] heart failure*), I13.2 (*Hypertensive heart and renal disease with both [congestive] heart failure and renal failure*), I25.5 (*Ischemic cardiomyopathy*), I42.0 (*Dilated cardiomyopathy*), I42.5 (*Other*

restrictive cardiomyopathy), I42.8 (*Other cardiomyopathies*), I42.9 (*Cardiomyopathy, unspecified*), I50.0 (*Congestive heart failure*), I50.1 (*Left ventricular failure*) or I50.9 (*Heart failure, unspecified*) OR 3) hospitalization due to ICD-9 code: 425.4 (*Other primary cardiomyopathies*), 428.0 (*Congestive heart failure unspecified*), 428.1 (*Left heart failure*), 428.9 (*Heart failure unspecified*) BUT without self-reported "Hypertrophic cardiomyopathy" during nurse-led interview; NOR 5) hospitalization or death due to I42.1 (*Obstructive hypertrophic cardiomyopathy*) or I42.2 (*Other hypertrophic cardiomyopathy*). This definition was adopted from a previous genome-wide association study on phenotypic refinement of HF. As such it is broader than typically used epidemiological definitions of HF (usually I50.X alone or I50.X and I11.0) as includes cardiomyopathy diagnosis as synonymous HF (Aragam et al., 2019). Such approach can be explained by frequent presence of cardiomyopathies as a syndrome of HF and significant increase in statistical power by increase a in number of available cases (Seferović et al., 2019). However, this definition excluded hypertrophic cardiomyopathy due to its high genetic heritability.

CAD: 1) Self-reported "Myocardial infarction" or "Coronary artery angioplasty" or "triple heart bypass" during nurse-led interview; OR 2) hospitalization of death due to ICD-10: I20 (*Angina pectoris*), I21 (*Acute myocardial infarction*), I22 (*Subsequent myocardial infarction*), I23 (*Certain current complications following acute myocardial infarction*), I24 (*Other acute ischemic heart diseases*) or I25 (*Chronic ischemic heart disease*); OR 3) hospitalization due to ICD-9 code: 410 (*Acute myocardial infarction*), 411 (*Other acute or subacute forms of ischemic heart disease*), 412 (*Old* myocardial infarction), 413 (*Angina pectoris*) or 414 (Other forms of chronic ischemic heart disease), OR 4) hospitalization due to OPCS-4 code: K40 (*Saphenous vein graft replacement of coronary artery*), K41 (*Other autograft replacement of coronary artery*), K45 (Connection of thoracic artery to coronary artery) K49 (*Transluminal balloon angioplasty of coronary artery*), K50.2 (*Transluminal balloon angioplasty of coronary artery*), K75 (*Percutaneous transluminal balloon angioplasty and insertion of stent into coronary artery*). This definition was adopted from a previously published study (Aragam et al., 2019).

AF: 1) Self-reported "*Atrial fibrillation*", "*Atrial flutter*" or "*Cardioversion*" during nurse-led interview; OR 2) hospitalization or death due to ICD-10: I48 (*Atrial fibrillation and flutter*); OR 3) hospitalization due to ICD-9 code: 427.3 (*Atrial fibrillation and flutter*).

VHD: 1) Self-reported "Aortic stenosis", "Aortic valve disease", "Aortic regurgitation", "Aortic valve repair/replacement", "Mitral valve disease,", "Mitral stenosis", "Mitral valve prolapse", "Mitral regurgitation" or "Mitral valve repair/replacement" during nurse-led interview; OR 2) hospitalization or death due to ICD-10: I05 (*Rheumatic mitral valve diseases*), I06 (*Rheumatic aortic valve diseases*), I07 (*Rheumatic tricuspid valve diseases*), I08.0 (*Disorders of both mitral and aortic valves*), I08.1 (*Disorders of both mitral and tricuspid valves*), I08.2 (*Disorders of both aortic and tricuspid valves*), I08.3 (*Combined disorders of mitral, aortic and tricuspid valves*), I34 (*Nonrheumatic mitral valve disorders*) I35 (*Nonrheumatic aortic valve disorders*), I36 (*Nonrheumatic tricuspid valve disorders*), I37 (*Pulmonary valve disorders*) I39.0 (*Mitral valve disorders in diseases classified elsewhere*), I39.1 (*Aortic valve disorders in diseases classified elsewhere*), I39.3 (*Pulmonary valve disorders in diseases classified elsewhere*); OR 3) hospitalization due to ICD-9 code: 394.0 (*Mitral stenosis*), 394.2 (*Mitral stenosis with insufficiency*), 394.9 (*Other and unspecified diseases of mitral valve*), 424.0 (*Mitral valve disorders*) or 424.1 (*Aortic valve disorders*); OR 4) hospitalization due to OPCS-4 code: K25 (*Plastic repair of mitral valve*), K26 (*Plastic repair of aortic valve*), K27 (*Plastic repair of tricuspid valve*), K28 (*Plastic repair of pulmonary valve*), K30.1 (*Revision of plastic repair of mitral valve*), K30.2 (*Revision of plastic repair of aortic valve*), K30.3 (*Revision of plastic repair of tricuspid valve*), K30.4 (*Revision of plastic repair of pulmonary valve*), K34.1 (*Annuloplasty of mitral valve*), K34.2 (*Annuloplasty of tricuspid valve*), K35.1 (*Percutaneous transluminal mitral valvotomy*) or K35.7 (*Percutaneous transluminal pulmonary valve replacement*). This definition was adopted from a previously published study (Aragam et al., 2019).

DM: 1) Self-reported "Diabetes", "Type 1 diabetes" or "Type 2 diabetes" during nurse-led interview; OR 2) hospitalization or death due to ICD-10: E10 (*Insulin-dependent diabetes mellitus*), E11 (*Non-insulin-dependent diabetes mellitus*), E13 (*Other specified diabetes mellitus*), E14 (*Unspecified diabetes mellitus*); OR 3) hospitalization due to ICD-9 code: 250 (*Diabetes mellitus*); OR 4) Glycated hemoglobin (HbA1c) during baseline assessment ≥48 mmol/mol (6.5%); OR 5) Use of anti-diabetic medication according during verbal interview with a nurse.

All-cause mortality: Death irrespective of underlying or contributory cause

Cardiovascular mortality: Death due to ICD-10 IXX.X code (*Disease of the circulatory system*) as an underlying cause of death.

Cardiovascular magnetic resonance imaging protocol

The UK Biobank imaging protocol and procedures were published in detail previously (Petersen et al., 2013; Petersen et al., 2016; UK Biobank, 2015). Each participant underwent a 20-minute cardiovascular MRI scan without pharmacological stressor or contrast agent, as part of a 90-minute-long imaging protocol that was implemented during imaging visit as a part of ongoing effort to collected imaging for 100,000 UK Biobank participants that started in 2014.

Cardiovascular MRI scan was conducted using a 1.5 Tesla scanner MAGNETOM Aera, Syngo Platform VD13A (Siemens Healthcare, Germany). The UK Biobank's cardiac magnetic resonance acquisitions included piloting and partial coverage of the chest and abdomen in three different planes: sagittal, transverse, and coronal. For the assessment of cardiac function, three types of long axis motion pictures, including the horizontal long axis, vertical long axis, and the LV outflow tract in both sagittal and coronal orientations, were obtained. Additionally, a complete short axis stack of balanced steady state free precession cines, encompassing the LV and right ventricle (RV), was acquired.

All radiographers participated in a standardized central training program, ensuring that all scans were executed according to set standard operating procedures. Each scan's acquisition was meticulously evaluated in real time by the technician to ensure it was complete. Once the data was collected, a senior radiographer examined the cardiac magnetic resonance images to identify the quality of the images and detect any artifacts.

During the scan, Siemens InlineVF analysis software (Siemens Healthcare, Germany) was enabled allowing automated assessment of LV function. The software had automatically pinpointed LV landmarks at the LV base (mitral valve) and apex in long-axis cine acquisitions, identified both endocardial and epicardial contours at end-diastole and end-systole in every short-axis cine slice, and conducted volume calculations to ascertain ventricular function parameters. Given that UK Biobank provides LV end-diastolic volume (LVEDV), LV end-systolic volume (LVESV), LVEF, LV stroke volume (SV), cardiac output (CO) and cardiac index (CI). A previous study suggested that there was a moderate-to-good agreement between automatically derived LV parameters and those derived manually (Suinesiaputra et al., 2018). LV parameters were stored as six numeric variables (Data fields 22420 to 22425).

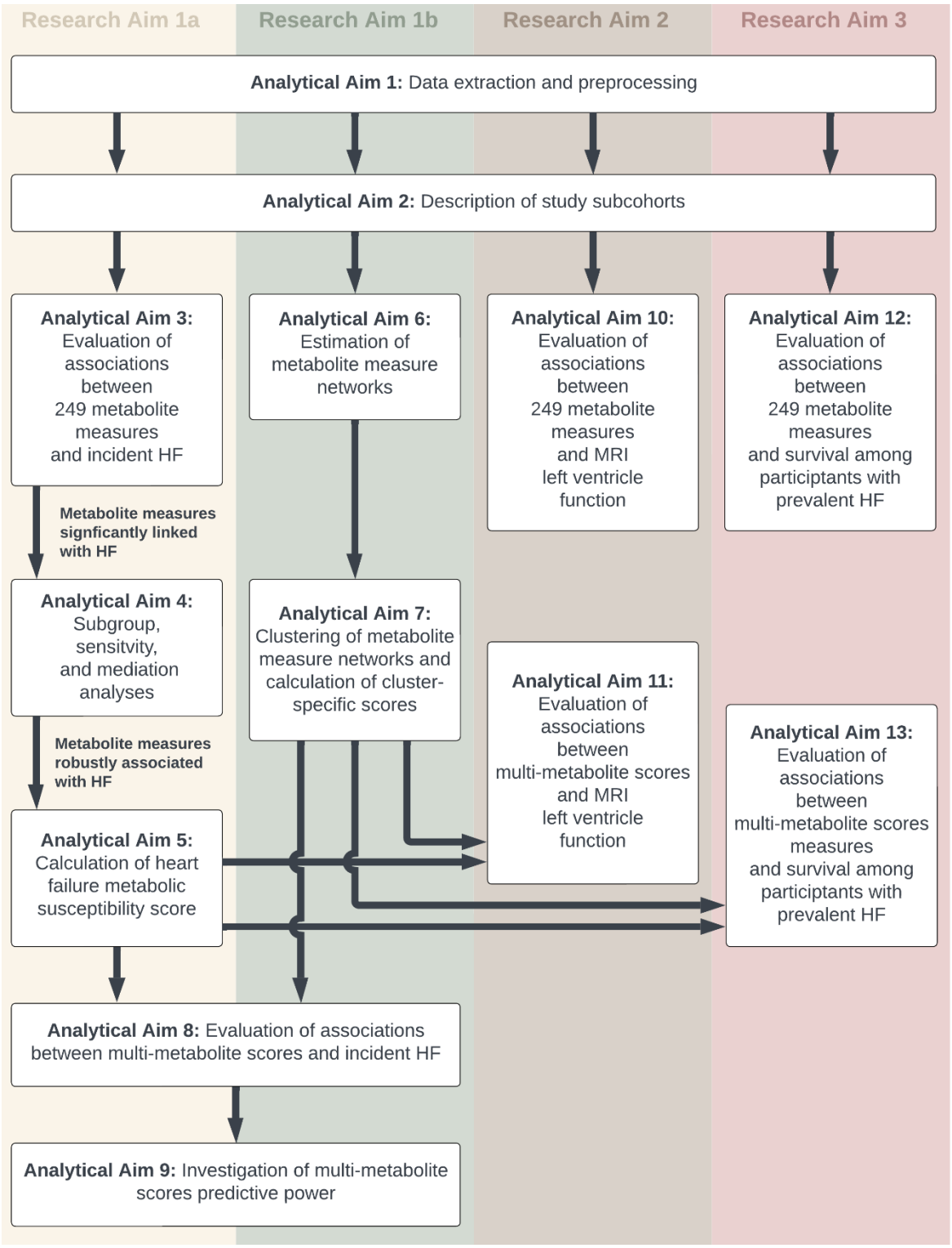
Moreover, LVESV, LVEDV and LV stroke volume index (SVI) were indexed on body surface area (Data field 22427). Rationale for this approach is strong determination of these parameters

by body surface area. Therefore, the list of parameters analyzed in context of link metabolite measures included five parameters: (1) CI [L/min/m^2], (2) LVEF [%], (3) LVESV index (LVESVI) [mL/m^2], (4) LVEDV index (LVEDVI) [mL/m^2]. and (5) SVI [mL/m^2].

Statistical analysis

Figure 2 depicts statistical analysis plan for this study translating selected research aims into analytical aims. Two initial steps (**Analytical aim 1** and **2**) were common for all research aims and involved extraction of exposure, outcome and covariate data, their preprocessing and description of data in context of selected study subcohorts. Aiming to identify metabolic pathways independently associated with the HF, two approaches were adopted classical step-wise and based on machine learning. As a part of classical step-wise approach first explored were association between 249 metabolite measures and incident IF (**Analytical aim 3**), then in subsequent step a series of subgroup, sensitivity and mediation analyses (**Analytical aim 4**) were conducted to identify metabolite measures robustly and independently associated with incident HF. Then, set of metabolite measures left was combined into a joint score reflecting HF metabolic susceptibility (**Analytical aim 5**). Simultaneously, network analysis and clustering were used to identify distinct metabolomic patterns (**Analytical aim 6 and 7**). Moreover, a cluster-specific scores were calculated to determine participants alignment with specific metabolomic patterns. In the next step, all multi-metabolite measure scores were evaluated for their association with HF risk (**Analytical aim 8**), and those significantly associated with HF predictive performance was tested (**Analytical aim 9**).

Figure 2. Flow-chart summarizing the main steps in the statistical analysis plan.



Regarding secondary outcomes, associations between 249 metabolite measures and LV parameters in cardiovascular MRI (**Analytical aim 10**), as well as survival among participants

with prevalent HF (**Analytical aim 12**), were evaluated. In subsequent steps, associations of previously identified multi-metabolite measures with secondary outcomes were further explored (**Analytical aim 11** and **13**).

Data preprocessing

With exception for HbA1c (4.5%), the proportion of missing covariate data for all other variables was <0.5% and the pattern of missing values seemed to be completely at random. Missing covariate data were imputed using k-nearest neighbors algorithm based on Gower distance assuming five nearest neighbors. The imputation was done separately in each of study subcohorts.

Missing rate for NMR metabolite measures did not exceed 5% (median missing rate: 1.1%). However, as missing values in NMR data were due to occurrence of QC flags or removal of outlier plates, they were clustered in specific participants and therefore missing not completely at random. Imputation of missing values was not required for single metabolite measure analyses as participants with missing values were excluded. For metabolite measure network estimation, no missing values were allowed. Therefore, for this analysis NMR data were imputed using iterative principal component analysis algorithm with two components used to predict missing values.

Metabolite measure concentrations were converted to z-scores (scaled to mean = 0 and standard deviation [SD] = 1). Normality of their distribution was evaluated using quantile-quantile plots. Majority of metabolite measures did not present highly skewed distributions (data not shown) At the initial analysis stage, different data transformations including logarithmic, square-root or inverse lang-normal were considered. However, data transformation did not significantly affected results, therefore consistently untransformed metabolite measures were used in this study.

Summary statistics

Baseline characteristics of study subcohorts were summarized using mean and SD (or median and interquartile range [IQR] in case of skewed variables) for continuous variables and number with percentage for categorical variables. Additionally, distribution for some parameters (e.g., LV function parameters) was presented using density plots.

Cox proportional hazards models

In the present study, Cox PH models were fitted for five purposes: (1) to evaluate associations between 249 metabolite measures and HF risk (**Analytical aim 3**), (2) as a major model used for subgroup/sensitivity/mediation analyses for significant metabolite measure-HF risk associations (**Analytical aim 4**), (3) to evaluate association between multi-metabolite measure scores and HF risk (**Analytical aim 8**), (4) to evaluate associations between 249 metabolite measures and mortality (all-cause and cardiovascular) in participants with prevalent HF (**Analytical aim 12**); (5) to evaluate the same relationship for multi-metabolite measure scores and HF survival (**Analytical aim 13**).

Cox PH models included one metabolite measure or multi-metabolite measure score at a time. Considering, that both types of exposure were converted to z-score, HR from these models can be interpreted as change in risk associated with one SD increase in either metabolite measure concentration or score value. Exception for this were analyses for HF susceptibility score where additionally HR for 10-point increase, as well as for quartiles of distribution were presented. For analytical aims 3, 4, and two sets of covariate adjustment were considered:

- **Model 1** adjusted for age (continuous), sex (male/female) and BMI (continuous);
- **Model 2** including covariates from Model 1, plus additionally self-reported ethnicity (Asian/Black/White/Mixed-Other), TDI (continuous), fasting time (continuous), smoking status (Current/Previous/Never), alcohol intake (continuous), eGFR (continuous), HbA1c (continuous), SBP (continuous), DBP (continuous), and use of anti-hypertensive medication (Yes /No).

For analytical aims 12 and 13 Model 2, additionally included following covariates:

- **Modified Model 2** including covariates from Model 2, plus additionally non-HDL-C, HDL-C, TG, use of lipid-lowering medication, and prevalent CAD/AF/VHD.

Results from Cox PH models were presented either in tabularized form, as volcano plots (scatter plot presenting HRs against log-transformed P-values) or interval plots.

P-values from single metabolite measure models were adjusted for multiple testing by controlling false discovery rate (P_{FDR}) and $P_{FDR}<0.05$ as statistically significant after the correction.

Test of proportionality was used to assess proportionality of hazards. The test checks if Pearson product-moment correlation between the scaled Schoenfeld residuals for each covariate and log-transformed time for each covariate is significantly different from 0 (indicating non-proportionality). All examined metabolite measures and scores did not show evidence for violation of proportionality assumption.

Subgroup and sensitivity analyses

Subgroup and sensitivity analyses were conducted for metabolite measures showing significant association with incident HF following multiple testing correction ($P_{FDR}<0.05$). Subgroup analyses involved splitting main analysis by a factor and fitting a Cox PH (including adjustment as in Model 2 excluding a given factor) in each subgroup. Additionally, a third model with multiplicate interaction term between the factor and metabolite measure to evaluate if effect is significantly different between groups. Following subgroup analyses were considered:

- Age: <65 vs. ≥65 years;
- Sex: Male vs. Female;
- Self-reported ethnicity: White vs. non-White;
- TDI: <median vs. ≥65 median (median TDI = -2.225);
- BMI: <30 vs. ≥30 kg/m^2;
- Alcohol intake: <17.1 vs. ≥17.1 g/day;
- Smoking: Current vs. Previous/Never;
- eGFR: <60 vs. ≥60 mL/min/1.73 m^2;
- SBP: <140 vs. ≥140 mmHg;
- HbA1c: <median vs. ≥65 median (median TDI = 34.7 mmol/mol).

Sensitivity analyses similarly involved refitting Model 2 for significant metabolite measures showing significant association with incident HF and additionally:

- Exclusion of first three years of follow-up to omit influence from high HF risk individuals;
- Use of participants' age as underlying time scale
- Further adjustment for healthy diet score;

- Further adjustment for physical activity level;
- Inclusion of model covariates showing evidence for nonlinearity with restricted cubic splines (with knots set at 10^{th}, 50^{th}, 90^{th} percentile of distribution).

Mediation analyses

Mediation analyses were conducted to evaluate the extent to which associations between metabolite measures and HF are mediated by an increased risk of predisposing events, as opposed to being directly and independently associated with HF. Interim diagnosis of CAD, AF, VHD, and DM were considered as potential mediators for link between metabolite measures and HF risk. In mediation analysis the total effect of exposure on outcome is assumed to be a sum of indirect effect through mediator and direct effect of exposure on outcome not explained by mediation. The ratio of indirect effect to total effect will be derived as attributable proportion of effect due to mediation. Indirect effect was calculated by subdividing coefficient for metabolite measure in a model adjusted for interim event from metabolite measure without such adjustment (total effect). Separately were estimated and presented Cox PH models for association between metabolite measures linked with HF and interim event risk, as well as Cox PH models for HF risk adjusted for interim event. Lastly, attributable proportion were calculated and summarized on a bar plot.

Additional hypothesis-driven models

Given problematic interpretation of ratio measures derived from highly correlated variables such as lipoprotein measures, to determine if those measure better inform HF risk than their components (nominator and denominator), Harrell's C-statistic was derived for a Cox PH model including covariates from Model 2 and separately given ratio, its nominator and denominator. Moreover, Pearson's correlation coefficients between the ratio and its denominator, as well as between the nominator and the denominator, were computed to assess if a given ratio introduces sparse correlation. Sparse correlation is defined here as a negative correlation between a ratio and its denominator, despite a high positive correlation between the nominator and denominator.

For lipoprotein classes that demonstrated a robust association with HF risk across all respective measures, a partial likelihood ratio statistic was calculated. A bootstrap resampling method with 1,000 iterations was employed to establish the relative ranking of these measures, along with their 95% CI. A higher value of the partial likelihood ratio can be interpreted as a better

predictor. Bootstrapped CI were calculated to determine if differences in rankings were statistically significant. This procedure aimed to discern whether an association for a given class is attributable to the number of particles (which drive its lipid content) or to a qualitative change in lipoprotein composition.

General linear models

General linear models (GLM) were used for two purposes (1) to evaluate associations of LV function parameters (CI, EF, LVESVI, LVEDVI, SVI) with 168 non-ratio metabolite measures (**Analytical aim 10**), as well as (2) with multi-metabolite measure scores (**Analytical aim 11**). GLM included one metabolite measure or multi-metabolite measure score at a time. Given that both kinds of predictors were standardized to z-scores, the regression coefficient (β) in these models can be interpreted as the difference in a specific LV function parameter corresponding to a 1-SD increase in the metabolite measure or score value. The covariate adjustments for the GLM were consistent with those specified earlier for the Cox PH models, denoted as Model 1 and Model 2.

Metabolite measure networks

Gaussian graphical model (GGM) is a probabilistic graphical model that represents the conditional independence structure between input variables. Nodes in the GGM are represented by metabolites, and the conditional dependence or independence between them, given the other metabolites, is indicated by the absence or presence of edges between nodes. The precision matrix, also known as the inverse covariance matrix, serves as the core mathematical foundation, and its off-diagonal elements are utilized to ascertain partial correlations between variables. In GGMs, zero entries in the inverse covariance matrix indicate conditional independence between variables. In high-dimensional data, the precision matrix often lacks zero entries, leading to dense, less informative graphs. To improve the model's accuracy and interpretability, regularization techniques like graphical least absolute shrinkage and selection operator (gLASSO) are used to enforce sparsity in the precision matrix. The gLASSO algorithm is a method that optimizes a penalized likelihood function, imposing constraints on the non-diagonal entries of the inverse covariance matrix. This constraint effectively reduces the estimated pairwise partial correlations, forcing minor or unreliable values to become zero, thereby leading to a sparse matrix that reflects direct relationships. Essentially, instead of simply maximizing the log likelihood, the algorithm optimizes a regularized form of the log likelihood to achieve a sparse structure. This regularization is accomplished by subtracting a

term from the log likelihood that consists of a nonnegative coefficient, λ, multiplied by the L1 norm. In this context, the L1 norm represents the sum of the absolute values of the elements in the inverse covariance matrix, and λ serves as a tuning parameter that controls the degree of shrinkage applied, also known as the regularization parameter. The aim of this is to identify a sparse graph where many of the pairwise conditional independencies are zero, effectively eliminating unnecessary edges in the corresponding graph. This results in a simpler, easier-to-interpret network of interactions among the variables, which is invaluable when dealing with high-dimensional data.

In this study, GGM derived with gLASSO algorithm was used to generate metabolite measure networks. The input for analysis included correlation matrix of 106 non-derived metabolite measures. The reason for dropping derived and ratio measures from analysis was justified by avoiding redundancy in dataset leading to less sparse, higher to interpret network. Value of λ was determined using 10-fold cross-validation with Extended Bayesian Information Criterium (EBIC) as the tuning parameter. Only edges representing partial correlation >|0.1| were retained in the network to improve biological interpretability of results.

On the top of derived networks walktrap algorithm was utilized to detect communities or clusters of nodes representing distinct metabolic pathways or classes. In brief, the algorithm does random walk on a graph starts from an initial node and moves to a randomly chosen neighboring node at each step. Algorithm does multiple random walks and then use the frequency of visits to each node during these walks to cluster nodes together, forming communities.

Multi-metabolite measure score calculation

In the present study, two distinct categories of multi-metabolite measure score calculations were performed.

First, the Heart Failure Metabolic Susceptibility Score (HFMSS) was computed by integrating metabolite measures that were significantly correlated with incident HF. Notably, these correlations were not largely attributed to underlying predisposing factors associated with HF. Two versions of the HFMSS – unweighted and weighted – were developed. In the unweighted version, the z-scores of the component metabolite concentrations were summed; positive values were assigned to components associated with higher HF risk, and negative values were assigned to those linked to lower HF risk. For the weighted version, a regularized Cox PH

model with ridge (L2) regularization was utilized. Ridge penalization was employed to ensure that all proposed score components were retained in the model, while simultaneously adjusting for intercorrelations among them. The optimal regularization parameter (λ) was identified through 10-fold cross-validation, with the objective being to maximize Harrell's C-statistic. The weighted HFMSS was then calculated as the sum of the component metabolites, each of which was weighted by their respective coefficients obtained from the regularized Cox PH model.

Second, cluster-specific susceptibility scores were developed using a regularized Cox PH model with elastic net penalization, which incorporated all metabolite measures identified in a given cluster. Unlike in the case of the HFMSS, it was not necessary for all components within a cluster to be predictive of HF. Elastic net regularization allowed for the coefficients used for weighting to be adjusted, while also enabling the exclusion of metabolite measures that were not informative for HF risk assessment. The procedures for estimating and calculating these cluster-specific scores were consistent with those used for the weighted HFMSS.

Predictive performance

The added benefit of multi-metabolite measurement scores in the prediction of HF was evaluated through a comparative analysis conducted using logistic regression models. In one model, conventional risk factors such as age, sex, BMI, self-reported ethnicity, TDI, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and the use of anti-hypertensive medications were included. In the second model, these factors were included along with an additional variable: the multi-metabolite measurement score.

For each pair of compared models receiver operating curves (ROC) were generated using linear prediction from both models. The difference between area under curve for two correlated ROCs was evaluated using DeLong test. As improvement in area under curve (AUC) in case of outcome where large proportion of variance is explained by clinical risk factor can be negligible, additionally net reclassification index (NRI), and integrated discrimination improvement (IDI) were considered.

NRI assesses reclassification of subjects into appropriate risk categories following addition of a new predictor compared to a baseline model. IDI is another metric that quantifies the improvement in model performance by comparing the average sensitivity and specificity

between the new and baseline models, thereby assessing the integrated sensitivity and specificity improvements across all possible cutoff points.

Statistical software

All statistical analyses were conducted using R version 4.2.2 (R Foundation for Statistical Computing, Austria) in RStudio Desktop version 2023.06.1+524 (Posit PBC, MA). Following additional packages were used for analyses: 'boot', 'bootstrap', 'broom', 'broom.helpers', 'car', 'carData', 'cluster', 'dplyr', 'FactoMineR', 'forcats', 'Formula', 'ggforce', 'ggforestplot', 'ggfortify', 'ggfun', 'ggplot2', 'ggplotify', 'ggpubr', 'ggrepel', 'ggsci', 'ggsurvfit', 'glasso', 'glmnet', 'gridExtra', 'gridGraphics', 'Hmisc', 'igraph', 'km.ci', 'lattice', 'latticeExtra', 'lubridate', 'MASS', 'Matrix', 'mice', 'missMDA', 'mixOmics', 'pROC', 'purrr', 'qgraph', 'RColorBrewer', 'readr', 'readxl', 'reshape', 'reshape2', 'rms', 'ROCR', 'SparseM', stringr, 'survival', 'survminer', 'tibble', 'tidyverse', 'timechange', 'ukbnmr'.

4. Results

Table 1. Comparison of participants' baseline characteristics between those with and without incident heart failure during follow-up.

Characteristic	Overall cohort	No incident HF during follow-up	Incident HF during follow-up
n	93,922	92,284	1,638
Age, years	55 (8)	55 (8)	61 (7)
Female sex	54,223 (58%)	53,506 (58%)	717 (44%)
Body mass index, kg/m^2	26.9 (4.6)	26.9 (4.5)	28.6 (5.4)
White ethnicity	89,270 (95%)	87,686 (95%)	1,584 (97%)
Townsend Deprivation Index	-1.40 (3.04)	-1.41 (3.04)	-0.95 (3.26)
High physical	31,635 (41%)	31,129 (41%)	506 (40%)
Current smokers	9,813 (10%)	9,519 (10%)	294 (18%)
Healthy Diet Score >7	19,600 (21%)	19,303 (21%)	297 (18%)
Fasting, hrs	3 (2, 4)	3 (2, 4)	4 (3, 5)
eGFR, mL/min/1.73 m^2	87 (77, 98)	87 (77, 98)	85 (75, 97)
Alcohol intake, g/day	13 (1, 28)	13 (1, 28)	14 (0, 35)
HbA1c, mmol/mol	34.7 (32.4, 37.0)	34.7 (32.4, 37.0)	35.6 (33.3, 38.1)
Systolic blood pressure, mmHg	135 (124, 148)	135 (124, 148)	144 (132, 156)
Diastolic blood pressure, mmHg	82 (75, 89)	82 (75, 89)	84 (77, 92)
Use of anti-hypertensive medication	10,965 (12%)	10,525 (11%)	440 (27%)

Values are mean (standard deviation) or median (interquartile range) for continuous variables, and number of participants (%) for categorical variables.
Abbreviations: eGFR - estimated glomerular filtration rate, HbA1c - hemoglobin A1c, HDL-C - high-density lipoprotein cholesterol, HF - heart failure,

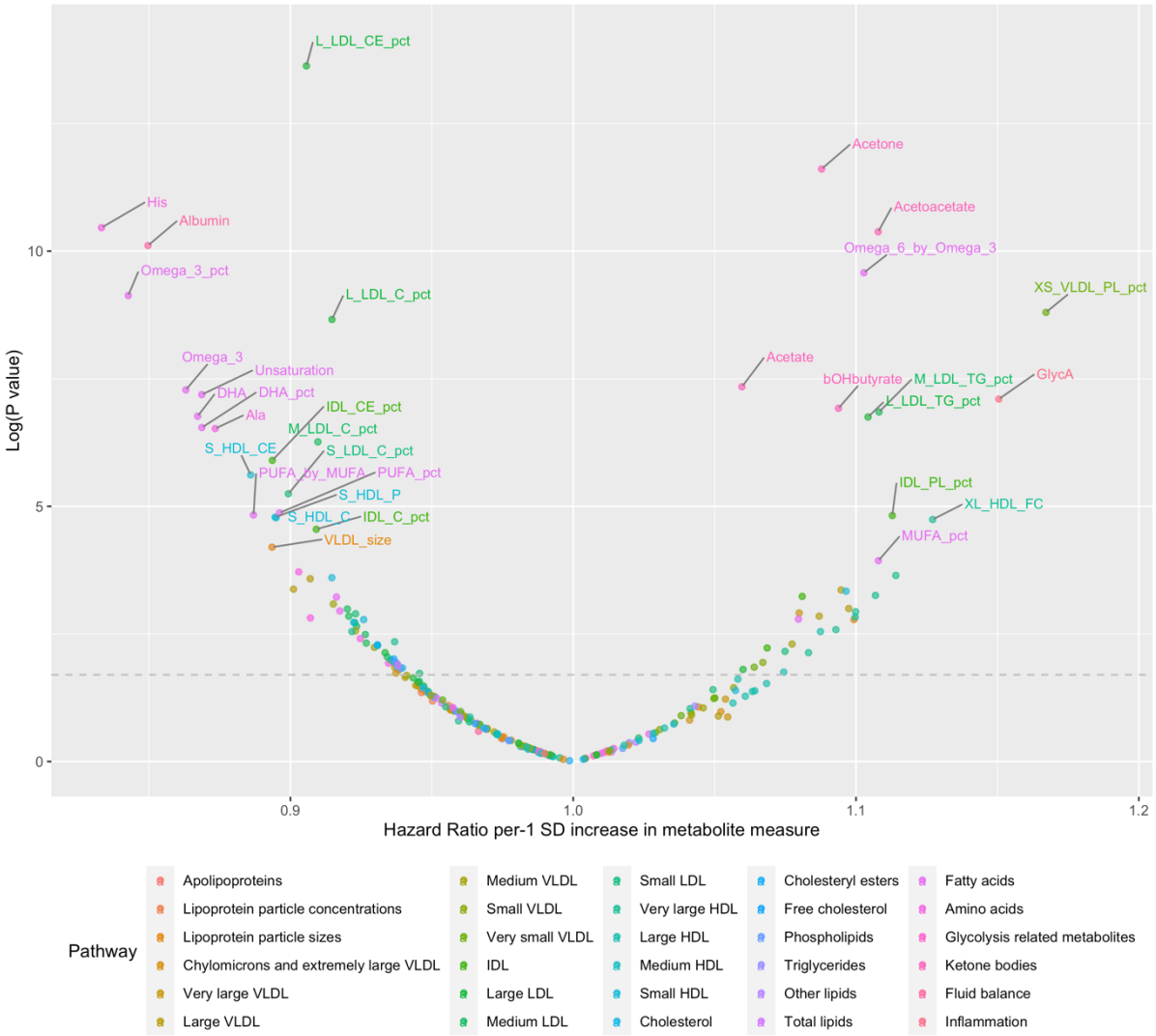
After a median follow-up of 12.1 years (IQR: 11.4-12.8), 1,638 participants were diagnosed with incident HF. When comparing baseline characteristics (**Table 1**), those with incident HF were typically older, predominantly male, had a higher BMI, and were more often of White ethnicity. Additionally, they were more likely to be current smokers, had a poorer diet quality,

longer fasting times, and higher HbA1c levels. Elevated SBP and DBP were also observed among these participants, and there was a greater chance of them being on antihypertensive medication.

Single metabolite markers and heart failure risk

Metabolome-wide analysis with basic covariate adjustment

Figure 3. Volcano plot summarizing associations between 249 metabolite measures with incident heart failure adjusted for age, sex, and body mass index.



Estimates are derived from a Cox proportional hazards model including one metabolite measure at a time and further adjusted for age, sex, and BMI.
Dashed line represents false discovery rate adjusted P (P_{FDR}) of 0.05.
Labeled are metabolite measures with strongest positive and negative associations with the endpoint.

The results of the explorative metabolome-wide analysis of incident HF are visually represented in **Figure 3** as a volcano plot. Out of the 249 metabolite measures initially available in the NMR platform, nearly half (120 measures) were found to be nominally associated with

incident HF. Among these 120 nominally significant measures, approximately one-third (41 measures) showed a positive association with HF risk, while the remaining two-thirds (79 measures) exhibited an inverse association with HF risk.

The magnitude of the nominally significant associations per 1-SD increase in metabolite measures ranged from a HR of 0.83 to 0.95 for measures linked with lower HF risk and from 1.05 to 1.17 for those associated with higher HF risk. Further analysis revealed that out of the 120 nominally significant measures, 97 remained significant even after adjusting the P-values for false discovery rate ($P_{FDR}<0.05$). Notably, these metabolite measures belonged, among others, to amino acids, FAs, ketone bodies, and cholesterol-enriched apoB-containing lipoproteins (very small VLDL, IDL, LDL), as well as small HDL pathways.

In the amino acids class, alanine, histidine, and valine were inversely associated with HF risk, with each 1-SD increase being linked to a 13-17% lower HF risk. Conversely, all metabolite measures related to the ketone bodies pathway were associated with a higher risk of HF. Specifically, a 1-SD increase in levels of acetone, acetate, acetoacetate, and 3-OHB was associated with a 4-6% higher HF risk.

Within the FA class, inverse associations with HF risk were observed for metabolite measures linked with PUFA, especially omega-3 FAs. Concentrations of PUFA, omega-3 FAs, and DHA (both crude and relative to the content of a higher tier trait) were associated with a 10-14% lower HF risk. In contrast, the ratio of omega-6 to omega-3 FAs and the relative content of PUFA were associated with a higher HF risk.

Concerning cholesterol-enriched apoB-containing lipoproteins (very small VLDL, IDL, LDL), the relative overall and esterified cholesterol content (as a percentage of total lipid content per particle) exhibited an inverse trend with HF risk, showing a 9-11% lower risk. On the other hand, the relative content of triglycerides and phospholipids in those lipoproteins was associated with a 10-17% higher HF risk. Additionally, when analyzing HDL particle subclasses, the direction of association varied with size. Concentrations of small HDL particles, overall cholesterol, and esterified cholesterol were associated with a lower HF risk (all 11%), whereas overall cholesterol in very large HDL particles was associated with a higher risk.

The inflammation and fluid balance classes were represented by a single metabolite measure each. Fluid balance pathway-related albumin was inversely associated with HF risk (HR: 0.85,

95%CI: 0.81-0.89, P=7.80 x 10^{-11}), while inflammation pathway-related glycoprotein acetyls were positively associated with HF risk (HR: 1.15, 95%CI: 1.09-1.21, P=7.9 x 10^{-8}).

Metabolome-wide analysis with essential covariate adjustment

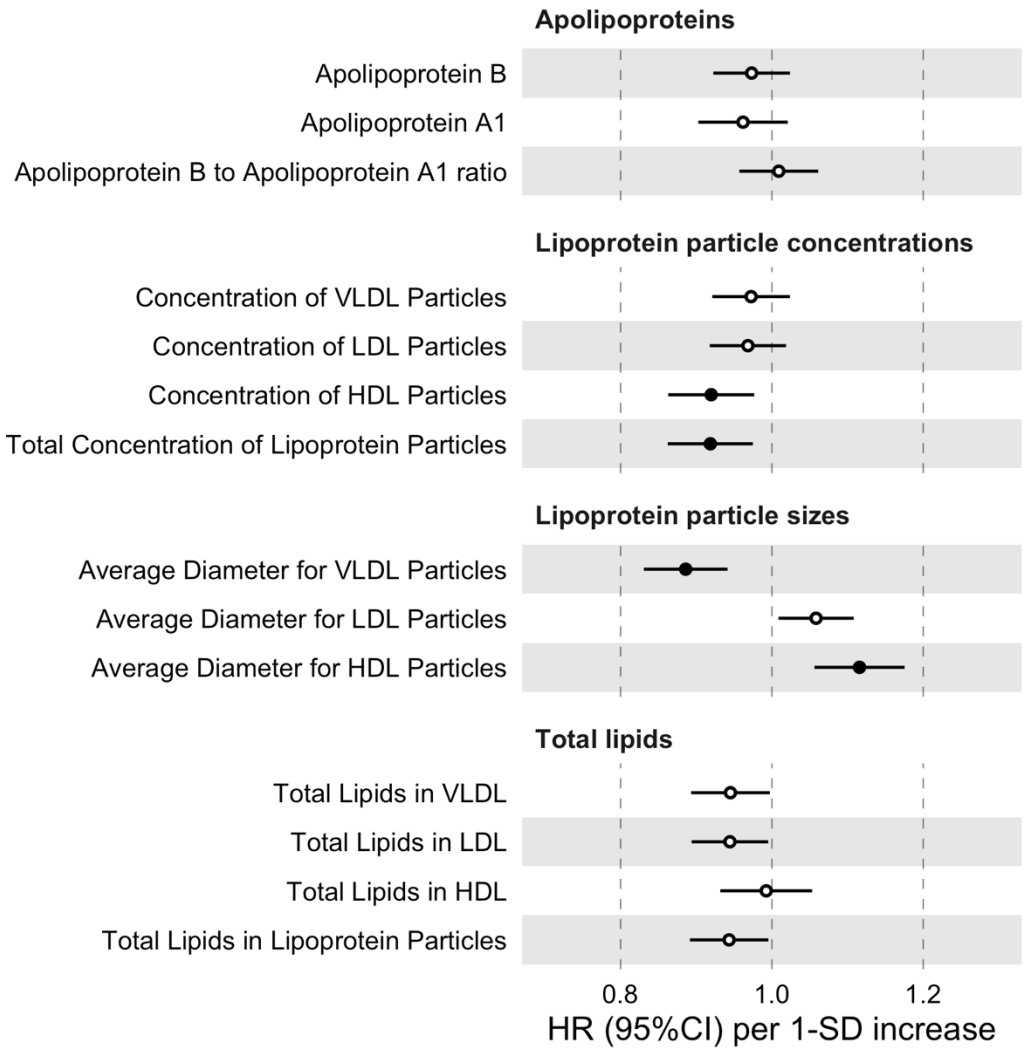
Figures 4-10 provide a summary of the estimates from Cox PH models for incident HF, including one metabolite measure at a time, and subsequently adjusted for an essential covariate set. The adjustment comprised age, sex, BMI, self-reported ethnicity, TDI, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication. Tabularized data supporting above-mentioned figures are available in the Supplementary Appendix.

The association of apolipoproteins, lipoprotein count, size and total lipid content with HF risk are presented in **Figure 4**. In the apolipoproteins class, none of the metabolite measures, namely, apoB (HR: 0.97, 95%CI: 0.93-1.02, P_{FDR}=0.388), apoA1 (HR: 0.96, 95%CI: 0.91-1.02, P_{FDR}=0.278) nor the ratio of apoB to apoA1 (HR: 1.01, 95%CI: 0.96-1.06, P_{FDR}=0.817), exhibited a significant association with incident HF.

Each 1-SD increase in concentration of HDL and total concentration of lipoproteins were both associated with 8% lower HF risk (both 95%CI: 0.87-0.97, P_{FDR} =0.017 and P_{FDR}=0.014 respectively). Conversely, the concentration of VLDL particles (HR: 0.97, 95%CI: 0.92-1.02, P_{FDR}=0.383) and LDL particles (HR: 0.97, 95%CI: 0.92-1.02, P_{FDR}=0.293) were not associated with HF risk.

For the lipoprotein particle sizes class, the average diameter for VLDL particles displayed an inverse association (HR: 0.89, 95%CI: 0.84-0.94, P_{FDR}=2.6 x 10^{-4}), while the average diameter for HDL particles displayed a positive association (HR: 1.12, 95%CI: 1.05-1.18, P_{FDR} = 0.002). Average diameter for LDL particles showed a nominal association with HF risk (HR: 1.06, 95%CI: 1.01-1.11), which became non-significant after multiple-testing-correction (P_{FDR}=0.066).

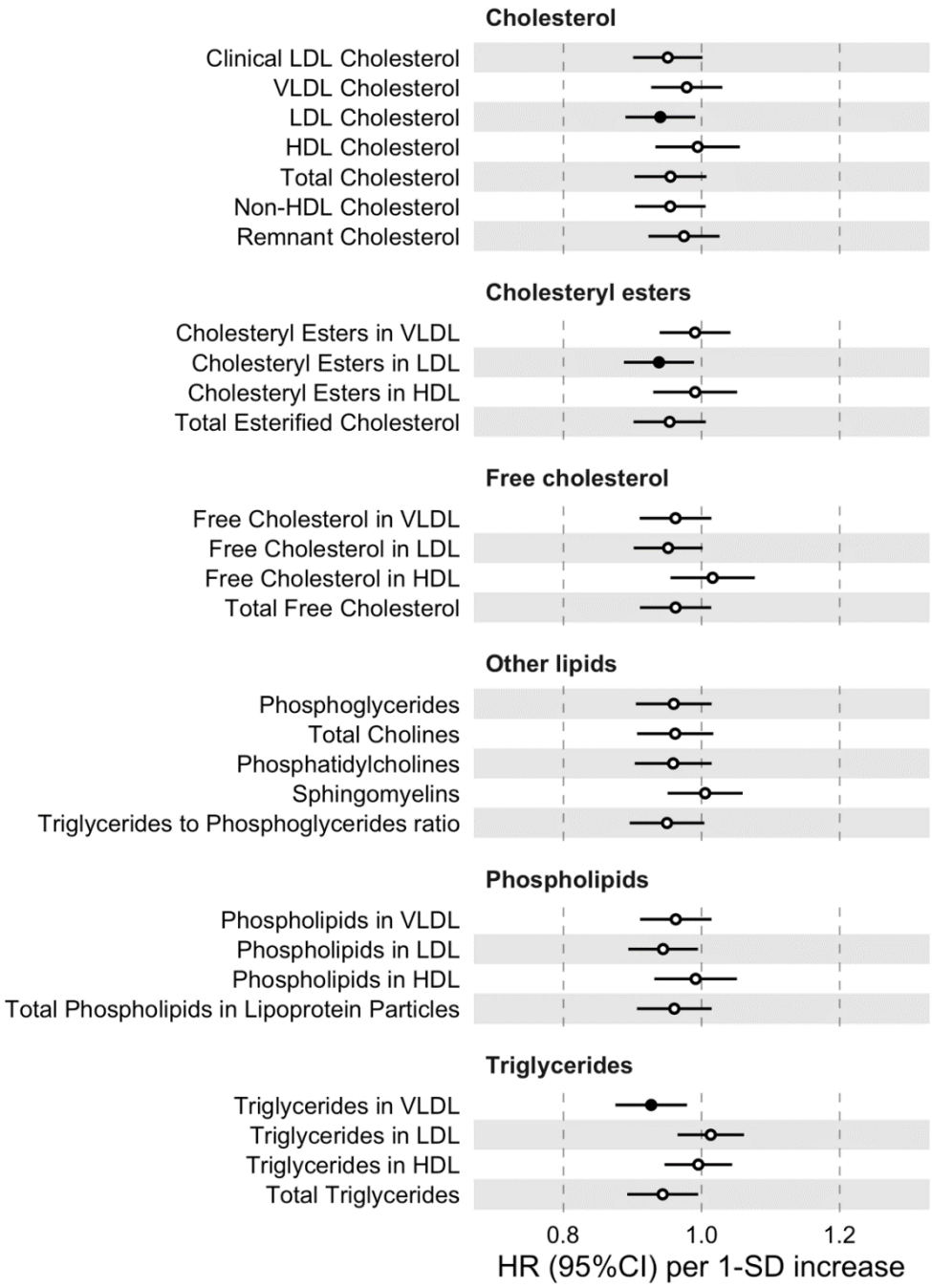
Figure 4. Multivariable-adjusted hazard ratios (95%CI) for heart failure associated with metabolic measures from apolipoprotein, lipoprotein particle concentrations, lipoprotein particle sizes and total lipids classes.



Estimates are derived from a Cox proportional hazards model including one metabolite measure at a time and further adjusted for age, sex, body mass index, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication.
Closed circles indicate metabolite measures significant at false discovery rate adjusted P<0.05, while open those which did not pass this threshold.

Total lipid content across all lipoprotein classes (HR: 0.94, 95%CI: 0.90-0.99), as well as lipid content of VLDL (HR: 0.95, 95%CI: 0.90-1.00) and LDL (HR: 0.94, 95%CI: 0.90-0.99) showed nominally significant inverse associations with incident HF. However, these associations became nonsignificant following multiple testing correction (P_{FDR}=0.068, P_{FDR} =0.079 and P_{FDR}=0.066, respectively). Total lipids in HDL were not linked with HF risk (HR: 0.99, 95%CI: 0.93-1.05, P_{FDR}=0.877).

Figure 5. Multivariable-adjusted hazard ratios (95%CI) for heart failure associated with metabolic measures from overall cholesterol, cholesteryl esters, free cholesterol, other lipids, phospholipids, and triglycerides classes.



Estimates are derived from a Cox proportional hazards model including one metabolite measure at a time and further adjusted for age, sex, body mass index, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication.
Closed circles indicate metabolite measures significant at false discovery rate adjusted $P<0.05$, while open those which did not pass this threshold.

Associations with incident HF for overall cholesterol, cholesteryl esters, free cholesterol, other lipids, phospholipids, and triglycerides are shown in **Figure 5**.

The majority of metabolite measures across lipid-related classes showed a trend towards a lower HF risk. In the overall cholesterol class, clinical (calculated from formula) LDL cholesterol (HR: 0.95, 95%CI: 0.90-1.00) and (direct) LDL cholesterol (HR: 0.94, 95%CI: 0.89-0.99) were associated with a lower HF risk, but the relationship for clinical one was no longer significant following multiple testing correction (P_{FDR}=0.108). No significant association was observed for total, remnant, non-HDL, and HDL cholesterols (P_{FDR}: 0.140 to 0.888).

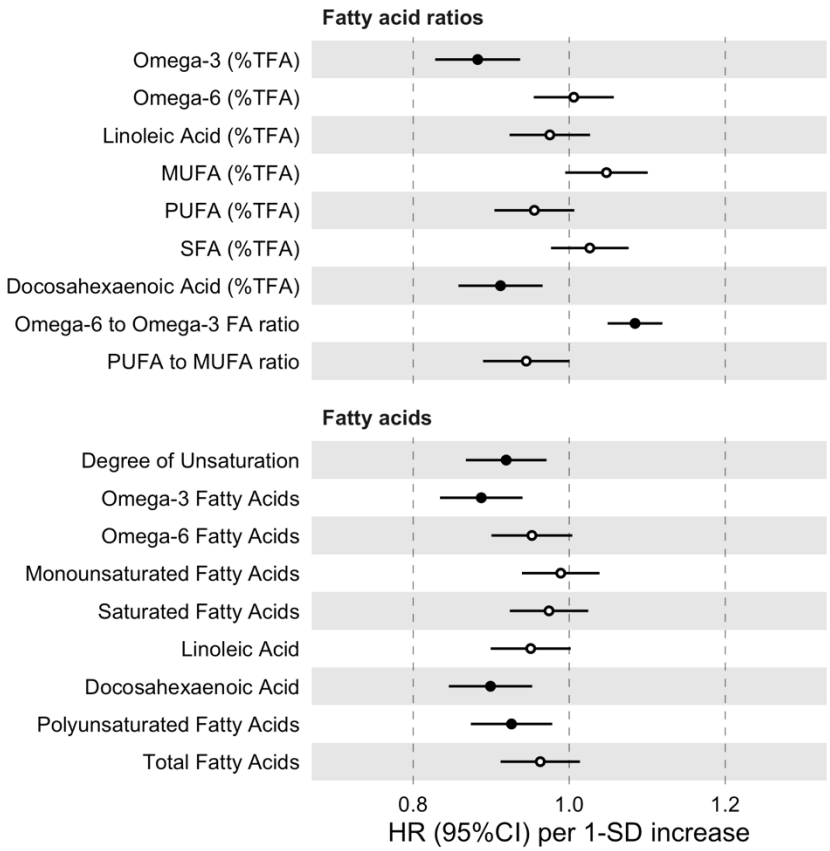
Direction and magnitude of association after looking separately at free and esterified cholesterol was consistent with overall cholesterol across majority of compartments. In LDL particles, esterified (HR: 0.94, 95%CI: 0.89-0.99, P_{FDR}=0.044), but not free cholesterol (HR: 0.95, 95%CI: 0.91-1.00, P_{FDR}=0.108) showed significant link with a lower HF risk. Total free cholesterol and cholesteryl esters, as well as, in VLDL and HDL were equally not related to HF risk (P_{FDR}: 0.140 to 0.841).

None of metabolite measures belonging to other lipids class phosphoglycerides (HR: 0.96, 95%CI: 0.91-1.01, P_{FDR}=0.22), total cholines (HR: 0.96, 95%CI: 0.91-1.02, P_{FDR}=0.244), PCs (HR: 0.96, 95%CI: 0.91-1.01, P_{FDR}=0.218), SMs (HR: 1.01, 95%CI: 0.95-1.06, P_{FDR}=0.888) nor triglycerides to phosphoglycerides ratio (HR: 0.95, 95%CI: 0.9-1, P_{FDR}=0.124) was linked with incident HF.

Only phospholipid content in LDL particles showed a nominal inverse association (HR: 0.94, 95%CI: 0.90-0.99) with incident HF, but it became insignificant following multiple testing correction (P_{FDR}=0.066). Total phospholipids, VLDL and HDL phospholipids were not linked with HF risk (P_{FDR}: 0.222 to 0.853).

Total triglycerides concentration showed a nominally significant inverse association with HF (HR: 0.94, 95%CI: 0.90-0.99), however it did not pass multiple testing correction (P_{FDR}=0.066). One SD increase in VLDL triglycerides was associated with 7% lower HF risk (95%CI: 0.88-0.98, P_{FDR}=0.017). However, triglycerides content in LDL (HR: 1.01, 95%CI: 0.97-1.06, P_{FDR}=0.691) and HDL (HR: 0.99, 95%CI: 0.93-1.05, P_{FDR}=0.853) were not associated with incident HF.

Figure 6. Multivariable-adjusted hazard ratios (95%CI) for heart failure associated with metabolic measures from fatty acids class and their ratios.



Estimates are derived from a Cox proportional hazards model including one metabolite measure at a time and further adjusted for age, sex, body mass index, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication.
Closed circles indicate metabolite measures significant at false discovery rate adjusted P<0.05, while open those which did not pass this threshold.

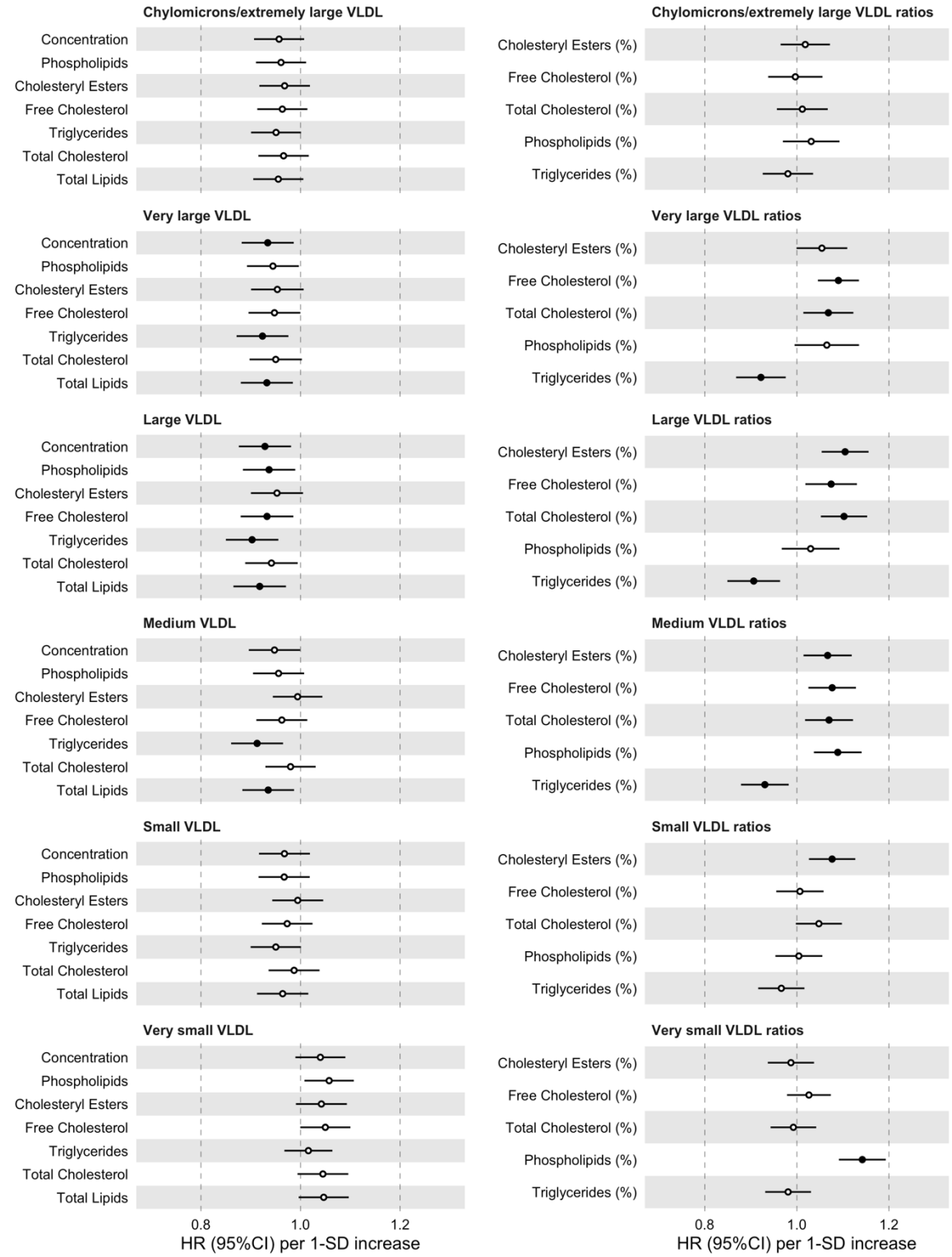
Figure 6 illustrates associations of metabolite measures belonging to FAs and their respective ratios with incident HF. The HF risk tended to decrease with higher degree of FA unsaturation (HR: 0.92, 95%CI: 0.87-0.97, P_{FDR}=0.008). Consistently, concertation of saturated (0.97, 95%CI: 0.93-1.02, P_{FDR}=0.400) and PUFA (HR: 0.99, 95%CI: 0.94-1.04, P_{FDR}=0.763), as well as their relative content as percentage of total FAs (HR: 1.03, 95%CI: 0.98-1.08, P_{FDR}=0.400 and HR: 1.05, 95%CI: 0.99-1.10, P_{FDR}=0.147, respectively) did not affect HF risk. In contrast, PUFA concentration showed an inverse association with HF risk (0.89, 95%CI: 0.84-0.94, P_{FDR}=1.70 x 10^{-4}). Subfamily of PUFA, omega-3 FAs concentration and their relative content (HR: 0.88, 95%CI: 0.84-0.93, P_{FDR}=1.50 x 10^{-4}) were linked with a lower HF risk. However, neither omega-6 FAs concentration (HR: 0.95, 95%CI: 0.90-1.00, P_{FDR}=0.124) nor relative content were associated with incident HF. Higher ratio of omega-6 to omega-3 FAs (HR: 1.08, 95%CI: 1.05-1.12, P_{FDR}=1.40 x 10^{-4}) showed a positive association with HF risk. The only

available individual omega-3 acid, DHA was tied with a lower HF risk (concentration HR: 0.90, 95%CI: 0.85-0.95, P_{FDR}=0.001 and relative content HR: 0.91, 95%CI: 0.86-0.96, P_{FDR}=0.005). No association was observed for the only directly measured omega-6 FA, linoleic acid (concentration HR: 0.95, 95%CI: 0.9-1.00, P_{FDR}=0.110 and relative content HR: 0.98, 95%CI: 0.93-1.03, P_{FDR}=0.431).

Associations of VLDL size subclasses counts and their lipid content and incident HF are presented in **Figure 7**. A general pattern of trend towards a lower HF risk could be observed for classes from chylomicrons and very large VLDL to small VLDL. In contrast, very small VLDL count, and their associated measures showed a trend towards a higher HF risk. Despite a trend towards lower HR risk, none of measures from chylomicrons and extremely large VLDL nor ratios were associated with HF risk (P_{FDR}: 0.108 to 0.918). In large VLDL size subclass, significant inverse associations with HF were found for concentration (HR: 0.93, 95%CI: 0.89-0.98, P_{FDR}=0.036), triglycerides (HR: 0.92, 95%CI: 0.88-0.97, P_{FDR}=0.013) and total lipids (HR: 0.93, 95%CI: 0.88-0.98, P_{FDR}=0.030). Overall cholesterol, cholesteryl esters, free cholesterol and phospholipids did not reach statistical significance (P_{FDR}: 0.075 to 0.140). Similar results were observed for large VLDL subclass with additional associations found for free cholesterol (HR: 0.93, 95%CI: 0.88-0.98, P_{FDR}=0.034) and phospholipids content (HR: 0.94, 95%CI: 0.89-0.99, P_{FDR}=0.047). In medium VLDL size subclass significant inverse associations with HF risk were found for triglycerides and total lipids content. From metabolite measures related to small VLDL particles, only triglyceride (HR: 0.95, 95%CI: 0.90-1.00) content was associated with HF risk but became non-significant after correction for multiple testing (P_{FDR}=0.106). Among very small VLDL lipoproteins, a trend towards increased HF risk was observed for all related measures. However, only phospholipids content reached nominal significance (HR: 1.06, 95%CI: 1.01-1.11, but was no longer significant following multiple testing correction (P_{FDR}=0.066).

Increase in relative triglyceride content at expense of other lipid fractions was linked with lower HF risk (7-9%). Moreover, the relative content of cholesteryl esters in small VLDL (HR: 1.08, 95%CI: 1.02-1.13, P_{FDR}=0.017) and the relative content of phospholipids in very small VLDL (HR: 1.14, 95%CI: 1.09-1.20, P_{FDR}=1.1 x 10^{-5}) both demonstrated a positive association with HF risk.

Figure 7. Multivariable-adjusted hazard ratios (95%CI) for heart failure associated with metabolic measures from VLDL size subclasses.

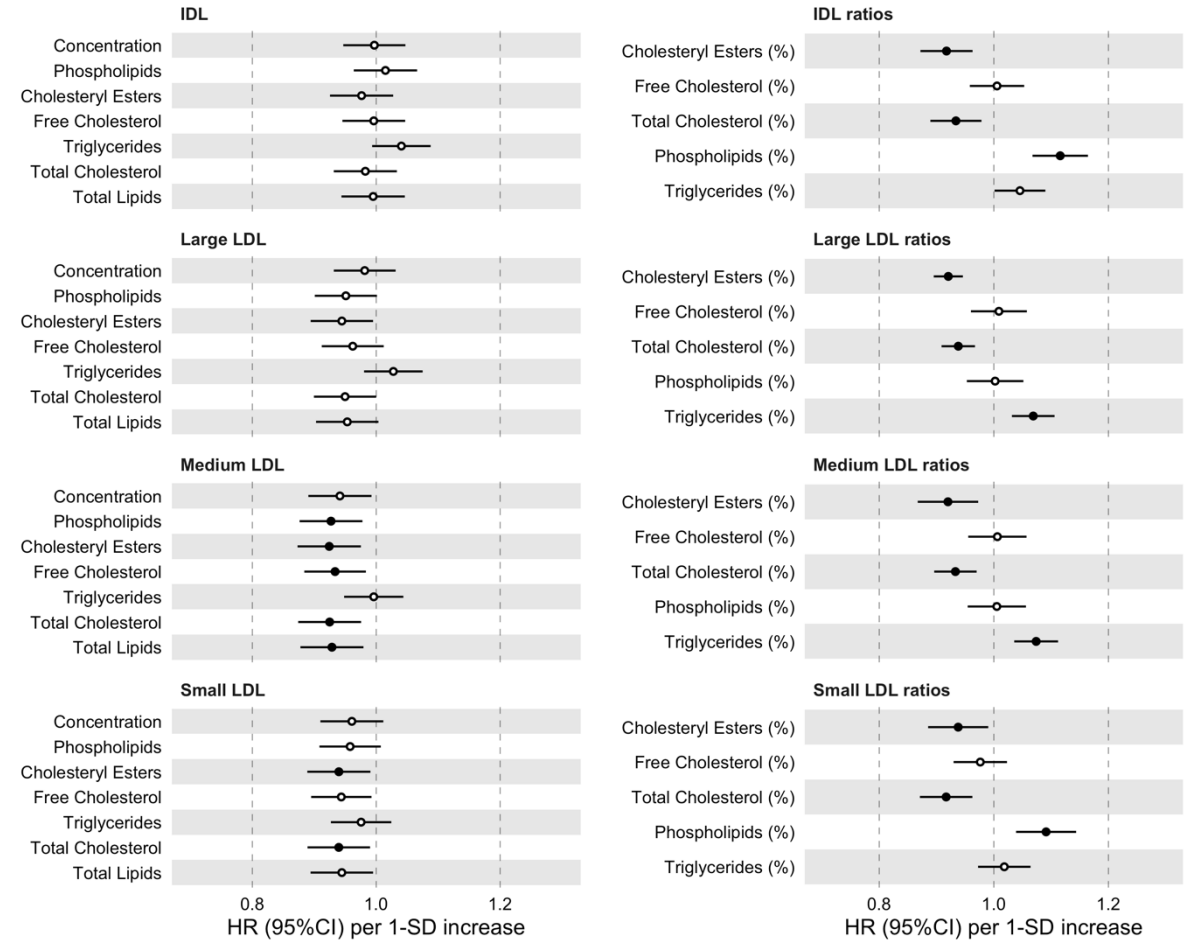


Estimates are derived from a Cox proportional hazards model including one metabolite measure at a time and further adjusted for age, sex, body mass index, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication.
Closed circles indicate metabolite measures significant at false discovery rate adjusted P<0.05, while open those which did not pass this threshold.

The HF risk associated with IDL and LDL size subclass measures is depicted in **Figure 8**. Metabolic measures of the IDL subclass demonstrated no association with incident HF, with the exception of triglycerides, where a trend towards an increased risk was evident. Conversely, metabolic measures across LDL size subclasses predominantly indicated a reduced HF risk. However, the triglycerides content did not appear to influence this risk.

Within the large LDL size subclass, only cholesteryl esters displayed a nominally significant inverse association (HR: 0.94, 95%CI: 0.90-0.99) with incident HF. This association was no longer evident after adjusting for multiple corrections (P_{FDR}=0.066).

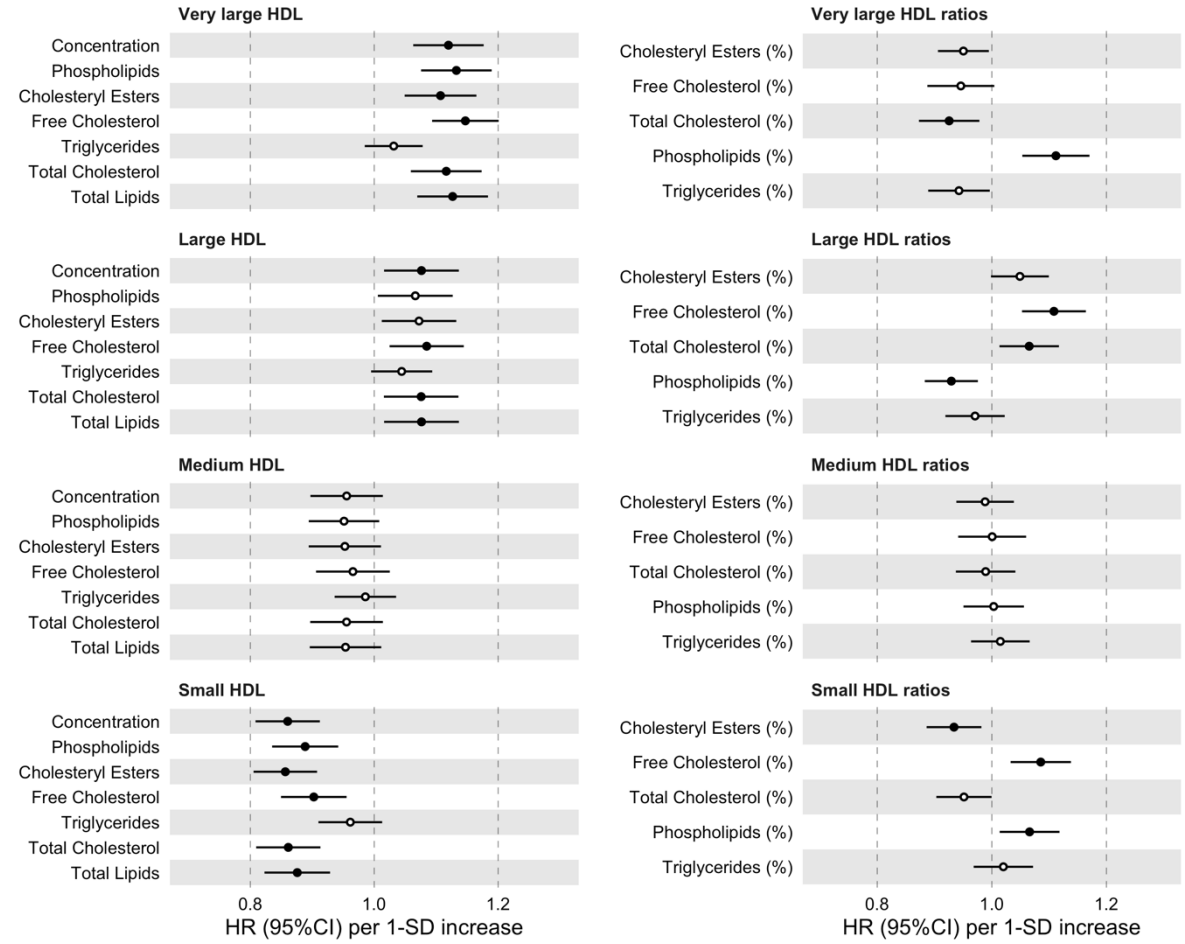
Figure 8. Multivariable-adjusted hazard ratios (95%CI) for heart failure associated with metabolic measures from IDL and LDL size subclasses.



Estimates are derived from a Cox proportional hazards model including one metabolite measure at a time and further adjusted for age, sex, body mass index, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication.
Closed circles indicate metabolite measures significant at false discovery rate adjusted P<0.05, while open those which did not pass this threshold.

For the medium LDL subclass, six out of the seven related measures – namely particle concentration (HR: 0.94, 95%CI: 0.89-0.99), total lipids (HR: 0.93, 95%CI: 0.88-0.98), total cholesterol (HR: 0.92, 95%CI: 0.88-0.97), esterified cholesterol (HR: 0.92, 95%CI: 0.88-0.97), free cholesterol (HR: 0.93, 95%CI: 0.89-0.98), and phospholipids – were associated with a reduced HF risk. Nonetheless, the association with particle concentration ceased to be significant (P_{FDR}=0.056) after multiple testing correction. In the small LDL subclass, a 1-SD increase in the concentration of four metabolite measures (total lipid content, total cholesterol, free cholesterol, and cholesteryl esters) corresponded to a 4% reduction in HF risk. However, only total cholesterol and cholesteryl esters remained significant (both P_{FDR}=0.048) post multiple testing correction.

Figure 9. Multivariable-adjusted hazard ratios (95%CI) for heart failure associated with metabolic measures from HDL size subclasses.



Estimates are derived from a Cox proportional hazards model including one metabolite measure at a time and further adjusted for age, sex, body mass index, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication.
Closed circles indicate metabolite measures significant at false discovery rate adjusted P<0.05, while open those which did not pass this threshold.

However, when lipid fractions were expressed relative to the total lipid content per particle, an increase in the content of both total and esterified cholesterol across all LDL size subclasses was associated with a 6-8% reduced HF risk. Additionally, a higher content of triglycerides in both large and medium LDL particles, as well as an increased content of phospholipids in IDL and small LDL particles, was linked to a higher HF risk.

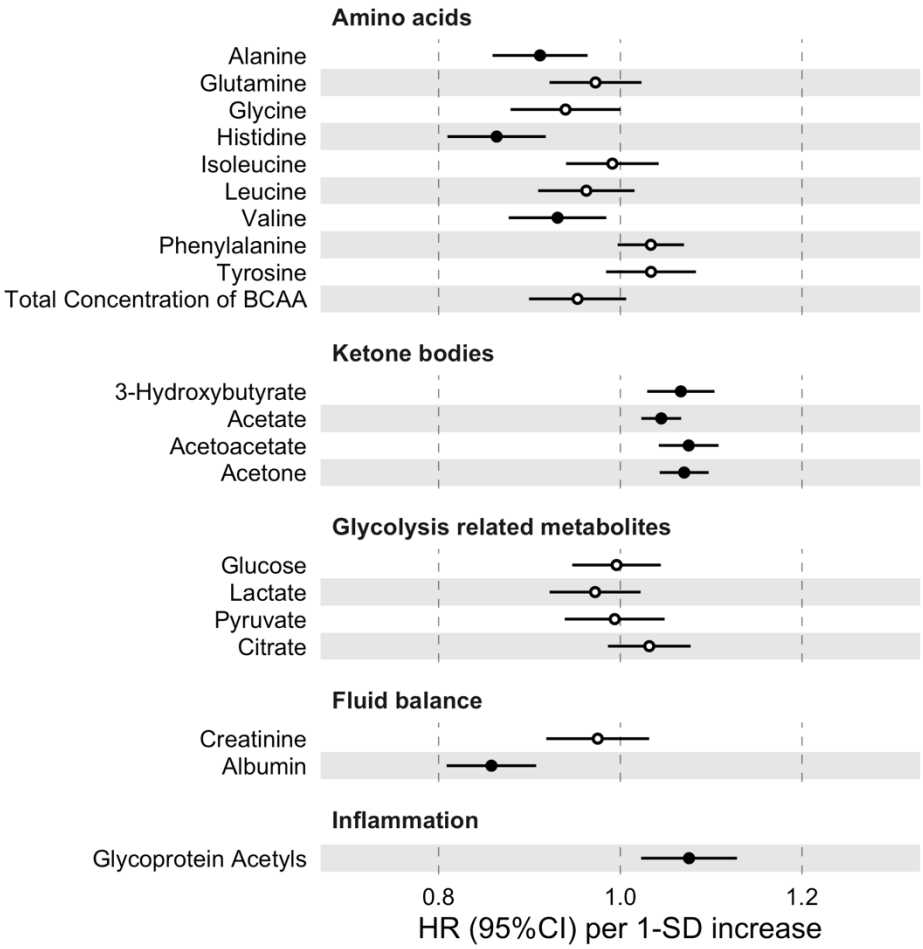
Figure 9 shows relationship between HDL subclasses measures and HF risk. There was a clear pattern where measures related to very large and large HDL particles showed a higher HF risk. On the other hand, measures related to medium and small HDL particles were associated with a lower HF risk. Triglycerides concentration was an exception, as no subclass triglyceride level showed a significant association with the risk.

In very large HDL subclass, significant positive associations with HF were observed for the concentration of particles (HR: 1.12, 95%CI: 1.06-1.19, P_{FDR}=0.001), phospholipids (HR: 1.13, 95%CI: 1.07-1.20, P_{FDR}=2.6 x 10^{-4}), cholesteryl esters (HR: 1.11, 95%CI: 1.04-1.17, P_{FDR}=0.004), free cholesterol (HR: 1.15, 95%CI: 1.09-1.21, P_{FDR}=1.7 x 10^{-5}), cholesterol (HR: 1.12, 95%CI: 1.05-1.18, P_{FDR}=0.001), and total lipids (HR: 1.13, 95%CI: 1.06-1.19, P_{FDR}=5.6 x 10^{-4}). Among large HDL particles, particle concentration (HR: 1.08, 95%CI: 1.01-1.14, P_{FDR}=0.048), total cholesterol (HR: 1.08, 95%CI: 1.01-1.14, P_{FDR}=0.048), free cholesterol (HR: 1.08, 95%CI: 1.02-1.15, P_{FDR}=0.028), and total lipids (HR: 1.08, 95%CI: 1.01-1.14, P_{FDR}=0.048) were linked to a higher HF risk. Despite a clear trend towards lower HF risk with higher concentration of medium HDL subclass-related measures, none of them was statistically significant (P_{FDR} 0.147 to 0.674). For the small HDL particles class, significant inverse associations were observed for concentration of particles (HR: 0.86, 95%CI: 0.82-0.91, P_{FDR}=1.0 x 10^{-6}), total cholesterol (HR: 0.86, 95%CI: 0.82-0.91, P_{FDR}=1.0 x 10^{-6}), cholesteryl esters (HR: 0.86, 95%CI: 0.81-0.9, P_{FDR}=4.0 x 10^{-7}), free cholesterol (HR: 0.90, 95%CI: 0.86-0.95, P_{FDR}=0.001), , and total lipids (HR: 0.88, 95%CI: 0.83-0.92, P_{FDR}=2.3 x 10^{-5}).

In very large HDL subclass, the relative content of cholesteryl esters was inversely associated with HF risk, while phospholipids to total lipids exhibited a positive. For large HDL particles, the free cholesterol to total lipids percentage indicated a positive HF risk association contrasting with an inverse association for phospholipids to total lipids. The medium HDL particles revealed no significant lipid composition percentages in relation to HF risk. In small HDL particles, cholesteryl esters to total lipids showed an inverse association with HF, while

both free cholesterol to total lipids and phospholipids to total lipids displayed positive HF risk associations.

Figure 10. Multivariable-adjusted hazard ratios (95%CI) for heart failure associated with metabolic measures from amino acids, ketone bodies, glycolysis-related, fluid balance and inflammation metabolite measures.



Estimates are derived from a Cox proportional hazards model including one metabolite measure at a time and further adjusted for age, sex, body mass index, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication.
Closed circles indicate metabolite measures significant at false discovery rate adjusted P<0.05, while open those which did not pass this threshold.
Creatinine model did not include adjustment for eGFR.

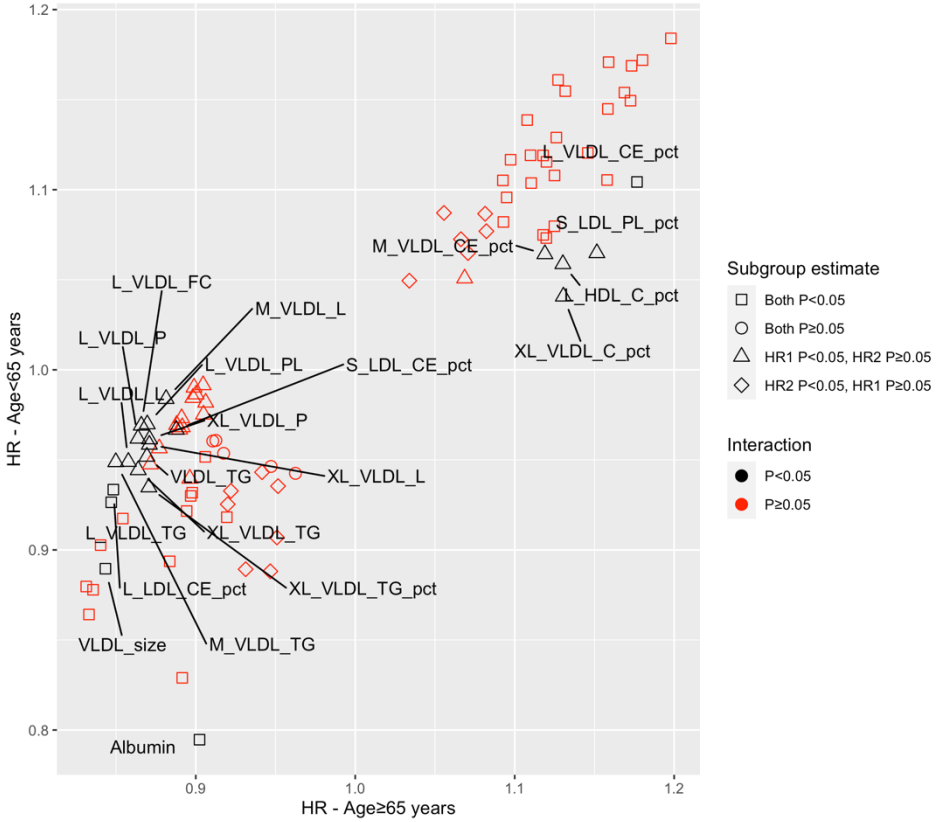
Associations between non-lipid-related metabolic measures and HF risk were presented in **Figure 10**. Among BCAAs, inverse association with incident HF was observed for valine (HR: 0.93, 95% CI: 0.88-0.98, P_{FDR}=0.031), but not for isoleucine (HR: 0.99, 95% CI: 0.94-1.04, P_{FDR}=0.817), leucine (HR: 0.96, 95% CI: 0.91-1.02, P_{FDR}=0.236) or total concentration of BCAA (HR: 0.95, 95% CI: 0.90-1.01, P_{FDR}=0.140). Both histidine and alanine were linked with lower HF risk (HR: 0.86, 95% CI: 0.82-0.91, P_{FDR}=4.8 x 10^{-6} and HR: 0.91, 95% CI: 0.87-

0.96, P_{FDR}=0.004, respectively). Both aromatic amino acids, phenylalanine, and tyrosine, showed nonsignificant association with HF risk (HR 1.03, 95% CI: 1.00-1.07, P_{FDR}=0.14 and HR: 1.03, 95% CI: 0.98-1.09, P_{FDR}=0.269, respectively). There was a nominally significant inverse association with HF for glycine (HR: 0.94, 95% CI: 0.88-1.00), which was no longer observed following correction for multiple testing (P_{FDR}=0.099). Regarding fluid balance, higher concentration of albumin was associated with a lower HF risk (HR: 0.86, 95% CI: 0.82-0.9, $P_{FDR}=2.9 \times 10^{-7}$). None of glycolysis-related metabolite measures was linked with HF risk (P_{FDR} 0.258 to 0.892). Higher concentration of glycoprotein acetyls belonging to the inflammation was associated with a higher HF risk (HR: 1.08, 95% CI: 1.02-1.13, P_{FDR}=0.025) All four metabolic measures belonging to ketone bodies class 3-OHB (HR: 1.07, 95% CI: 1.03-1.11, P_{FDR}=0.004), acetate (HR: 1.05, 95% CI: 1.02-1.07, P_{FDR}=0.001), acetone (HR: 1.07, 95% CI: 1.04-1.10, $P_{FDR}=2.2 \times 10^{-5}$), and acetoacetate (HR: 1.08, 95% CI: 1.04-1.11, $P_{FDR}=2.4 \times 10^{-4}$) showed a positive associations with HF risk.

To summarize, 90 metabolite measures showed significant association with incident HF after correction for multiple comparison testing. These included lipoprotein particle sizes (average diameter for VLDL particles, average diameter for HDL particles), FAs (degree of unsaturation, omega-3 FAs, DHA, PUFA), amino acids (alanine, histidine, valine), ketone bodies (3-OHB, acetate, acetoacetate, acetone), fluid balance (albumin), inflammation (glycoprotein acetyls), very large VLDL (concentration particles, triglycerides, total lipids), large VLDL (concentration particles, phospholipids, free cholesterol), medium VLDL (triglycerides, total lipids), medium LDL (phospholipids, cholesteryl esters, free cholesterol), small LDL (cholesteryl esters, cholesterol in), very large HDL (concentration of particles, phospholipids, cholesteryl esters), large HDL (concentration of particles, free cholesterol), small HDL (concentration of particles, phospholipids, cholesteryl esters), cholesterol (LDL cholesterol, cholesterol in medium LDL), lipoprotein particle concentrations (concentration of HDL particles, total concentration of lipoprotein particles), triglycerides (triglycerides in VLDL), cholesteryl esters (cholesteryl esters in LDL), and 37 relative metabolite measures including lipoprotein and FA ratios.

Subgroup analyses for 90 heart failure related metabolite measures

Figure 11. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by baseline age of participants.

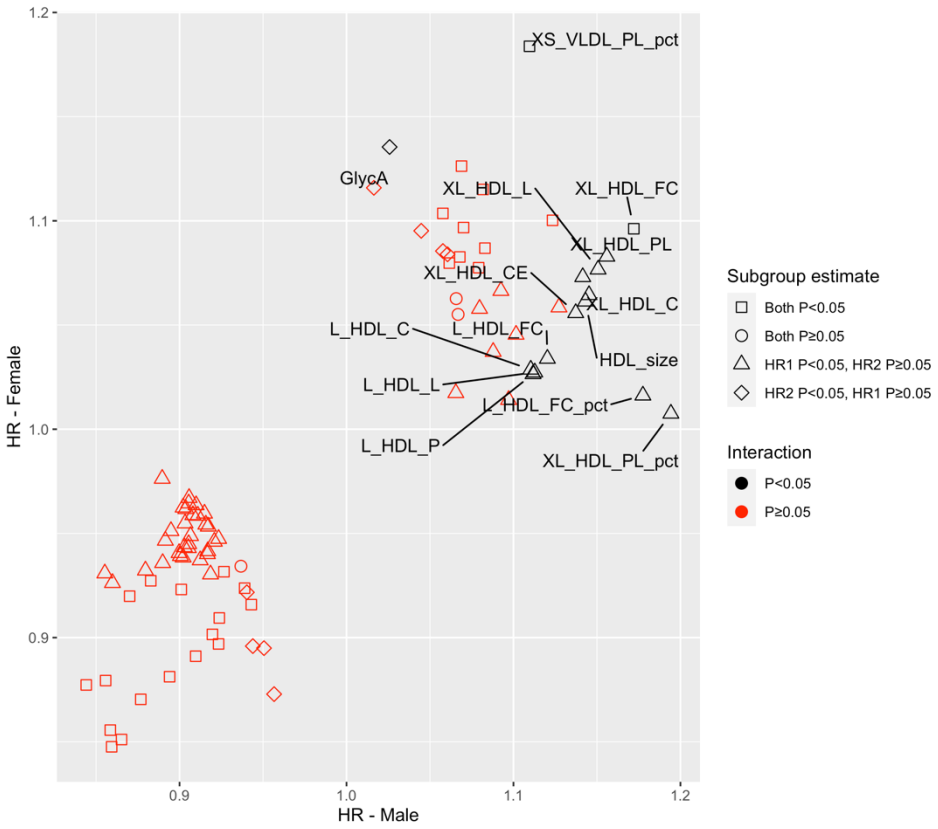


Stratified estimates (HR) are derived from a multivariable-adjusted Cox proportional hazards model excluding age.
P for interaction was calculated for multiplicative interaction term between a given metabolite and age subgroup.
Labelled are top 10% metabolite measures with lowest P for interaction.

Baseline age-stratified analyses for 90 metabolite measures predictive for HF are presented in **Figure 11**. Of 79,957 participants aged <65 years and 13,965 participants aged ≥65 at baseline, respectively 1064 (1.3%) and 574 (4.1%) developed incident HF during follow-up. Of 90 metabolite measures, 21 showed a significant interaction with a baseline age of participants (P<0.05). These measures mostly were related to extremely large, very large and large VLDL particle counts and lipid fractions, but also average diameter for VLDL, triglyceride content, albumin as well as relative cholesterol content across various lipoprotein subclasses. Significant inverse association for HF observed with largest size subclasses (10-16% lower HF risk) were present in participants aged ≥65 years, but this association was attenuated in younger participants. Similarly, relative cholesterol content measures were linked to a higher HF risk (12-18% higher HF risk) in older but not in younger participants. Inverse associations for

albumin and average VLDL diameter were replicated in subgroups but differed in magnitude (<65 years vs. ≥65 years: 0.79 vs. 0.90 for albumin and 0.89 vs. 0.84).

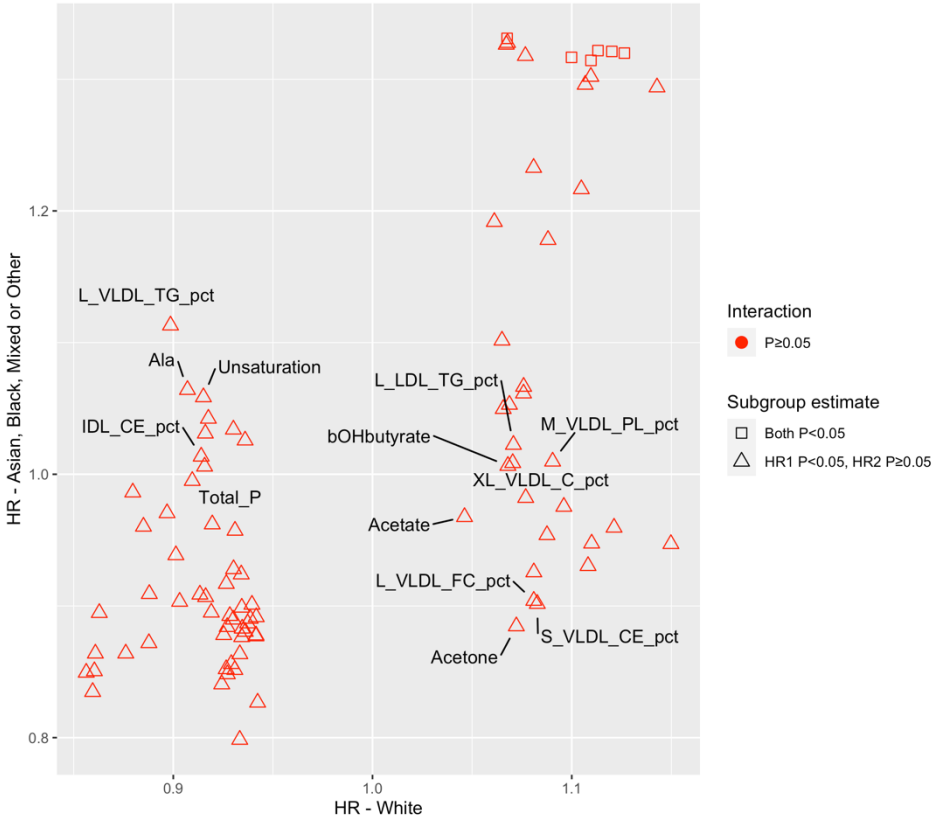
Figure 12. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' sex.



Stratified estimates (HR) are derived from a multivariable-adjusted Cox proportional hazards model excluding sex.
P for interaction was calculated for multiplicative interaction term between a given metabolite and sex subgroup.
Labelled are top 10% metabolite measures with lowest P for interaction.

Sex-stratified analyses for significant metabolite measures-HF associations are presented in **Figure 12**. 912 out of 39,699 male (1.7%) and 717 out of 54,223 female (1.3%) UK Biobank participants developed incident HF during follow-up. Fifteen of 90 metabolite measures showed inconsistent associations with HF between male and female participants. Most differences were observed for very large and large HDL metabolite measures, with increasing concertation of both particles and lipid fractions (total lipids, cholesterol) linked with a higher HF risk in men (11-17% higher HF risk), but not in women, with exception for very large HDL cholesterol associated in both sexes (male vs. female: 1.17 vs. 1.10). Glycoprotein acetyls were associated with a higher HF risk in female (HR: 1.14, 95%CI: 1.05-1.23, P=0.002), but not in male participants HR: 1.03, 95%CI: 0.96-1.10, P=0.471).

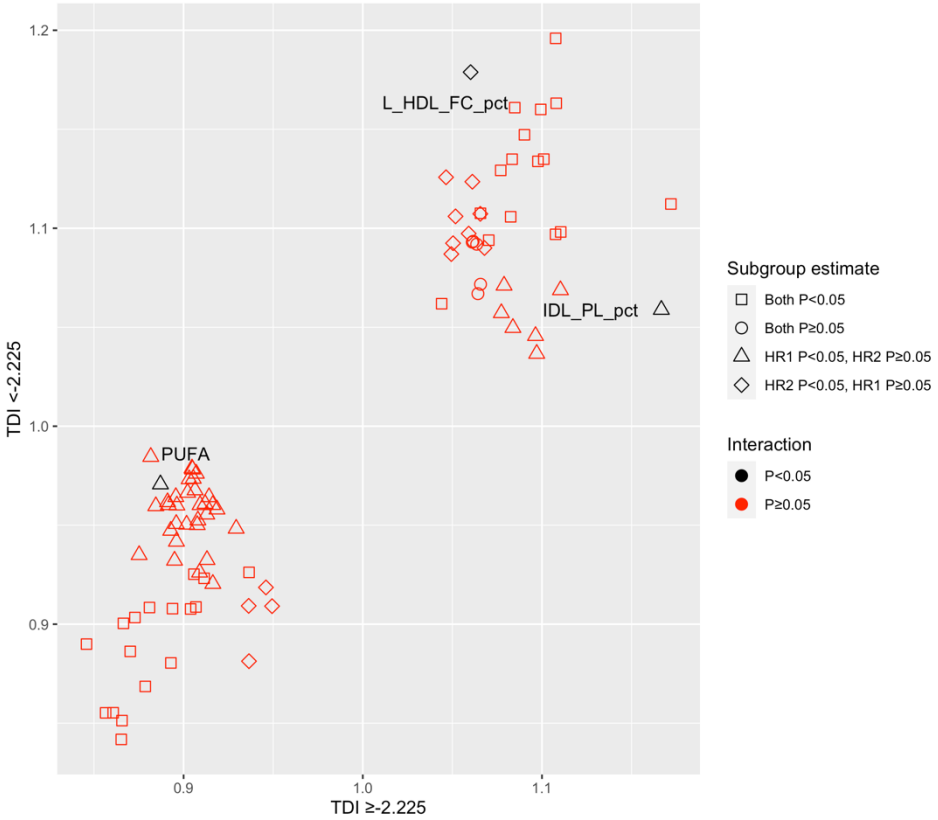
Figure 13. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' self-reported ethnicity.



Stratified estimates (HR) are derived from a multivariable-adjusted Cox proportional hazards model excluding self-reported ethnicity.
P for interaction was calculated for multiplicative interaction term between a given metabolite and self-reported ethnicity subgroup.
Labelled are top 10% metabolite measures with lowest P for interaction.

Subgroup analyses stratified for self-reported ethnicity of participants are presented in **Figure 13**. Among 89,270 participants of White ethnicity, 1584 participants developed incident HF. There were 54 incident HF cases in 4652 participants of non-White ethnicities (20 cases in 1854 Asian individuals, 20 cases in 1406 Black individuals and 14 cases in 1392 individuals of mixed or other ethnicities). None of 90 metabolite measures showed a significant interaction with dichotomized (White vs. Asian, Black, Mixed/Other), however these models were likely highly underpowered due to very small number of incident HF cases in non-White individuals. Therefore, with exception for six metabolite measures related very large HDL, majority of associations did not reach statistical significance. Most heterogenous estimates (HR>1 for White individuals and HR<1 for other ethnicities or opposite) were observed for metabolite measures from FAs (degree of unsaturation), ketone bodies (all four), inflammation (glycoprotein acetyls), very large, large VLDL and IDL (relative lipid content). However, due to low number of participants these differences might be due to imprecision.

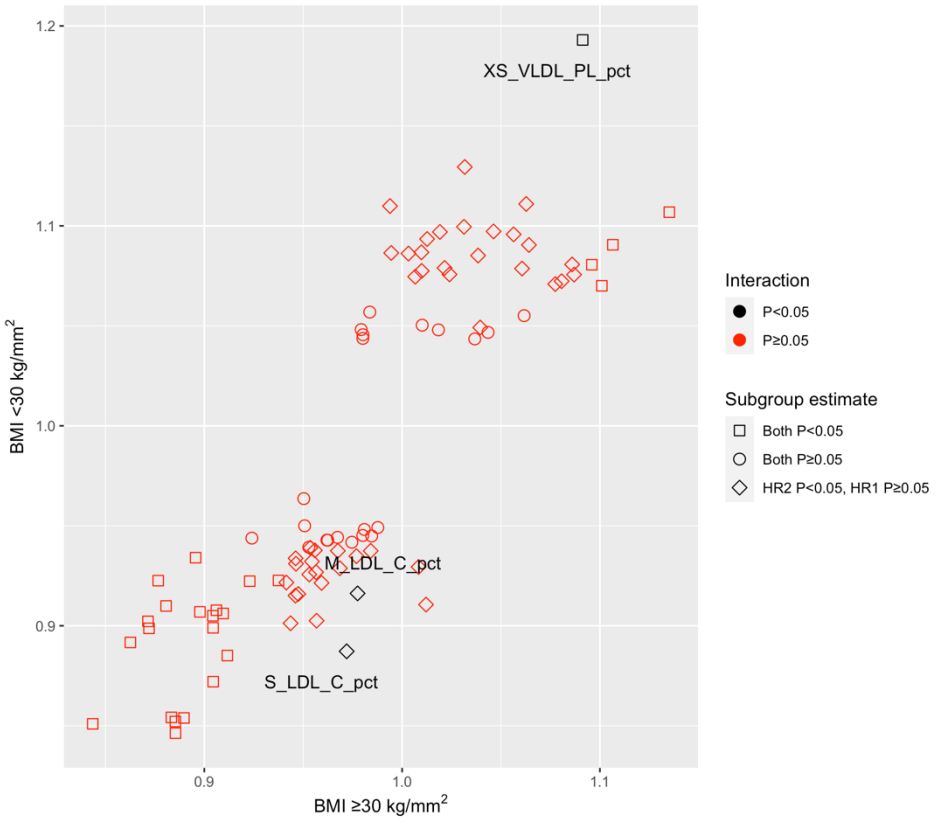
Figure 14. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' baseline deprivation index.



Stratified estimates (HR) are derived from a multivariable-adjusted Cox proportional hazards model excluding age.
P for interaction was calculated for multiplicative interaction term between a given metabolite and age subgroup.
Labelled are top 10% metabolite measures with lowest P for interaction.

Figure 14 presents analyses for 90 significant metabolite measures stratified for median baseline deprivation index. There were respectively 742 and 896 incident HF cases in groups below and above median TDI (median: -2.225 – 46,961 participants <median and ≥median). Of 90 compared metabolite measures, 87 showed consistent associations with no significant interaction with socioeconomic status approximated by deprivation index. In contrast, three measures – PUFA, relative content of phospholipids in IDL and relative content of free cholesterol in large HDL showed differential association with incident HF dependent on baseline deprivation score. Higher concentration of PUFA was associated with HF risk in less but not in more deprived individuals. In contrast, free cholesterol content in large HDL was linked with higher HF risk in more, but not in less deprived group.

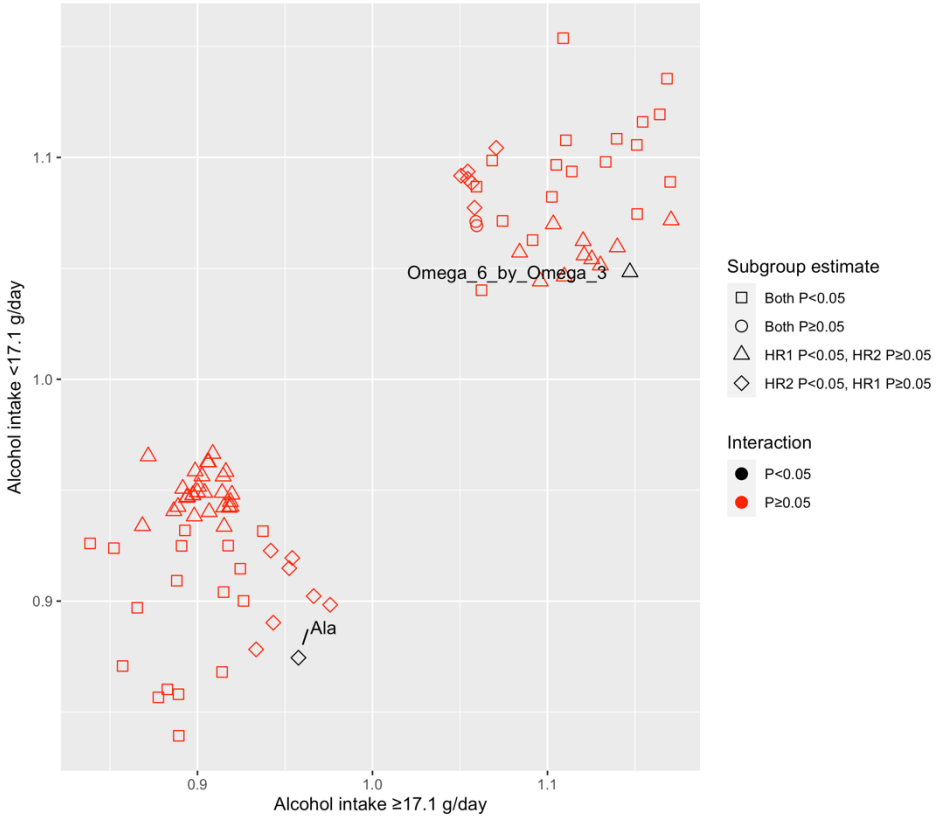
Figure 15. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' baseline body mass index.



Stratified estimates (HR) are derived from a multivariable-adjusted Cox proportional hazards model excluding age.
P for interaction was calculated for multiplicative interaction term between a given metabolite and age subgroup.
Labelled are top 10% metabolite measures with lowest P for interaction.

Subgroup analyses stratified for baseline BMI are presented in **Figure 15**. There were 74,429 participants who had normal weight or overweight (BMI <30 kg/m^2) and 19,493 participants who had obesity at baseline (BMI ≥30 kg/m^2). Among them respectively, 1094 (1.5%) and 544 (2.8%) participants developed incident HF during follow-up. Eighty-seven metabolite measures showed consistent associations with HF risk in both groups, while in case of three a significant association with BMI was observed (P<0.05). Relative content of cholesterol in medium (HR: 0.92, 95%CI: 0.88-0.95, P=4.0 x 10^{-6}) and small LDL (HR: 0.89, 95%CI: 0.83-0.94, P=2.0 x 10^{-5}) particles were associated with lower HF risk in participants who had normal weight and overweight, but not in those with obesity. Association with relative content of phospholipids in very small VLDL was more pronounced among normoweight and overweight participants.

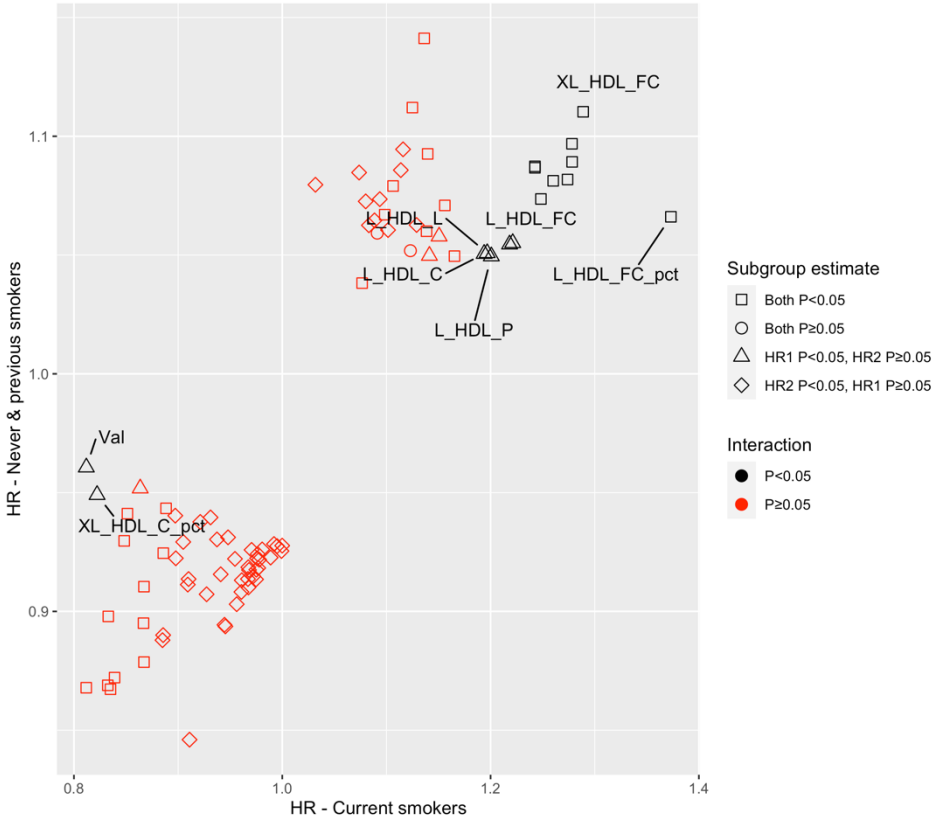
Figure 16. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' baseline alcohol intake.



Stratified estimates (HR) are derived from a multivariable-adjusted Cox proportional hazards model excluding age.
P for interaction was calculated for multiplicative interaction term between a given metabolite and age subgroup.
Labelled are top 10% metabolite measures with lowest P for interaction.

Figure 16 presents metabolite-HF associations stratified for participants baseline alcohol intake. At baseline, 55,057 participants declared alcohol intake <17.1 g/day and 38,865 participants declared alcohol intake ≥17.1 g/day. Respectively, 901 and 737 of them developed incident HF during follow-up. Two metabolite measures showed significant interaction with baseline alcohol intake. Higher concentration of alanine was associated with a lower HF risk in a lower (HR: 0.87, 95%CI: 0.81-0.94, P=2.4 x 10^{-4}), but not in a higher alcohol intake group (HR: 0.96, 95%CI: 0.89-1.03, P=0.268). Contrarily, ratio of omega 6 to omega 3 FAs was linked with a higher HF risk in higher intake (HR: 1.15, 95%CI: 1.09-1.21, P=8.8 x 10^{-7}), but not in a lower intake group (HR: 1.05, 95%CI: 0.99-1.11, P=0.094).

Figure 17. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' baseline smoking status.

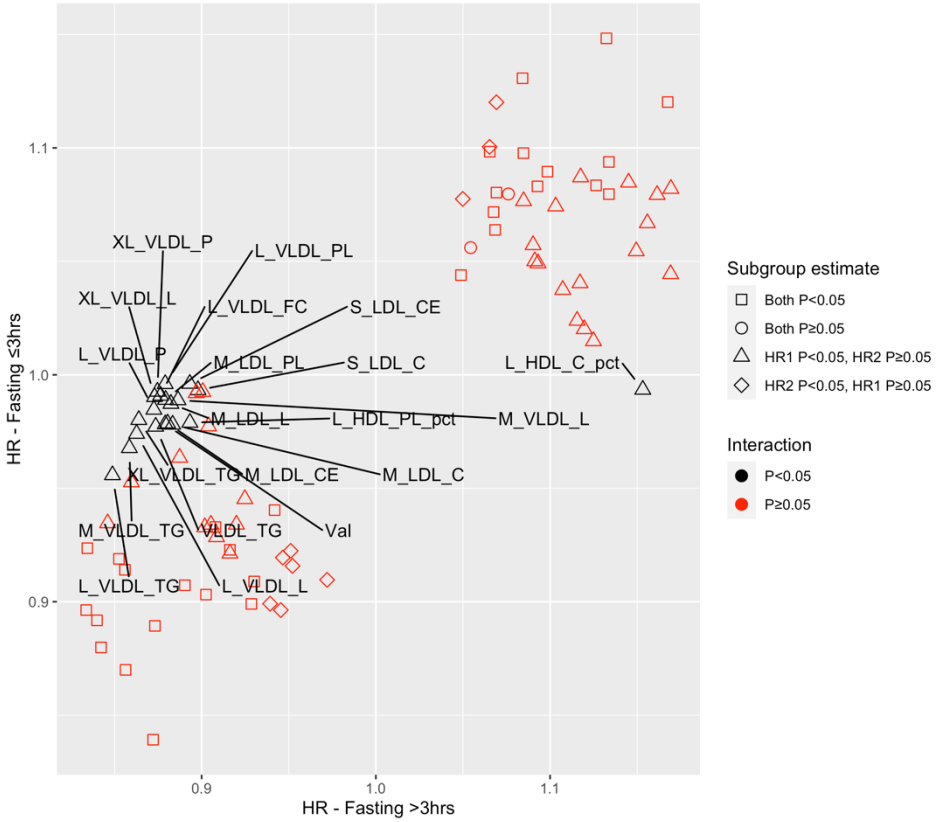


Stratified estimates (HR) are derived from a multivariable-adjusted Cox proportional hazards model excluding age.
P for interaction was calculated for multiplicative interaction term between a given metabolite and age subgroup.
Labelled are top 10% metabolite measures with lowest P for interaction.

Subgroup analyses stratified for smoking status are presented in **Figure 17**. There were 9813 declared current smokers and 84,109 previous or never smokers at the study baseline. Of them, respectively 1344 and 294 were diagnosed with HF during follow-up. Sixteen metabolite measures showed significant interaction (P<0.05) with dichotomized smoking status (current smokers vs. non-smokers). Lipid content (HR: 1.20, 95%CI: 1.05-1.36, P=0.006), cholesterol (HR: 1.19, 95%CI: 1.05-1.35, P=0.006), free cholesterol (HR: 1.22, 95%CI: 1.08-1.38, P=0.002), particle count of large HDL (HR: 1.20, 95%CI: 1.06-1.36, P=0.005) and relative free cholesterol in small HDL particles (HR: 1.22, 95%CI: 1.09-1.36, P=3.9 x 10^{-4}) was associated with a higher HF risk in current smokers, but not in non-smokers. Similarly, inverse associations for HF with higher levels of valine (HR: 0.81, 95%CI: 0.71-0.93, P=0.003) and relative content of cholesterol in very large HDL (HR: 0.82, 95%CI: 0.72-0.94, P=0.004) were replicated only in current smokers. Metabolite measures related with very large HDL particle (particle count, lipid content, cholesterol, cholesteryl esters, free cholesterol and phospholipids

and relative phospholipid content), average diameter for HDL and free cholesterol content in large HDL particles presented stronger positive associations with HF risk in the current smokers (24-37% higher HF risk) than in non-smokers (7-11% higher HF risk).

Figure 18. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' fasting time during sample collection.

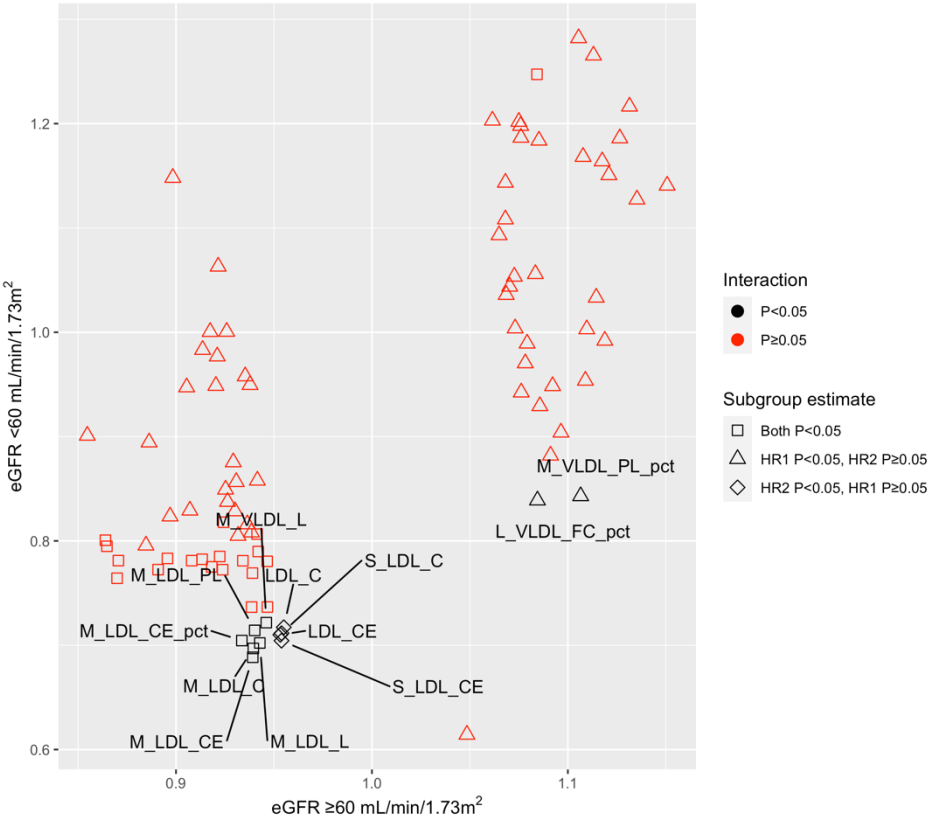


Stratified estimates (HR) are derived from a multivariable-adjusted Cox proportional hazards model excluding age.
P for interaction was calculated for multiplicative interaction term between a given metabolite and age subgroup.
Labelled are top 10% metabolite measures with lowest P for interaction.

Associations between significant metabolite measures and incident HF stratified by fasting time are presented in **Figure 18**. Of 52,513 participants with fasting time during sample collection <3 hours and 41,409 participants who fasted ≥3 hours, respectively 786 (1.5%) and 852 (42.1%) developed incident HF before the end of follow-up. Twenty of 90 metabolite measures showed a significant interaction with participants fasting time (P<0.05). Metabolic measures related to very large VLDL (particle count, lipid content, triglycerides), large VLDL (particle count, lipid content, phospholipids, free cholesterol and triglycerides), medium VLDL (lipid content, triglycerides), medium LDL (lipid content, cholesterol, cholesteryl esters and phospholipids), small LDL (cholesterol, cholesteryl esters), VLDL triglycerides, valine, as well

as relative content of phospholipids in large HDL particles showed an inverse association (10-15% lower HF risk) with incident HF among participants with longer fasting time, but not in those with shorter fasting time during sample collection. Similarly, relative content of cholesterol in large HDL was linked with a higher HF risk in longer (HR: 1.15, 95%CI: 1.07-1.24, P=0.038), but not in shorter fasting time group (HR: 0.99, 95%CI: 0.93-1.07, P=0.857).

Figure 19. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' baseline estimated glomerular filtration rate.

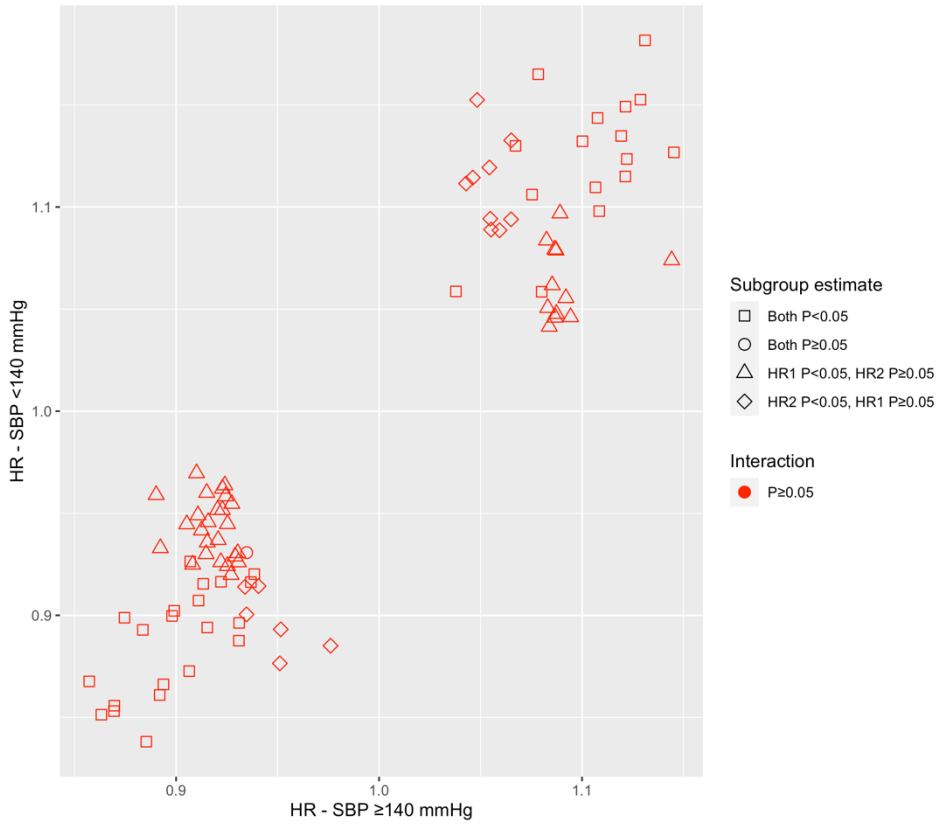


Stratified estimates (HR) are derived from a multivariable-adjusted Cox proportional hazards model excluding age.
P for interaction was calculated for multiplicative interaction term between a given metabolite and age subgroup.
Labelled are top 10% metabolite measures with lowest P for interaction.

Figure 19 presents analyses for 90 HF-associated metabolite measures stratified for estimated eGFR. Among 91,428 participants with preserved kidney function at baseline (≥60 mL/min/1.73m^2), there were 1546 incident HF cases (1.7%). Decreased kidney function (<60 mL/min/1.73m^2) was observed in 2494 participants of which 92 had incident HF (3.7%). Twelve of 90 metabolite measures showed significant interaction with estimated eGFR. Concentration of total lipids, total cholesterol, cholesteryl esters and phospholipids in medium LDL particles showed more profound association with HF in the group with decreased (28-

31% lower HF risk) then preserved kidney function (3-5% lower HF risk). Higher relative content of free cholesterol in large VLDL (HR: 1.08, 95%CI: 1.03-1.15, P=0.005) and phospholipids in medium VLDL particles (HR: 1.11, 95%CI: 1.05-1.17, P=2.0 x 10^{-4}) were associated with HF risk in preserved, but not in decreased kidney function group. An opposite pattern of association was observed for LDL cholesterol HR: 0.72, 95%CI: 0.56-0.92, P=0.003) and cholesteryl esters (HR: 0.71, 95%CI: 0.55-0.92, P=0.002), as well as small LDL cholesterol (HR: 0.71, 95%CI: 0.56-0.91, P=0.002) and cholesteryl esters (HR: 0.70, 95%CI: 0.54-0.91, P=0.002) showing an inverse association with HF risk in the decreased, but not preserved kidney function. A large proportion of estimates in a group with decreased kidney function did not reach statistical significance likely due to low statistical power.

Figure 20. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' baseline systolic blood pressure.

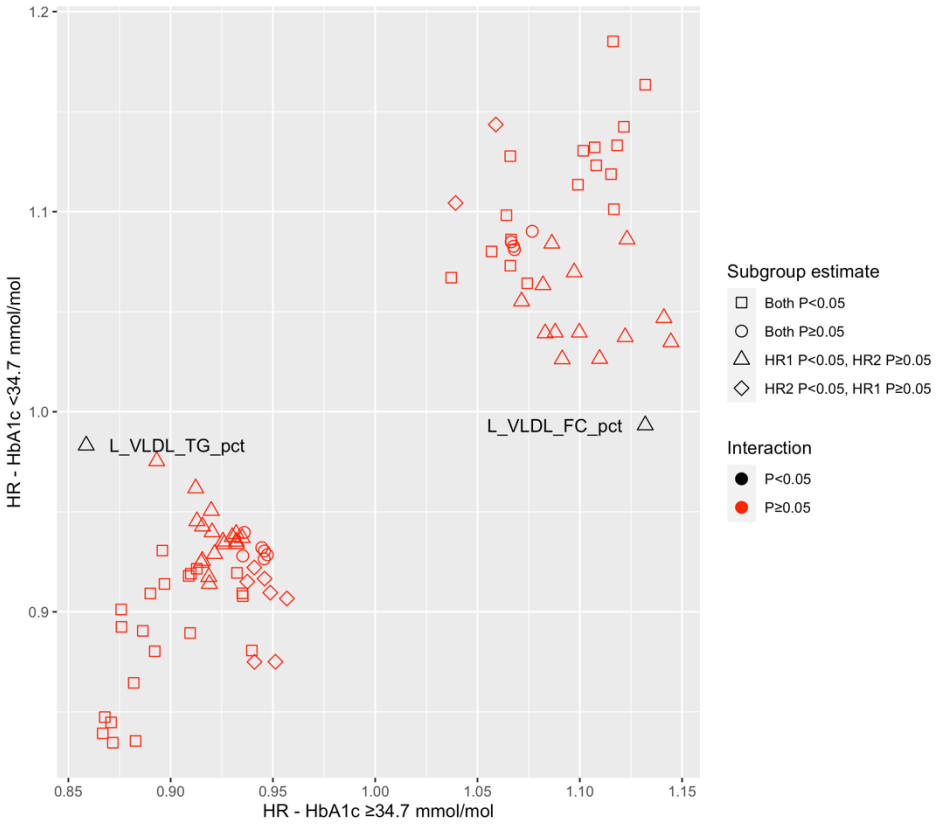


Stratified estimates (HR) are derived from a multivariable-adjusted Cox proportional hazards model excluding age.
P for interaction was calculated for multiplicative interaction term between a given metabolite and age subgroup.
Labelled are top 10% metabolite measures with lowest P for interaction.

Analyses stratified for baseline SBP are presented in Figure **20**. There were 56,319 participants who had normal SBP (<140 mmHg) and 37,603 participants who had elevated SBP at baseline

(≥140 mmHg). Among them respectively, 679 (1.2%) and 959 (2.6%) were diagnosed with HF during follow-up. All 90 metabolite measures showed consistent associations with HR risk, as significant interaction term was observed in none of them (P>0.05).

Figure 21. Scatter plot comparing hazard ratios for heart failure associated with 90 significant metabolic measures by participants' baseline glycated hemoglobin level.



Stratified estimates (HR) are derived from a multivariable-adjusted Cox proportional hazards model excluding age.
P for interaction was calculated for multiplicative interaction term between a given metabolite and age subgroup.
Labelled are top 10% metabolite measures with lowest P for interaction.

Associations for significant metabolite measures-HF stratified for baseline glycated hemoglobin levels are presented in **Figure 21**. 638 out of 46,671 (1.4%) and 1000 out of 47,251 (2.1%) participants with glycated hemoglobin respectively <34.7 and ≥34.7 mmol/mol developed incident HF during follow-up. Among 90 metabolite measure two differed depending on baseline long-term glycemia. Relative content of free cholesterol in large VLDL particles was associated with a higher, while relative content of triglycerides in the same size subclass with a lower HF risk among participants with higher, but not those with lower baseline long-term glycemia.

Sensitivity analyses for 90 heart failure related metabolite measures

Figure 22 presents sensitivity analysis for 90 significant metabolite-measure HF associations after exclusion of first three years of follow-up, in order to minimize influence of early incident HF in participants who had a high baseline risk of HF. Almost all metabolite measures with exception for relative content of cholesteryl esters in small LDL particles, and relative content of free cholesterol in small HDL particles, maintained stable and significant association with incident HF.

Reproduction of main findings using age as underlying time scale in Cox PH model is presented in **Figure 23**. All 90 metabolite measures showed exactly same results with main analysis.

Figure 24 summarizes sensitivity analyses including further adjustment for self-reported physical activity level at baseline, which had substantial missing rate and there was not imputed and included in main covariate adjustment set. Associations for ten metabolite measures, glycoprotein acetyls, relative content of cholesterol in very large VLDL, relative content free cholesterol and triglycerides in large VLDL, relative content of total cholesterol, cholesteryl esters, free cholesterol, phospholipids and triglycerides in medium VLDL, as well as relative content of cholesteryl esters in small VLDL were no longer significant. Moreover, small-to-moderate attenuation of associations could be observed for a number of crude and relative measures related to various VLDL, LDL and HDL particles.

Sensitivity analyses including further adjustment for baseline healthy diet score are presented in **Figure 25**. All 90 metabolite measures showed consistent results with main analysis. Small attenuation of association was observed for various measures related HDL particles and PUFA.

Figure 26 reproduces associations for 90 metabolite measures following inclusion of non-linear terms for BMI, fasting time, eGFR and alcohol intake (all as restricted cubic splines with three knots set 10^{th} , 50^{th} and 90^{th} percentile of distribution). All association with exception LDL cholesterol remained significant with largely comparable magnitude of association. Small attenuation of association was observed for small HDL particles and their lipid content.

Figure 22. Sensitivity analysis for metabolite measures significantly associated with incident heart failure after exclusion of first three years of follow-up.



Figure 23. Sensitivity analysis for metabolite measures significantly associated with incident heart failure using age as underlying time scale for regression models.

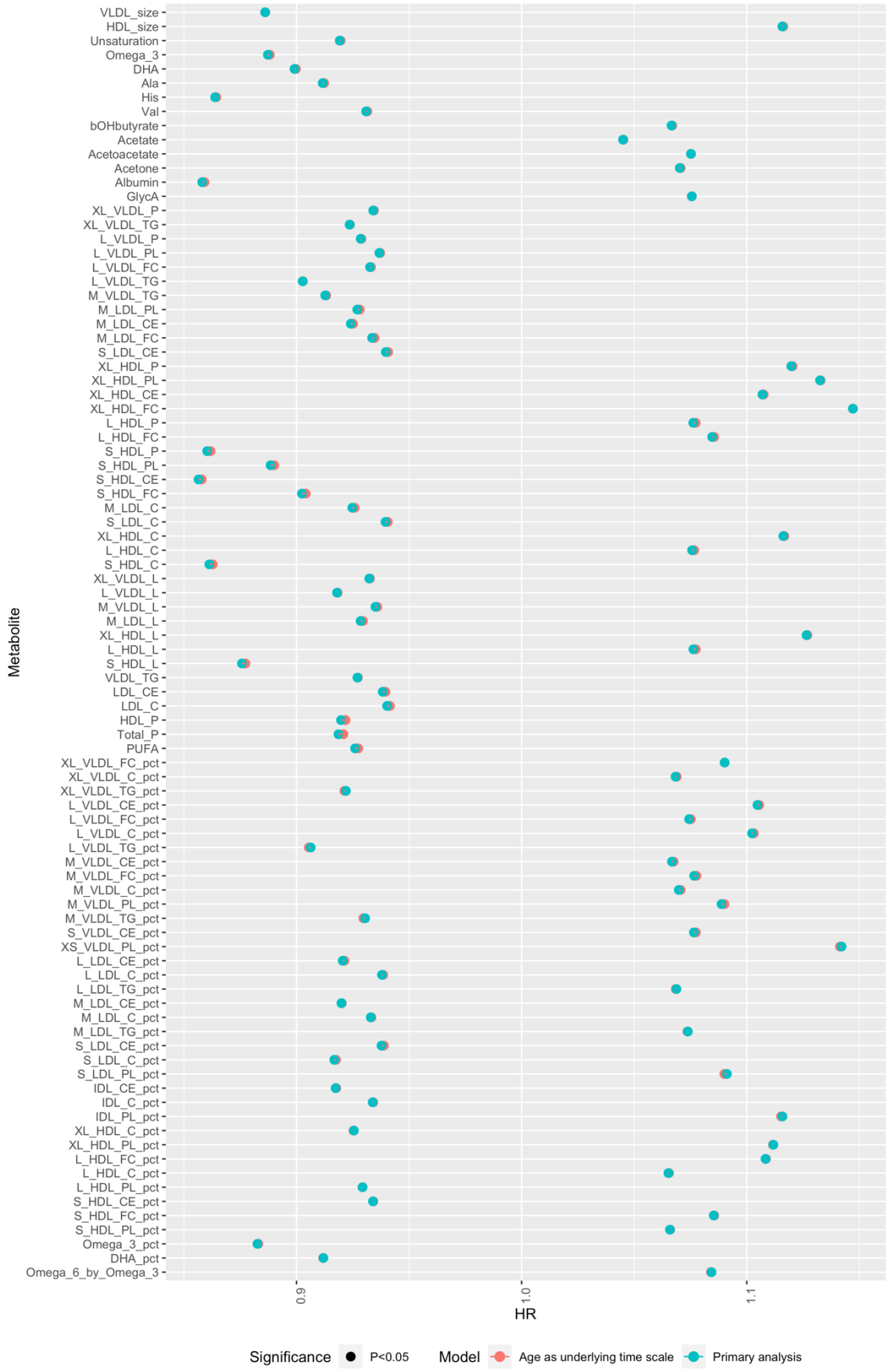


Figure 24. Sensitivity analysis for metabolite measures significantly associated with incident heart failure after further adjustment for baseline self-reported physical activity level.

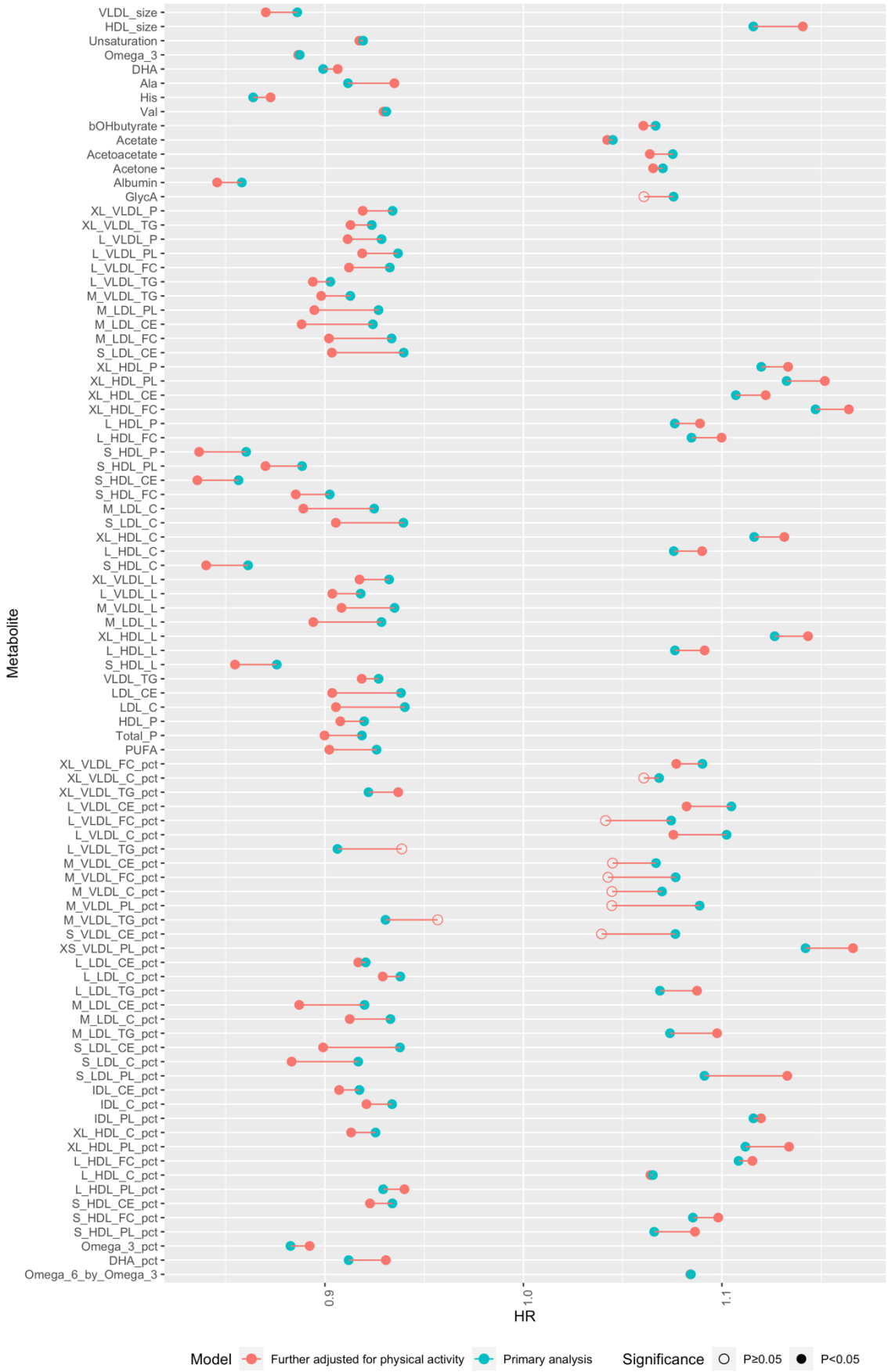


Figure 25. Sensitivity analysis for metabolite measures significantly associated with incident heart failure after further adjustment for baseline healthy diet score.

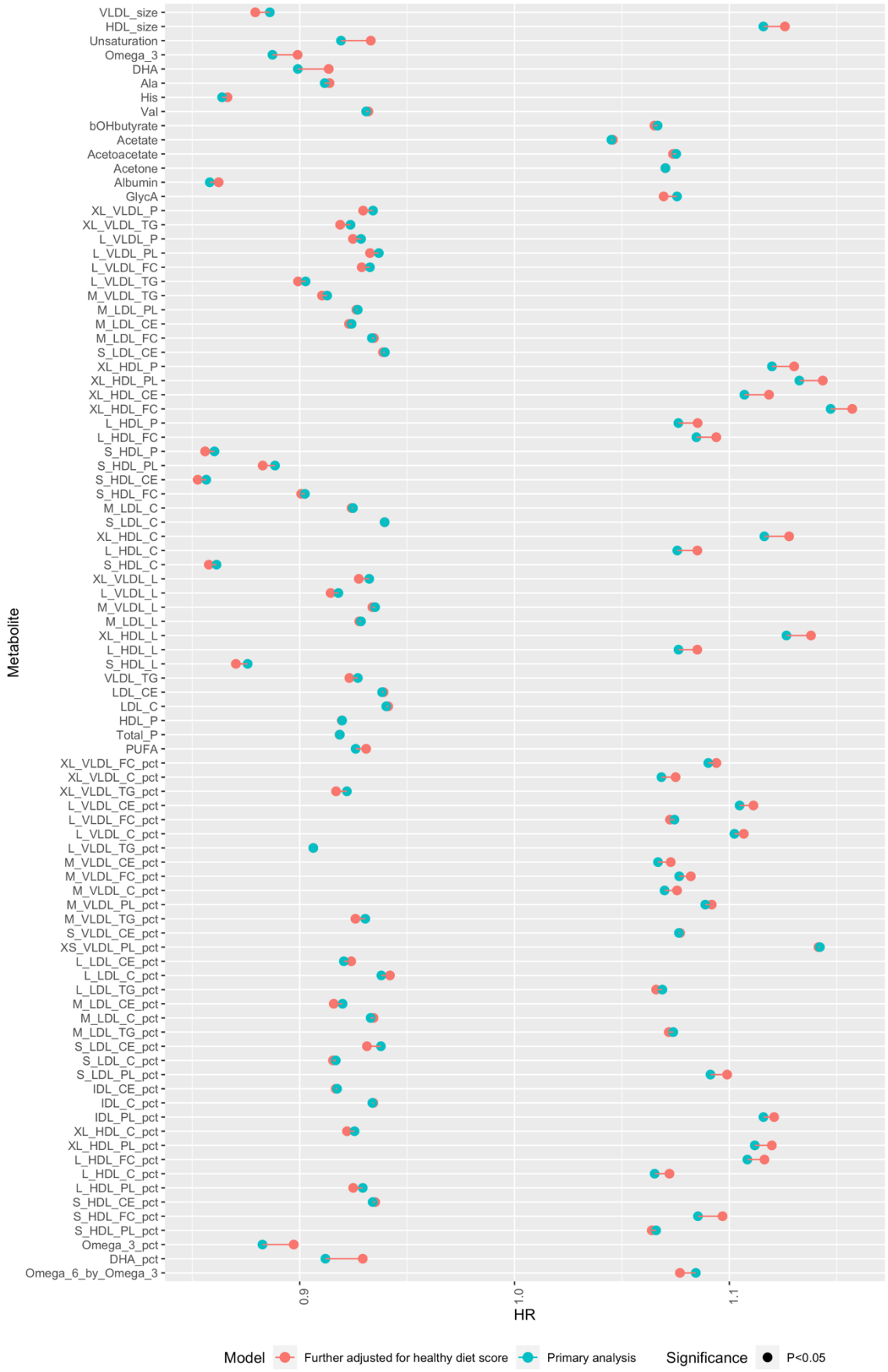


Figure 26. Sensitivity analysis for metabolite measures significantly associated with incident heart failure after inclusion of non-linear covariate terms to regression models.



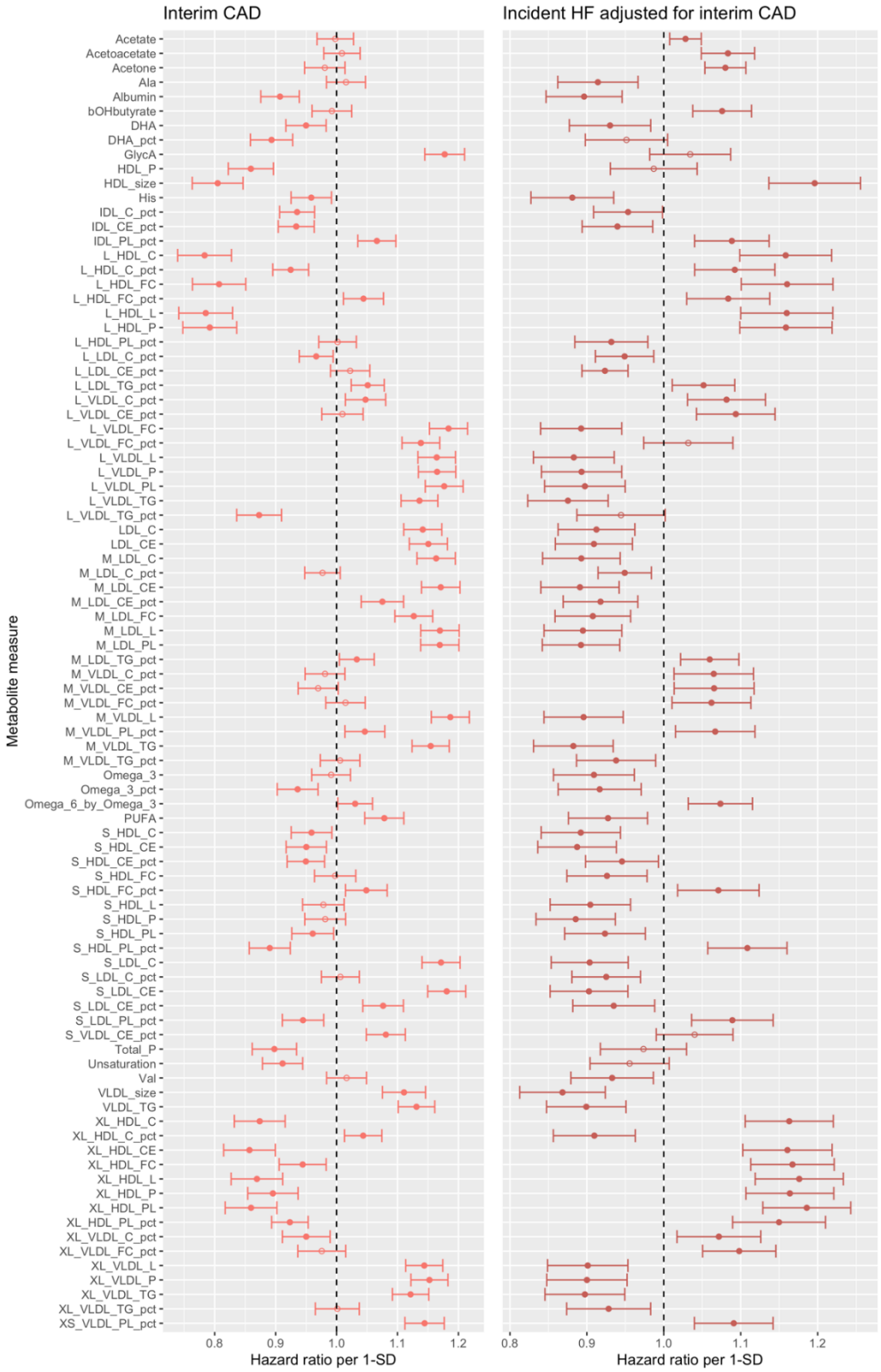
Mediation analyses for indirect effects of interim events on metabolite measure-HF associations

Figures 27-30 present associations between 90 metabolite measures linked with HF risk and occurrence of interim events (CAD, AF, VHD, and DM) that were diagnosed before incident HF, as well as effect of adjustment for particular event on association between metabolite measure and HF risk.

None of measures belonging to ketone bodies were associated with interim CAD (**Figure 27**). Among amino acids only histidine (HR: 0.96, 95%CI: 0.93-0.99, P=0.012) showed significant association with lower CAD risk. In FAs class, an inverse association with CAD could be observed for crude (HR: 0.95, 95%CI: 0.92-0.98, $P=2.3 \times 10^{-3}$) and relative concentration of DHA (HR: 0.89, 95%CI: 0.86-0.92 $P=1.8 \times 10^{-10}$), degree of unsaturation (HR: 0.91, 95%CI: 0.88-0.94, $P=3.7 \times 10^{-8}$) as well as relative concentration of omega-3 FAs (HR: 0.94, 95%CI: 0.91-0.97, $P=1.1 \times 10^{-4}$). Inflammation-related glycoprotein acetyl showed strong positive association with CAD risk (HR: 1.18, 95%CI: 1.14-1.22, P < 0.001). Higher concentration of albumin was associated with a lower CAD risk (HR: 0.91, 95%CI: 0.88-0.94, $P=1.5 \times 10^{-9}$). The direction of associations for lipoproteins depended on class, with trend towards higher risk for measures related to VLDL and LDL particles and trend towards a lower risk for HDL-related measures. Magnitude of association with CAD for crude metabolite measures was consistent across different VLDL and LDL size subclasses (HR: 1.12-1.17). However, a larger average diameter of VLDL showed a trend towards higher CAD risk (HR: 1.11, 95%CI: 1.07-1.15, $P=6.0 \times 10^{-9}$). For HDL related measures, the association appeared to be weaker for small HDL (HR: 0.95-0.98) than for large and very large HDL (HR: 0.79-0.94). This was consistent with lower CAD risk with increasing average diameter of HDL (HR: 0.80, 95%CI: 0.77-0.84, $P=1.9 \times 10^{-24}$).

Following adjustment of incident HF models for interim CAD majority of metabolite measures showed relative consistency with main analyses results. Total lipoprotein count, HDL particle count, degree of unsaturation and glycoprotein acetyls were no longer associated with HF risk. Furthermore, previously significant associations were attenuated for relative content of free cholesterol and triglycerides in large VLDL, relative content of cholesteryl esters in small VLDL, as well as relative content of DHA.

Figure 27. Multivariable-adjusted hazard ratios (95%CI) for interim coronary artery disease (left pane) and incident heart failure after adjustment for this interim event (right pane).

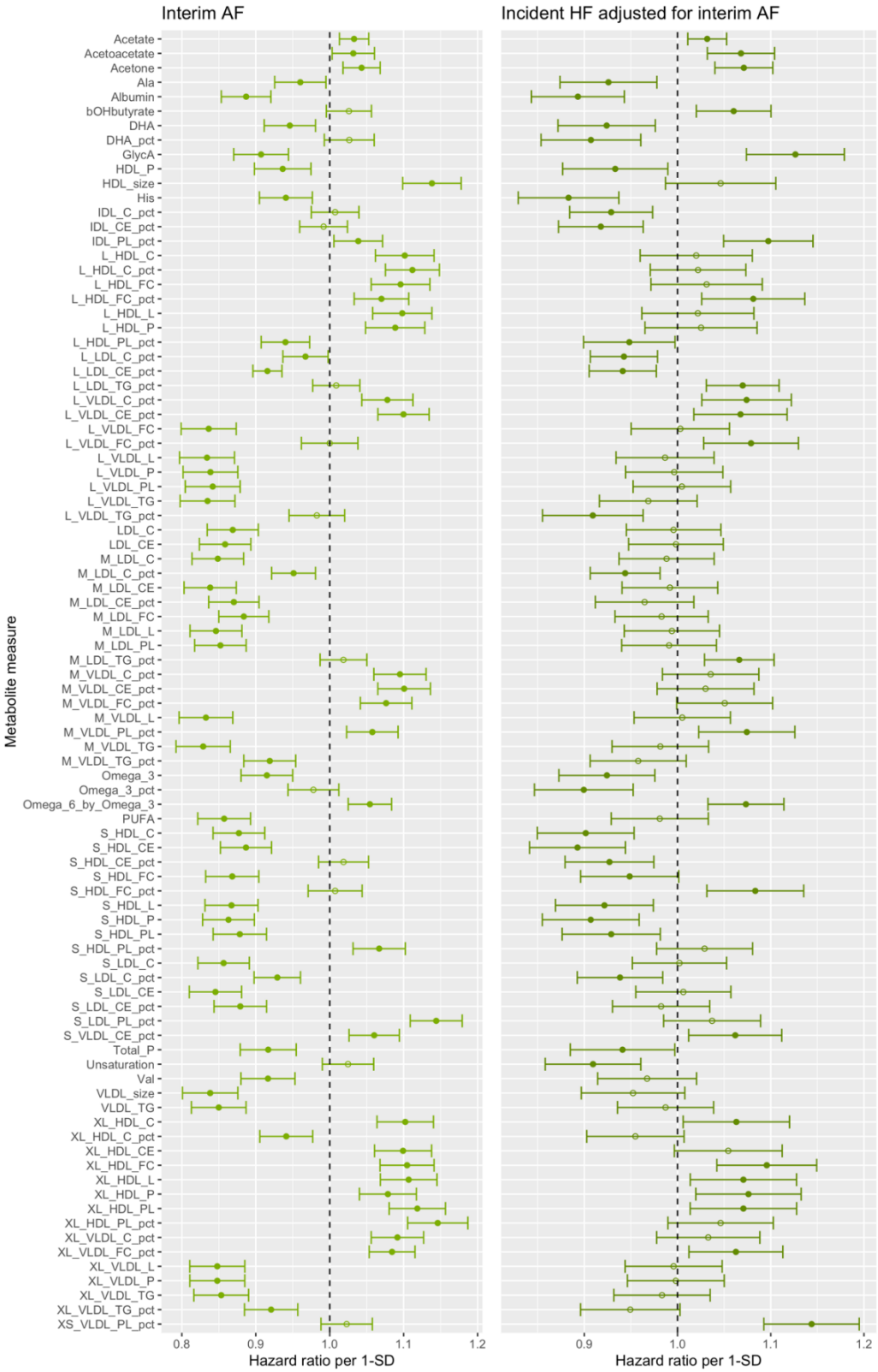


Estimates are derived from a Cox proportional hazards model including one metabolite measure at a time and further adjusted for age, sex, body mass index, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication. Additionally, models in the right pane included adjustment for interim CAD.

All three considered amino acids, alanine (HR: 0.96, 95%CI: 0.93-0.99, P=0.022), histidine (HR: 0.94, 95%CI: 0.91-0.97, P=8.2 x 10^{-4}) and leucine (HR: 0.92, 95%CI: 0.88-0.95, P=2.9 x 10^{-6}) showed an inverse association with interim AF. In ketone bodies class, acetone (HR:1.04, 95%CI: 1.02-1.07, P=0.001), acetate (HR: 1.03, 95%CI: 1.01-1.05, P=0.001) and acetoacetate (HR: 1.03, 95%CI: 1.00-1.06, P=0.031) showed a positive association with AF risk. Higher concentrations of both albumin (HR: 0.89, 95%CI: 0.86-0.92, P=2.5 x 10^{-12}) and glycoprotein acetyls (HR: 0.91, 95%CI: 0.87-0.94, P=2.5 x 10^{-7}) showed an inverse association with interim AF. In FAs class significant linked with AF risk was observed for concentration of DHA (HR: 0.91, 95%CI: 0.88-0.95, P=6.7 x 10^{-7}), total omega-3 FAs (HR: 0.95, 95%CI: 0.91-0.98, P=1.7 x 10^{-3}) and total concentration of PUFA (HR: 0.86, 95%CI: 0.83-0.89, P=2.6 x 10^{-17}). Triglyceride content in VLDL particles, as well as average diameter of VLDL (HR: 0.84, 95%CI: 0.80-0.87, P=2.8 x 10^{-20}) showed an inverse association with risk of interim AF. Consistently various metabolite measures related to considered VLDL size subclasses were associated with lower risk of interim AF. Similarly, LDL cholesterol (HR: 0.87, 95%CI: 0.84-0.90, P=1.46 x 10^{-15}) and measures related to different LDL size subclasses were inversely associated with interim AF. The magnitude of association between HDL particles and interim AF turned out to depend on size. Measures related to small HDL particles (HR: 0.86-0.88) were associated with lower risk of interim AF, while measures related to very large HDL with a higher risk of interim AF (HR: 1.09-1.10). These findings were confirmed by positive trend for AF risk with larger HDL diameter (HR: 1.14, 95%CI: 1.09-1.18, P=1.3 x 10^{-10}).

After adjustment for interim AF, only 48 metabolite measures out of 90 showed consistent associations with HF compared to main analysis. Concertation of PUFA and valine was no longer associated with incident HF. Inclusion of interim AF, attenuated associations between total LDL cholesterol and other crude metabolite measures across LDL subclasses. Moreover, an inverse association with HF for triglyceride content in VLDL was no longer significant following that adjustment. Consistently, associations for all crude metabolite measures related to VLDL size, as well as the average diameter for VLDL were attenuated. Furthermore, metabolite measures related to large HDL particle, as well as average diameter for HDL particles were not associated with HF after adjustment for interim AF. Moreover, non-significant relationship with HF was observed for 13 relative lipid contents across VLDL, LDL and HDL particle size subclasses.

Figure 28. Multivariable-adjusted hazard ratios (95%CI) for interim atrial fibrillation (left pane) and incident heart failure after adjustment for this interim event (right pane).

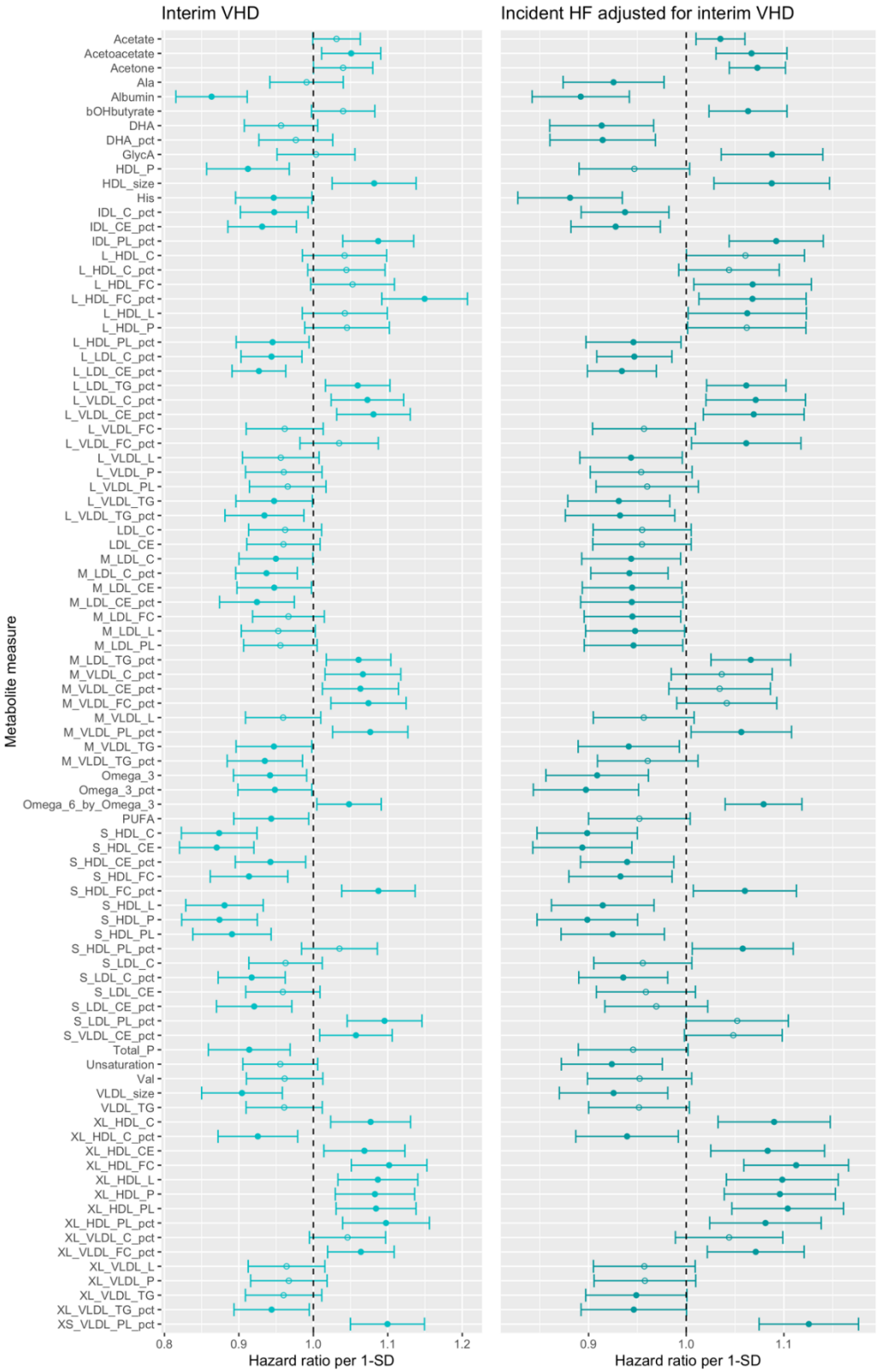


Estimates are derived from a Cox proportional hazards model including one metabolite measure at a time and further adjusted for age, sex, body mass index, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication. Additionally, models in the right pane included adjustment for interim AF.

Only one of amino acids histidine was associated with an interim VHD (HR: 0.95, 95%CI: 0.90-1.00, P=0.037). Among ketone bodies, a positive association with VHD risk was observed for acetoacetate (HR: 1.05, 95%CI: 1.01-1.09, P=0.014), but did not reach significance for acetone (HR: 1.04, 95%CI: 1.00-1.08, P=0.054), acetate (HR: 1.03, 95%CI: 1.00-1.06, P=0.060) or 3-OHB (HR: 1.04, 95%CI: 1.00-1.09, P=0.070). Glycoprotein acetyls were not associated with interim VHD risk (HR: 1.00, 95%CI: 0.95-1.06, P=0.890). In FAs class, inverse associations with VHD risk were observed for the concentration of omega-3 FAs (HR: 0.94, 95%CI: 0.90-0.99, P=0.017) and total PUFA (HR: 0.94, 95%CI: 0.90-0.99, P=0.024). Moreover, a lower risk of VHD was linked with a higher concentration of albumin (HR: 0.86, 95%CI: 0.82-0.91, P=1.8 x 10^{-9}). Total cholesterol (HR: 0.96, 95%CI: 0.92-1.01, P=0.120) and cholesteryl esters (HR: 0.96, 95%CI: 0.91-1.01, P=0.100) were not associated with interim VHD. None of crude metabolite measures related to LDL size subclasses with exception for cholesteryl esters in medium LDL particles (HR: 0.95, 95%CI: 0.90-1.00, P=0.035). Triglyceride content of VLDL particles was not linked with interim VHD (HR: 0.96, 95%CI: 0.91-1.01, P=0.130). In contrast, average diameter for VLDL was associated with a lower risk of interim VHD (HR: 0.90, 95%CI: 0.86-0.95, P=2.7 x 10^{-4}). Total concentration of HDL particles was associated with lower (HR: 0.91, 95%CI: 0.86-0.96, P=0.001), while average diameter of HDL particles with a higher VHD risk (HR: 1.08, 95%CI: 1.02-1.14, P=0.006).

Adjustment of incident HF models for interim VHD did not alter associations between non-lipoprotein-related metabolite measures. However, associations for metabolite measures belonging to lipoprotein classes were considerably attenuated. Associations with incident HF for major lipid fraction in VLDL and LDL, triglycerides, and cholesterol were no longer significant after adjustment for interim VHD. Furthermore, inclusion of VHD in the model nullified inverse associations for PUFA and valine. Among VLDL size subclasses, additional adjustment attenuated relationship for very large and large VLDL particle count, as well as for total lipids in very large VLDL, phospholipids and free cholesterol content in large VLDL. Inverse association for total HDL particle was no longer observed after adjustment for interim VH, while trend towards lower HF risk for measures related to small HDL particles and trend towards higher risk for large and very HDL particles. Moreover, for cholesterol content in small and large HDL, as well as for count of HDL particles. Lastly, association with incident HF were nullified fore between nine relative lipid fractions belonging to VLDL, LDL and HDL size subclasses.

Figure 29. Multivariable-adjusted hazard ratios (95%CI) for interim valvular heart disease (left pane) and incident heart failure after adjustment for this interim event (right pane).



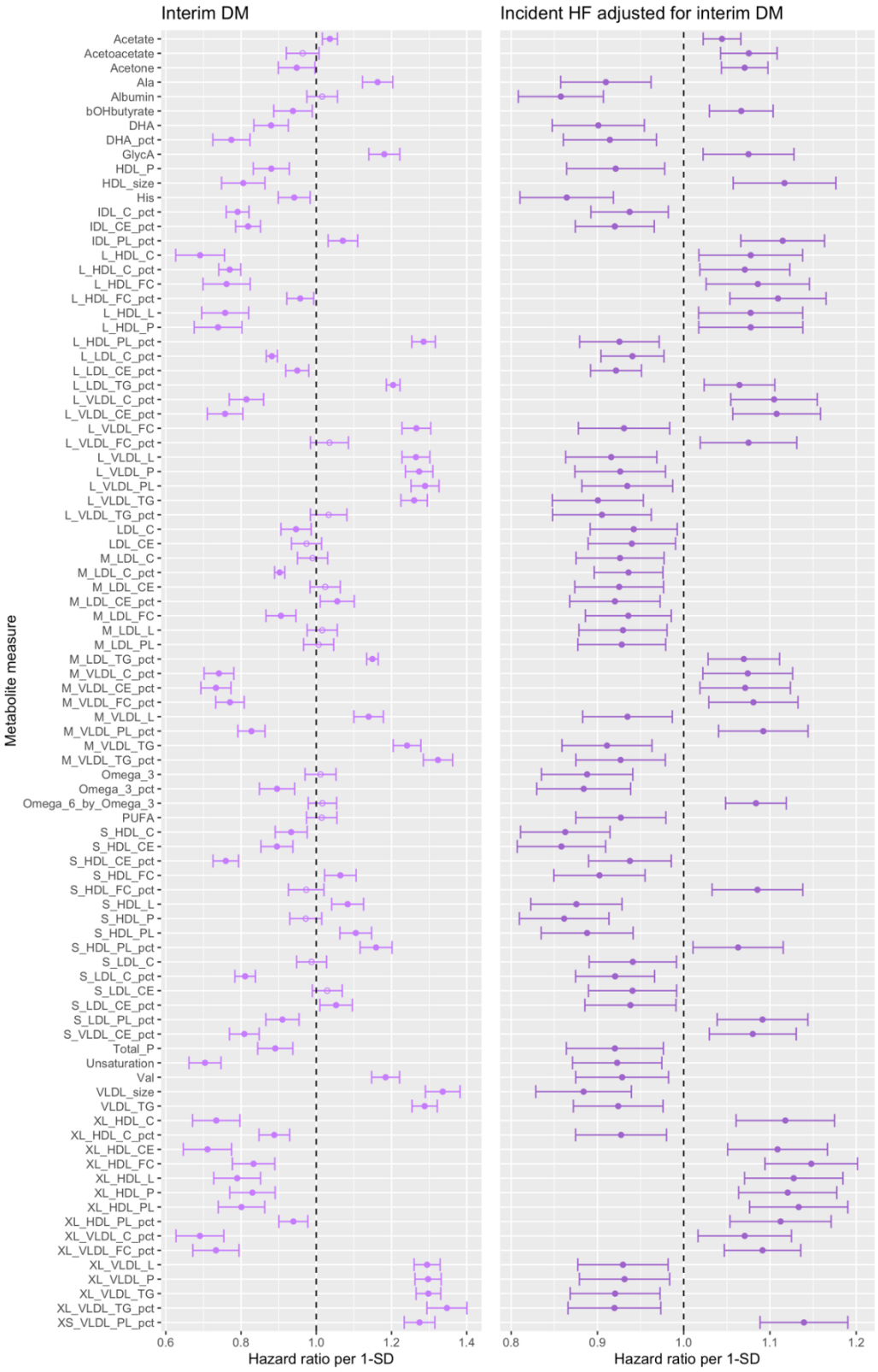
Estimates are derived from a Cox proportional hazards model including one metabolite measure at a time and further adjusted for age, sex, body mass index, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication. Additionally, models in the right pane included adjustment for interim VHD.

Higher concertation of alanine (HR: 1.16, 95%CI: 1.12-1.21, P=1.7 x 10^{-13}) and valine (HR: 1.18, 95%CI: 1.14-1.23, P=3.4 x 10-19) were associated with higher risk of DM, while histidine (HR: 0.94, 95%CI: 0.90-0.98, P=0.005) with a lower risk of DM. In ketone bodies class, acetate was positively associated with interim DM (HR: 1.04, 95%CI: 1.02-1.06, P=5.0 x 10^{-4}), while 3-OHB inversely associated with interim DM (HR: 0.94, 95%CI: 0.89-0.99, P=0.014). Inflammation-relation glycoprotein acetyls were associated with a higher risk of DM (HR: 1.18, 95%CI: 1.13-1.23, P=2.5 x 10^{-15}). In contrast, albumin was not associated with interim DM (HR: 1.02, 95%CI: 0.98-1.06, P=0.450). Among measures related to FAs, strong inverse associations with interim DM were observed for degree of unsaturation (HR: 0.70, 95%CI: 0.67-0.73, P=1.2 x 10^{-58}) and DHA (HR: 0.88, 95%CI: 0.84-0.92, P=4.2 x 10^{-8}). VLDL-related measures including total triglyceride content (HR: 1.29, 95%CI: 1.25-1.33, P=1.4 x 10^{-49}) and average diameter for VLDL (HR: 1.34, 95%CI: 1.28-1.40, P=6.1 x 10^{-35}) were associated with a higher risk of interim DM. The magnitude of associations across VLDL size subclasses was comparable (HR: 1.24-1.30). Total cholesterol in LDL particles was associated with lower risk of interim DM (HR: 0.95, 95%CI: 0.91-0.98, P=0.007), however, cholesterol measures across LDL size subclasses did not show significant associations with DM. Total count (HR: 0.88, 95%CI: 0.84-0.92, P=1.8 x 10-7), as well as average diameter for HDL particles (HR: 0.81, 95%CI: 0.76-0.85, P=2.0 x 10-13) were linked to lower risk of interim DM. Total lipid content (HR: 1.08, 95%CI: 1.04-1.13, P=2.1 x 10^{-4}), phospholipids (HR: 1.11, 95%CI: 1.06-1.15, P=3.4 x 10^{-6}) and free cholesterol (HR: 1.06, 95%CI: 1.02-1.11, P=0.003) in small HDL were positively associated with interim DM, In contrast, cholesteryl esters (HR: 0.90, 95%CI: 0.86-0.93, P=3.6 x 10^{-7}) and total cholesterol (HR: 0.93, 95%CI: 0.89-0.97, P=0.001) in small HDL were inversely associated with an interim DM. Measures related to very large and large HDL were consistently linked to a lower risk of DM (HR: 0.69-0.83).

Further adjustment of models for association between significant metabolites measure and incident HF did not influence main analysis findings as magnitude and direction of estimates was not changed.

Figure 28 summarizes proportion of association between metabolite measures and incident HF attributable to occurrence of interim CAD, AF, VHD or DM prior to the diagnosis of HF. If the attributable proportion is larger than 100%, this means that the estimated indirect effect mediated by an interim event is larger than the direct effect of metabolite measure on HF.

Figure 30. Multivariable-adjusted hazard ratios (95%CI) for interim diabetes mellitus (left pane) and incident heart failure after adjustment for this interim event (right pane).



Estimates are derived from a Cox proportional hazards model including one metabolite measure at a time and further adjusted for age, sex, body mass index, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication. Additionally, models in the right pane included adjustment for interim DM.

Negative attributable proportion can be observed when the direct and indirect effects have an opposite sign of association with incident HF. It suggests that the mediator suppresses, rather than explains, the direct effect of the independent variable on the dependent variable. In other words, the effect of the independent variable on the dependent variable would be stronger if the mediator was not present. For several metabolite measures there was a pattern for opposite sign of attributable proportion for mediation by CAD and AF, as well as to a lesser extent, VHD. This pattern was observed in case of metabolite measures that were associated with a higher risk of interim CAD, but simultaneously associated with a lower risk of AF and VHD. Then assuming causality of both association between metabolite measure and interim event, the true effect of metabolite over HF risk is null and if uncontrolled for interim events it is biased towards a sign of association for a more prevalent interim effect.

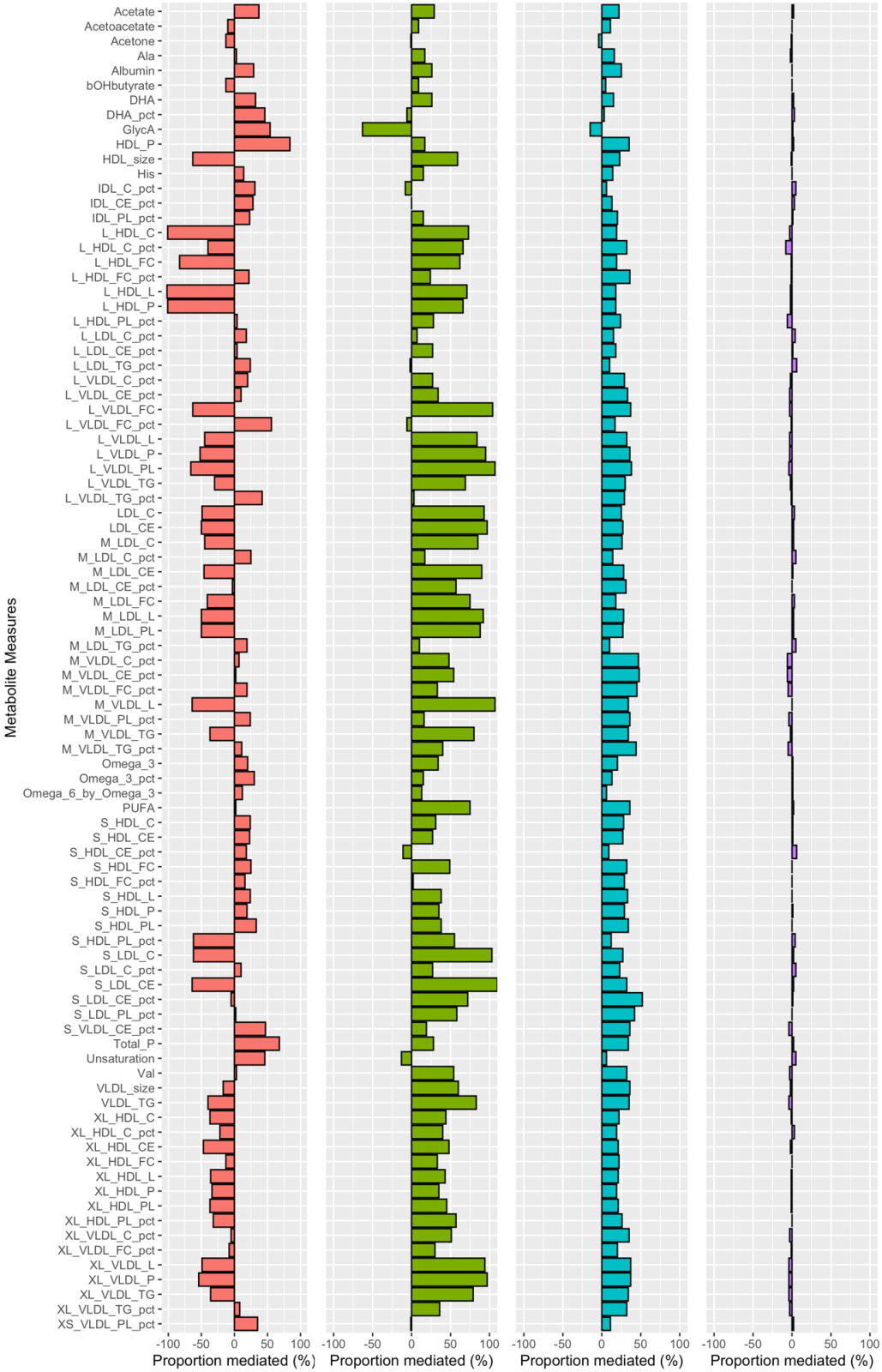
Interim CAD explained 84% of inverse association between total HDL particle count and incident. Moreover, it considerably mediated association with glycoprotein acetyls (54%) and total lipoprotein count (68%). Interim CAD also explained a lesser proportion of relationship between degree of unsaturation (46%), acetate (37%), DHA (32%), albumin (29%), omega-3 FAs (20%), and measures related to small HDL particles (19-33%).

Occurrence of AF prior to HF diagnosis explained respectively 93% and 83% of inverse associations of LDL cholesterol and VLDL triglycerides with incident HF. Moreover, interim AF explained 84-100% of crude measures related to VLDL size subclasses as well as 75-100% of crude measures related to LDL size subclasses. It also mediated a considerable part of the relationship with PUFA (74%) and valine (54%). Attributable proportion of effect mediated by interim VHD ranged from 10 to 40% for the majority of metabolite measures.

Interestingly, while interim CAD, AF and VHD individually explained substantial proportion of many metabolites measure-HF association (over 75% of total effect), effect mediated by interim DM for almost all metabolite measures did not exceed 5%.

To summarize, 23 metabolite measures showed independent association with incident HF that was not attenuated after adjustment nor indirectly explained by occurrence of interim predisposing events. These included amino acids (alanine, histidine), ketone bodies (acetate, acetoacetate, acetone, 3-hydroxyybutyrate), fluid balance (albumin), FAs (DHA, omega-3 FAs, relative content of omega-3 FAs, ratio of omega-6 to omega-3 FAs), very large HDL (free cholesterol content), small LDL (particle count, total lipid content, total cholesterol, cholesteryl

Figure 31. Proportion of association between metabolite measures and incident heart failure attributable to occurrence of interim events.



The far-left pane represents the proportion mediated by interim CAD, the center-left pane represents mediation by interim AF, the center-right pane represents the proportion mediated by interim VHD, and the far-right pane represents mediation by interim DM.

esters, phospholipids) and lipoprotein ratios (relative cholesterol content in large VLDL, relative phospholipids content in very small VLDL, relative cholesteryl esters content in large LDL, relative content of cholesteryl esters and phospholipids in IDL particles).

Assessment of relative metabolite measure validity

Relative metabolite measures (percentage of total lipid fraction, ratios) represented a substantial number of significant metabolite measures for incident HF identified in main, subgroup, and sensitivity analyses. However, many relative measures available in the Nightingale platform showed an opposite direction with HF risk compared to the crude measure. This might be counterintuitive especially in the content of lipoproteins as usually nominator (lipid fraction) and denominator (sum of all lipid fraction per class) show substantial collinearity as they change proportionally with an increasing number of lipoprotein particles. In such scenario, the ratio represents a theoretical situation when of the lipid fractions is changing its concentration while keeping the sum constant. Nevertheless, such situation is largely theoretical as variability of relative concentration is substantially smaller than variability of crude lipid fraction concentration. Moreover, calculation of a ratio between two highly correlated variables instead of creation of a new quality indicator that is independent from sum of lipid fractions might present significant spurious correlation with it. Therefore, it is crucial to determine if relative metabolite measures are more informative than crude measures constituting them alone or whether association between relative measure with incident HF can be explained by presence of spurious correlation.

Table 2 summarizes HR for relative metabolite measures which previously were significantly associated with incident HF, together with comparison of goodness of fit statistic (Harrell's C) for the relative measure together with the measure for analogous model fitted with nominator and denominator. Additionally, correlation between nominator and denominator were compared with correlation between respective relative measure and its denominator.

With exception for relative content of phospholipids in very small VLDL and phospholipids and free cholesterol in large HDL, Harrell's C was equal or higher for nominator and denominator of a relative measure than for the actual ratio. However, even for these two relative measures the goodness of fit was comparable and therefore use of ratio had no clear superiority over variables constructing those ratios. Most relative measures showed substantial collinearity (r_{ND}: 0.90-0.99), and newly create relative measures despite hypothetical independence from sum of lipid fractions (denominator) showed significant correlation with it (r_{RD}). This indicates

that relative measures are still substantially dependent on the denominator meaning that their variability is explained by changes in denominator. That is why in most of the cases, the direction of correlation between relative measure and denominator would denote concordance of association between crude and relative measure. For example, total lipid content in very large VLDL associated with lower a HF risk. Relative content of free cholesterol was negatively associated with total lipid content (r_{RD}=-0.549), with positive association between it and HF (HR 1.09).

A conclusion from this analysis is that relative measures are not more informative than crude measures contributing to them, and significant association observed for them might be in most cases explained by presence of spurious correlation. Therefore, in subsequent stages of this study, a focus will be reduced to 168 non-ratio metabolite measures.

Table 2. Goodness of fit statistics for ratio measures relevant for heart failure risk, their nominators, and denominators.

Metabolite measure	HR	Harrell's C-statistic			r_{ND}	r_{RD}
		Ratio	Nominator	Denominator		
Free Cholesterol in Very Large VLDL (%)	1.09	0.765	0.765	0.765	0.992	-0.549
Cholesterol in Very Large VLDL (%)	1.07	0.765	0.765	0.765	0.971	-0.733
Triglycerides in Very Large VLDL (%)	0.92	0.765	0.765	0.765	0.993	0.581
Cholesteryl Esters in Large VLDL (%)	1.10	0.766	0.765	0.766	0.942	-0.527
Free Cholesterol in Large VLDL (%)	1.07	0.765	0.765	0.766	0.993	0.171
Cholesterol in Large VLDL (%)	1.10	0.766	0.765	0.766	0.978	-0.368
Triglycerides in Large VLDL (%)	0.91	0.766	0.766	0.766	0.989	-0.076
Cholesteryl Esters in Medium VLDL (%)	1.07	0.765	0.764	0.765	0.689	-0.367
Free Cholesterol in Medium VLDL (%)	1.08	0.765	0.765	0.765	0.948	-0.125
Cholesterol in Medium VLDL (%)	1.07	0.765	0.764	0.765	0.828	-0.310
Phospholipids in Medium VLDL (%)	1.09	0.765	0.765	0.765	0.982	0.014
Triglycerides in Medium VLDL (%)	0.93	0.765	0.766	0.765	0.925	0.248
Cholesteryl Esters in Small VLDL (%)	1.08	0.765	0.765	0.765	0.948	0.024
Phospholipids in Very Small VLDL (%)	1.14	0.767	0.765	0.765	0.994	0.103

Cholesteryl Esters in Large LDL (%)	0.92	0.765	0.765	0.764	0.983	0.344
Cholesterol in Large LDL (%)	0.94	0.765	0.765	0.764	0.988	0.376
Triglycerides in Large LDL (%)	1.07	0.765	0.764	0.764	0.983	-0.291
Cholesteryl Esters in Medium LDL (%)	0.92	0.766	0.765	0.765	0.997	0.465
Cholesterol in Medium LDL (%)	0.93	0.765	0.765	0.765	0.997	0.452
Triglycerides in Medium LDL (%)	1.07	0.765	0.765	0.765	0.557	-0.324
Cholesteryl Esters in Small LDL (%)	0.94	0.765	0.765	0.765	0.992	0.481
Cholesterol in Small LDL (%)	0.92	0.765	0.765	0.765	0.998	0.388
Phospholipids in Small LDL (%)	1.09	0.765	0.765	0.765	0.665	-0.241
Cholesteryl Esters in IDL (%)	0.92	0.765	0.765	0.765	0.986	0.462
Cholesterol in IDL (%)	0.93	0.765	0.765	0.765	0.995	0.484
Phospholipids in IDL (%)	1.12	0.765	0.765	0.765	0.969	-0.44
Cholesterol in Very Large HDL (%)	0.93	0.766	0.766	0.766	0.989	-0.704
Phospholipids in Very Large HDL (%)	1.11	0.766	0.766	0.766	0.996	0.900
Free Cholesterol in Large HDL (%)	1.11	0.766	0.765	0.765	0.995	0.653
Cholesterol in Large HDL (%)	1.07	0.765	0.765	0.765	0.991	0.651
Phospholipids in Large HDL (%)	0.93	0.765	0.765	0.765	0.997	-0.455
Cholesteryl Esters in Small HDL (%)	0.93	0.765	0.767	0.767	0.864	-0.144
Free Cholesterol in Small HDL (%)	1.09	0.765	0.766	0.767	0.91	-0.120
Phospholipids in Small HDL (%)	1.07	0.765	0.766	0.767	0.985	0.121
Omega-3 Fatty Acids (%TFA)	0.88	0.765	0.765	0.764	0.559	0.144
Docosahexaenoic Acid (%RFA)	0.91	0.765	0.765	0.764	0.313	-0.231
Omega-6 to Omega-3 Fatty Acids ratio	1.08	0.765	0.764	0.765	0.476	-0.938

HR is a multivariable association between a given ratio and incident HF (as in Figure 4-10). Harrel's C-statistic was calculated for Cox proportional hazards model including a ratio, nominator and denominator and adjusted for age, sex, body mass index, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication.

R_{ND} – Pearson's correlation between nominator and denominator, r_{RD} – Pearson's correlation between ratio and denominator.

Comparison of particle count and lipid content predictive power

Among metabolite measures that were interpedently associated with incident HF there were two lipoprotein classes small and very large HDL particles. While for very large HDL only free cholesterol remained significantly associated, six measures for small HDL particles were all significantly associated with HF risk. High intercorrelation between them makes obstructs their simultaneous use in joint metabolite measure score or multivariable analysis. Moreover, it cannot be ruled out that selection of only free cholesterol in HDL but not other measures from this class could be a false negative. Therefore, bootstrap resampling was used over all crude measures belonging to very large and small HDL to re-run multivariable adjusted Cox models and determine if any of lipid fractions. Partial likelihood ratio (95% CI) was used to compare the importance of measures. Results of this analysis are presented in **Table 3**. Despite that total lipid content was ranked as the most important very large HDL measure for HF risk and particle count as the most important small HDL measure for HF risk, bootstrapped CI suggested that superiority of none of the measures can be suggested. That is why in subsequent chapter of this study particle count of those two classes will be used as proxy for these two lipoprotein groups.

Table 3. Comparison of relative importance of metabolite measures in small and very large HDL class.

Metabolite measure	Partial Likelihood Ratio	Bootstrapped 95% CI
Very large HLDL		
Concentration of Very Large HDL	164.9	22.3 - 164.9
Cholesterol in Very Large HDL	302.8	71.5 - 302.8
Free Cholesterol in Very Large HDL	182.2	24.2 - 182.2
Cholesteryl Esters in Very Large HDL	195.3	32.3 - 195.3
Phospholipids in Very Large HDL	201.3	35.9 - 201.3
Triglycerides in Very Large HDL	263.4	60.3 - 263.4
Total lipids in Very Large HDL	421.1	153.1 - 421.1
Small HDL		
Concertation of Small HDL	670.5	318.7 - 670.5
Cholesterol in Small HDL	669.7	318.1 - 669.7
Free Cholesterol in Small HDL	648.3	300.2 - 648.3
Cholesteryl Esters in Small HDL	647.7	311.4 - 647.7
Phospholipids in Small HDL	631.9	283 - 631.9
Triglycerides in Small HDL	248	46.7 - 248
Total lipids in Small HDL	749	347 - 749

Multi-metabolite signatures and incident heart failure

Heart failure metabolic susceptibility score

Based on 10 metabolite measures which association with HF was not explained by association with risk of predisposing interim events, a HFMSS was calculated to represent a joint metabolic risk of incident HF. As both DHA and total omega-3 FAs showed similar magnitude of association with HF risk, for calculation of score DHA was used as a more specific marker. Regarding very large and small HDL particles, considering comparable predictive ability of particle concentration and lipid content, the first one was used for calculation of the score. Two types of HFMSS were calculated – unweighted by summing standardized concentration of contributing metabolite measures (measures inversely associated with HF risk were summed with negative sign) and weighted, where weights were derived from a Cox PH model with ridge penalization. The aim of this approach was to adjust weights, while forcing all variables to remain in the model. HFMSS scores for consistency were rescaled to 0-100. Calculations of both HFMSS versions is summarized in **Table 4**.

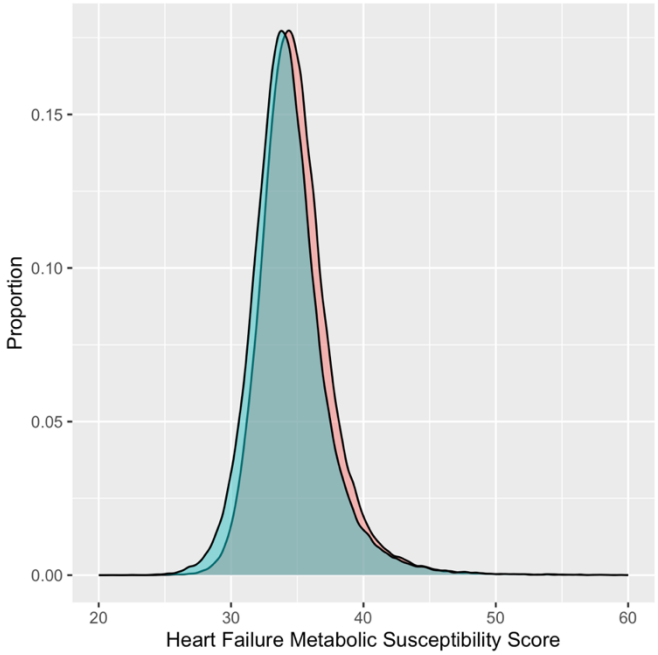
Table 4. Composition and coefficients of heart failure metabolic susceptibility scores.

Metabolite measure	Unweighted HFMSS	Weighted HFMSS
Alanine	-1	-0.003
Acetate	+1	0.020
Acetoacetate	+1	0.028
Acetone	+1	0.021
3-hydroxybutyrate	+1	0.017
Albumin	-1	-0.041
Concentration of large HDL particles	+1	0.011
Concentration of small HDL particles	-1	-0.009
Docosahexaenoic acid	-1	-0.015
Histidine	-1	0.024

Score are calculated by summing standardized concentration of metabolite measures multiplied by coefficients from a table. Weighted score coefficients are derived from Cox proportional hazards regression with ridge penalization.
Abbreviations: HDL – high density lipoprotein, HFMSS – heart failure susceptibility score.

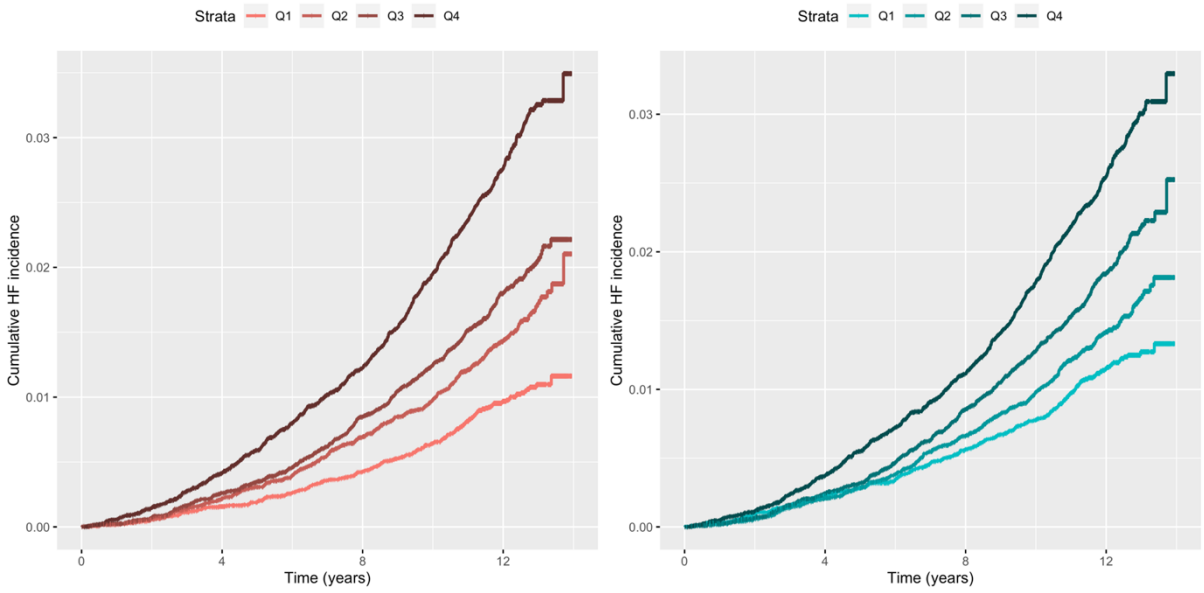
Distribution of unweighted and weighted HFMSS is presented in **Figure 32**. The median value of unweighted HFMSS was 34.1 (IQR: 32.6-35.7) and for weighted HFMSS was 34.6 (IQR: 33.2-36.2) points. Cumulative incidence of HF stratified by quartiles of unweighted and weighted HFMSS is presented in **Figure 33**. Incidence of HF gradually increased across quartiles of both scores.

Figure 32. Histogram presenting distribution of unweighted and weighted heart failure metabolic susceptibility scores.



Blue color denotes unweighted scores, while red color denotes weighted heart failure susceptibility scores.

Figure 33. Kaplan-Meier curve of cumulative incidence of heart failure stratified by quartiles of weighted (left pane) and (right pane) unweighted heart failure metabolic suspectibility scores.



Associations between HFMSS and incident HF are presented in **Table 5**. In Model 1 and Model 2, a 10-point increase in unweighted HFMSS was respectively associated with a 77% (95%CI: 58%-97%, P<2 x 10^{-16}) and 51% (95%CI: 34%-70%, P=2.4 x 10^{-11}) higher HF risk. For the

fourth versus first quartile, the HRs were 1.75 (95%CI: 1.51-2.02, P=9.2 x 10^{-14}) and 1.48 (95%CI: 1.27-1.72, P=2.7 x 10^{-7}) in Model 1 and 2 respectively.

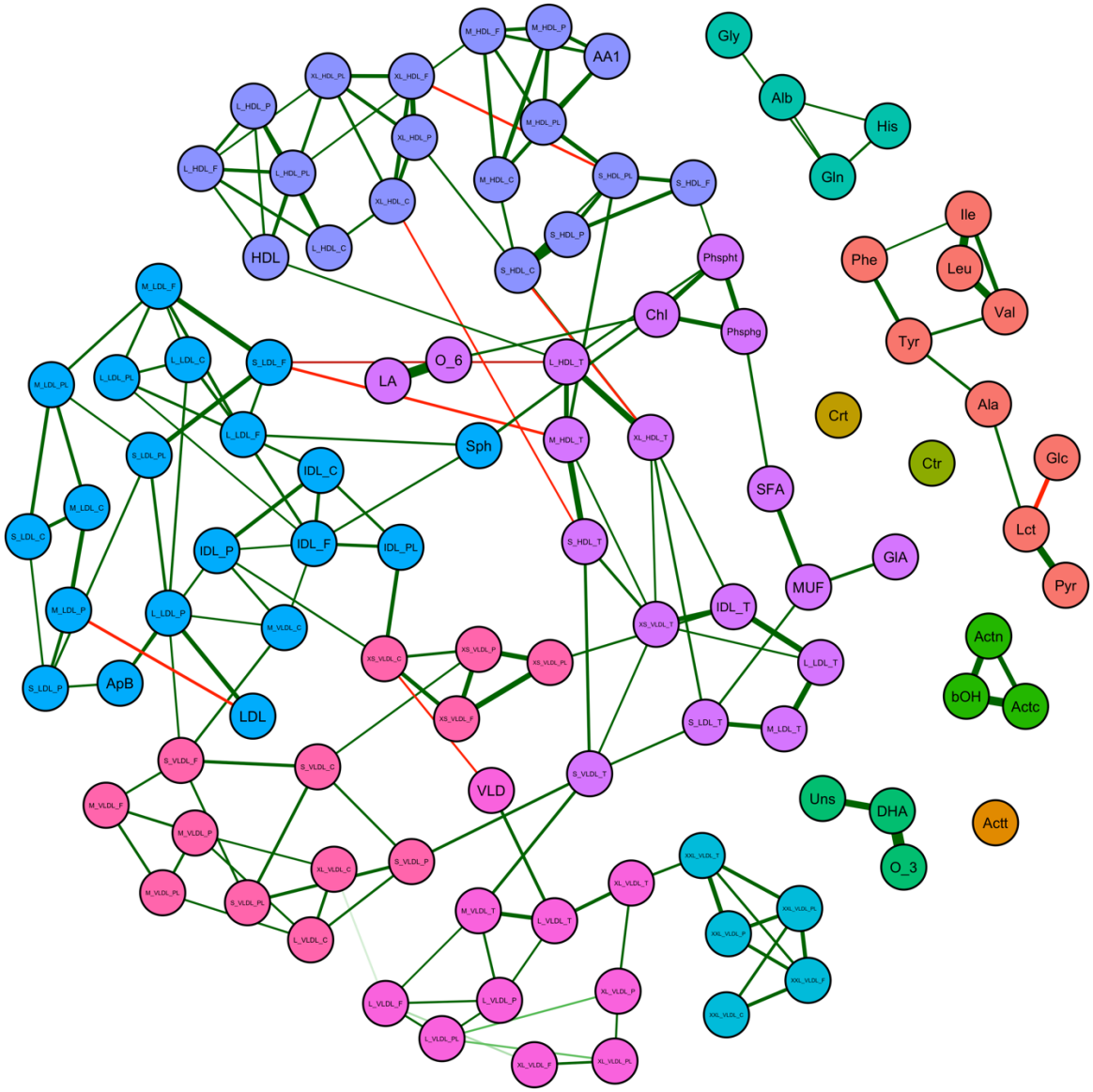
The magnitude of the association was larger for weighted HFMSS. A 10-point increase in weighted HFMSS was associated with a 94% and 69% higher HF risk in Model 1 and 2, respectively. When comparing the fourth versus the first quartile, the HRs were 2.12 (95%CI: 1.82-2.47, P<<2 x 10^{-16}) and 1.88 (95%CI: 1.61-2.20, P=1.1 x 10^{-11}).

Table 5. Associations of unweighted and weighted heart failure susceptibility scores with incident heart failure.

Score	Model 1			Model 2		
	HR	**95% CI**	**P-value**	**HR**	**95% CI**	**P-value**
Unweighted HFMSS						
Continuous						
Per 10-point increase	1.77 (1.58-1.97)		<2 x 10^{-16}	1.51 (1.34-1.70)		2.4 x 10^{-11}
Per 1-SD increase	1.19 (1.15-1.23)		<2 x 10^{-16}	1.13 (1.09-1.17)		2.4 x 10^{-11}
Categorical						
[0,32.6]	Ref.		—	Ref.		—
(32.6, 34.1]	1.11 (0.94-1.31)		0.200	1.07 (0.91-1.25)		0.400
(34.1, 35.7]	1.33 (1.14-1.55)		3.0 x 10^{-4}	1.21 (1.04-1.41)		0.016
(35.7,100]	1.75 (1.51-2.02)		9.2 x 10^{-14}	1.48 (1.27-1.72)		2.7 x 10^{-7}
Weighted HFMSS						
Continuous						
Per 10-point increase	1.94 (1.75-2.14)		<2 x 10^{-16}	1.69 (1.52-1.88)		<2 x 10^{-16}
Per 1-SD increase	1.21 (1.17-1.24)		<2 x 10^{-16}	1.16 (1.13-1.20)		<2 x 10^{-16}
Categorical						
[0,33.2]	Ref		—	Ref.		—
(33.2,34.7]	1.35 (1.14-1.59)		5.0 x 10^{-4}	1.31 (1.11-1.55)		0.002
(34.7,36.3]	1.47 (1.25-1.73)		3.2 x 10^{-6}	1.40 (1.19-1.65)		4.5 x 10^{-5}
(36.3,100]	2.12 (1.82-2.47)		<2 x 10^{-16}	1.88 (1.61-2.20)		1.1 x 10^{-11}

Estimates are derived from a Cox proportional hazards model including score at a time and further adjusted.
Model 1 was adjusted for age, sex, body mass index.
Model 2 included covariates from Model1 and further self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication.

Figure 34. Networks of metabolite measures derived by Gaussian Graphical Model.



Nodes represent different metabolite measures, while edges represent conditional associations measured by partial correlation coefficients. Thickness of edges is proportional to the magnitude of partial correlation coefficients. Green edges represent positive partial correlations, while red negative partial correlations. Partial correlations coefficients $|p| < 0.1$ are not plotted. Colors denote cluster membership derived by walt-trap algorithm run on the top of derived networks.
Abbreviations: AA1 – Apolipoprotein A1, Actc – Acetoacetate, Actn – Acetone, Actt - Acetate, Alb – Albumin, Ala – Alanine, ApB – Apolipoprotein B, bOH – 3-hydroxybutyrate, Chl – Cholines, Crt – Creatinine, Ctr – Citrate, GlA – Glycoprotein Acetyls, Glc- Glucose, Gly – Glycine, Gln – Glutamine, His – Histidine, Ile – Isoleucine, Lct – Lactate, Leu – Leucine, LDL – LDL size, MUF – monounsaturated fatty acids, O_3/6 – Omega-3/6 Fatty Acids, Phe – Phenylalanine, Phsphg – phosphoglycerides, Phspht – Phosphatidylcholines, Pyr – Pyruvate, SFA – saturated fatty acids, Sph – Sphingomyelins, Tyr – Tyrosine, Val – Valine, VLD – VLDL size. *Lipoprotein measures*: (1)X_(2)LDL_(3)X – 1) size subclass (XXL, XL, L, M, S, XS), 2) lipoprotein class (VLDL, LDL, HDL), 3) fraction (_C – cholesteryl esters – free cholesterol, _P – particles, _PL – phospholipids, _T – triglycerides.

Figure 34 presents partial correlation network of 106 primary metabolite measures derived by GGMs, further dived into ten metabolite measure clusters using walk-trap algorithm. Five metabolite measure networks were produced, and additionally three metabolite measures (Acetate, Citrate and Creatinine) were left as standalone nodes.

The largest network included 85 metabolite measures related to various lipoproteins and lipid classes and was further diced into six clusters. The first cluster identified as "*Cholesterol-enriched Lipoproteins*" comprised 20 metabolite measures encompassing average diameter for LDL, apoB, SMs, and various sub-components of lipoproteins. Specifically, these included measures related to IDL/large/medium/small LDL (particle concentration, phospholipids, cholesteryl esters, and free cholesterol), and medium VLDL (cholesteryl esters). The second cluster, identified as *"Triglyceride-rich Lipoprotein Remnants"* comprised 13 metabolite measures related to extra-large and large VLDL (cholesteryl esters), and medium/small/very small VLDL (particle concentration, phospholipids, cholesterol esters, and free cholesterol. The third cluster, identified as *"Triglyceride-rich Lipoproteins*" comprised 10 metabolite measures including VLDL particle size and measures related to extra-large/large VLDL (particle concentration, phospholipids, free cholesterol, and triglycerides) and triglycerides in medium VLDL. The fourth cluster identified as *"Chylomicrons and Extremely Large VLDL"* comprised five metabolite measures related to the XXL VLDL subclass - particle concentration phospholipids cholesterol esters free cholesterol and triglycerides. The fifth cluster, identified as "*High-Density Lipoproteins*" comprised 18 metabolite measures and captured HDL size, apoA1 concentration, extra-large/large/medium/small/very small HDL (particle concentration, phospholipids, cholesteryl esters, and free cholesterol). The sixth cluster, identified as "*Triglycerides, Fatty Acids and Other Lipids*," encompassed 18 metabolite measures. This cluster featured phosphoglycerides, cholines, PCs, omega-6 FAs, linoleic acid, PUFA, SFA, and glycoprotein acetyls. Furthermore, it captured triglyceride metrics across various lipoprotein subclasses, specifically in small/extra-small VLDL, IDL, large/medium/small LDL and extra-large/large/medium small HDL.

Each of four remaining networks represented separate individual clusters. The second largest network and cluster identified as "*Amino acids-Energy metabolism*" included nine metabolite measures – among them BCAAs (isoleucine, leucine, valine) aromatic amino acids (phenylalanine, tyrosine), and energy-metabolism measures (alanine, glucose, lactate, pyruvate). Another network including four metabolite measures albumin, histidine, glutamine,

and glycine was identified as "*Albumin-Amino acids*" cluster. Lastly, two three component networks were identified as "*Ketone bodies*" (comprised acetoacetate, acetone and 3-hydrobutyrate) and "*Omega-3 Fatty Acids*" (comprised: DHA, unsaturation degree and omega-3 FAs).

In the next step, we entered metabolite measures belonging to specific clusters as predictors for Cox regression model with elastic net regularization to select the most important metabolite measures for HF risk as derive cluster-specific weighted metabolite measure scores. **Table 6** presents metabolite measures selected for each of cluster along with their weights. Cluster-specific scores were then calculated as weighted sum of standardized metabolite measure levels.

Based on results from Cox models with elastic net regularization, the "Amino acid-Energy metabolism" cluster-specific score was reduced to four metabolite measures. It showed positive loadings for tyrosine, glucose, and lactate, and negative loadings for alanine. The "Ketone bodies", "Omega-3 Fatty Acid", and "Chylomicrons and extremely large VLDL" cluster-specific scores were based on single metabolite measures: acetoacetate, degree of unsaturation, and cholesteryl esters content, respectively. The "Cholesterol-enriched Lipoproteins" cluster-specific score comprised 14 metabolite measures. It was negatively correlated with the average diameter for LDL, cholesteryl esters in medium VLDL, IDL, and small LDL, while it was positively correlated with lipid content in the remaining lipoprotein subclasses. The "Albumin-Amino acids" cluster-specific score encompassed three metabolite measures: albumin, glycine, and histidine, all of which negatively correlated with the score. The "High-density Lipoproteins" cluster-specific score, consisting of 12 metabolite measures, was positively correlated with free cholesterol in very large/small HDL, cholesteryl esters in medium HDL, and phospholipids in small HDL. Conversely, it was negatively correlated with the average diameter for HDL and the remaining HDL lipid fractions. The "Triglycerides, Fatty Acids and Other Lipids" cluster-specific score included 16 metabolite measures, showing positive loadings for saturated and PUFA, glycoprotein acetyls, and triglycerides in very small VLDL, IDL, large LDL, and HDL, while showing negative loadings for other lipids and triglycerides in the remaining lipoprotein fractions. The "Triglyceride-rich Lipoproteins" cluster-specific score, which included five metabolite measures, showed positive loadings for the concentration of very large VLDL particles, as well as their phospholipids and free cholesterol content. In contrast, it exhibited negative loadings for triglycerides in large and medium VLDL. Lastly,

the “Triglyceride-rich Lipoprotein Remnants” cluster-specific score, which encompassed 11 metabolite measures, was positively correlated with cholesteryl esters in very large/small VLDL, as well as with concentration and free cholesterol in very small VLDL. This score was negatively correlated with other measures, primarily related to medium VLDL particles.

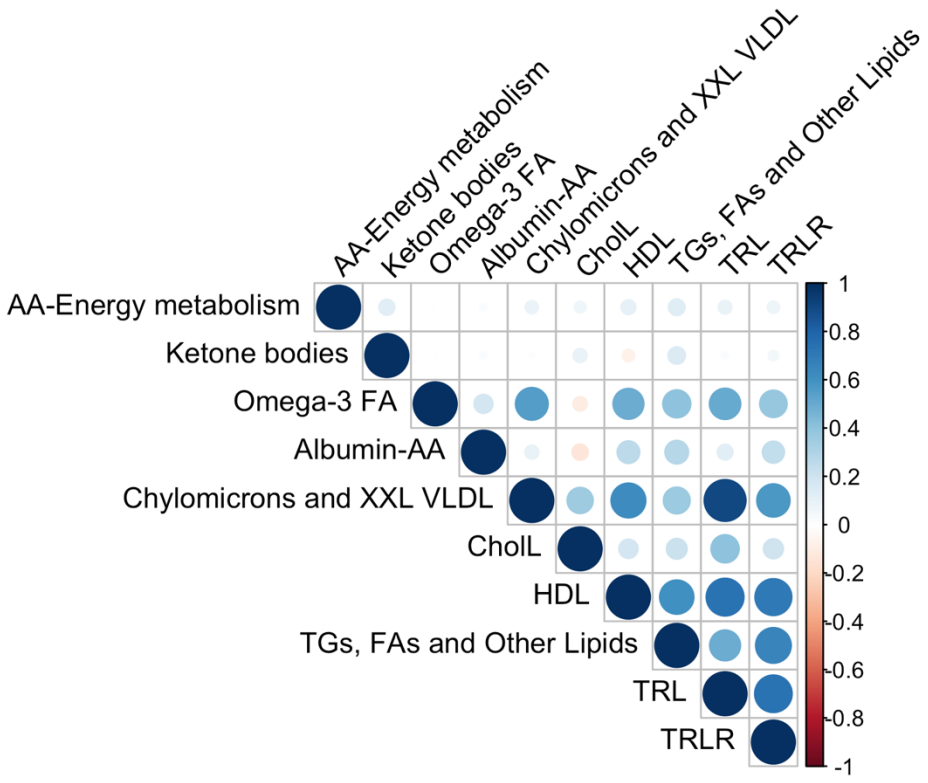
Table 6. Weights for cluster-specific scores derived from Cox regression model with elastic net regularization.

Cluster	**Metabolite measure**	**EN weight**
Amino acids-Energy metabolism	Alanine	-0.059
	Tyrosine	0.112
	Glucose	0.056
	Lactate	0.002
Ketone bodies	Acetoacetate	0.029
Omega-3 Faty acids	Degree of Unsaturation	-0.042
Albumin-Amino acids	Glycine	-0.145
	Histidine	-0.049
	Albumin	-0.163
Chylomicrons and extremely large VLDL	Cholesteryl Esters in Chylomicrons and Extremely Large VLDL	0.021
Cholesterol-enriched Lipoproteins	Average Diameter for LDL Particles	-0.05
	Sphingomyelins	0.332
	Apolipoprotein B	0.065
	Cholesteryl Esters in Medium VLDL	-0.478
	Concentration of IDL Particles	0.647
	Phospholipids in IDL	0.167
	Cholesteryl Esters in IDL	-0.877
	Free Cholesterol in IDL	0.499
	Concentration of Large LDL Particles	0.502
	Phospholipids in Medium LDL	0.191
	Cholesteryl Esters in Medium LDL	0.098
	Free Cholesterol in Medium LDL	0.167
	Phospholipids in Small LDL	0.035
	Cholesteryl Esters in Small LDL	-0.253
High-density Lipoproteins	Average Diameter for HDL Particles	-0.133
	Concentration of Very Large HDL Particles	-0.015
	Cholesteryl Esters in Very Large HDL	-0.574
	Free Cholesterol in Very Large HDL	0.678
	Cholesteryl Esters in Large HDL	-0.190
	Free Cholesterol in Large HDL	0.001
	Phospholipids in Medium HDL	-0.009
	Cholesteryl Esters in Medium HDL	0.170

	Free Cholesterol in Medium HDL	0.001
	Phospholipids in Small HDL	0.139
	Cholesteryl Esters in Small HDL	-0.236
	Free Cholesterol in Small HDL	0.057
Triglycerides, Fatty Acids and Other Lipids	Phosphoglycerides	-0.214
	Total Cholines	-0.664
	Phosphatidylcholines	-0.021
	Omega-6 Fatty Acids	0.055
	Monounsaturated Fatty Acids	0.688
	Saturated Fatty Acids	0.756
	Linoleic Acid	-0.027
	Glycoprotein Acetyls	0.209
	Triglycerides in Small VLDL	-0.149
	Triglycerides in Very Small VLDL	0.037
	Triglycerides in IDL	0.452
	Triglycerides in Large LDL	0.204
	Triglycerides in Small LDL	-1.302
	Triglycerides in Large HDL	0.268
	Triglycerides in Medium HDL	-0.315
	Triglycerides in Small HDL	-0.006
Triglyceride-rich Lipoproteins	Concentration of Very Large VLDL Particles	0.054
	Phospholipids in Very Large VLDL	0.177
	Free Cholesterol in Very Large VLDL	0.172
	Triglycerides in Large VLDL	-0.182
	Triglycerides in Medium VLDL	-0.099
Triglyceride-rich Lipoprotein Remnants	Cholesteryl Esters in Very Large VLDL	0.559
	Cholesteryl Esters in Large VLDL	-0.102
	Concentration of Medium VLDL Particles	-0.076
	Phospholipids in Medium VLDL	-0.469
	Free Cholesterol in Medium VLDL	-0.074
	Concentration of Small VLDL Particles	-0.326
	Cholesteryl Esters in Small VLDL	0.239
	Concentration of Very Small VLDL Particles	0.022
	Phospholipids in Very Small VLDL	0.101
	Cholesteryl Esters in Very Small VLDL	-0.410
	Free Cholesterol in Very Small VLDL	0.656

Cluster-specific scores are calculated by summing standardized concentration of metabolite measures multiplied by coefficients from a table. Weighted score coefficients are derived from Cox proportional hazards regression with elastic net penalization.

Figure 35. Heatmap of correlations between ten derived cluster-specific scores.



Abbreviations: AA – Amino acids, CholL = Cholesterol-enriched lipoproteins, FA – Fatty acids, TRL – Triglyceride-rich lipoproteins, TRLR – Triglyceride-rich lipoprotein remnants

Correlations between cluster-specific scores were presented in **Figure 35**. "Amino acids-Energy metabolism" and "Ketone bodies" clusters were not associated with other cluster scores. 'Albumin-Amino acids" and "Cholesterol-enriched Lipoproteins" clusters showed very low to low correlation with other clusters. Moderate to high positive correlation coefficients (p=0.4-0.9) were observed between "Omega-3 Fatty Acids", "Chylomicrons and Extremely Large VLDL", "High Density Lipoproteins", "Triglycerides, Fatty acids and Other lipids", "Triglyceride-rich Lipoproteins" and 'Triglyceride-rich Lipoprotein Remnants" cluster-specific scores.

Table 7. Hazard ratios (95% CI) for incident heart failure associated with one-standard deviation increase in metabolite measure cluster-specific scores.

Cluster-specific score	**Model 1**		**Model 2**	
	HR (95% CI)	**P-value**	**HR (95% CI)**	**P-value**
Amino acids-Energy metabolism	1.08 (1.03-1.13)	0.003	1.07 (1.02-1.12)	0.008
Ketone bodies	1.11 (1.07-1.14)	4.3×10^{-11}	1.08 (1.04-1.11)	1.6×10^{-5}
Omega-3 FA	1.15 (1.09-1.21)	7.1×10^{-8}	1.09 (1.03-1.14)	0.002

Albumin-Amino Acids	1.23 (1.16-1.29)	1.1×10^{-14}	1.19 (1.13-1.25)	8.6×10^{-11}
Chylomicrons and XXL VLDL	0.99 (0.94-1.04)	0.634	0.97 (0.92-1.02)	0.193
Cholesterol-enriched Lipoproteins	0.99 (0.94-1.04)	0.697	0.99 (0.94-1.04)	0.645
High-density Lipoproteins	1.12 (1.06-1.18)	2.5×10^{-5}	1.09 (1.04-1.15)	0.001
Triglycerides, Fatty Acids and Other Lipids	1.40 (1.34-1.47)	1.4×10^{-42}	1.29 (1.23-1.36)	8.5×10^{-23}
Triglyceride-rich Lipoproteins	1.05 (1.00-1.11)	0.036	1.03 (0.98-1.08)	0.279
Triglyceride-rich Lipoproteins Remnants	1.20 (1.14-1.25)	1.0×10^{-14}	1.15 (1.10-1.21)	3.5×10^{-9}

Estimates are derived from a Cox proportional hazards model including one cluster-specific score at a time and further adjusted.
Model 1 was adjusted for age, sex, body mass index.
Model 2 included covariates from Model1 and further self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication.

Associations between cluster-specific scores and incident HF are presented in **Table 7**. In model adjusted for age, sex and BMI, all scores with exception for "Chylomicrons and Extremely Large VLDL" (HR: 0.99, 95%CI: 0.94-1.04, P=0.634) and "Cholesterol-Enriched Lipoproteins" (HR: 0.99, 95%CI: 0.94-1.04, P=0.697) were associated with HF risk. Following adjustment for further covariates in Model 2, all associations remained statistically significant with exception for "Triglyceride-rich Lipoproteins" (HR: 1.03, 95%CI: 0.98-1.08, P=0.279). The strongest associations were observed for "Triglycerides, Fatty Acids and Other Lipids" (HR: 1.29, 95%CI: 1.23-1.36, $P=8.5 \times 10^{-23}$), "Albumin-Amino Acids" (HR: 1.19, 95%CI: 1.13-1.25), $P=8.6 \times 10^{-11}$) and "Triglyceride-rich Lipoproteins Remnants" (HR: 1.15, 95%CI: 1.10-1.21), $P=3.5 \times 10^{-9}$). 1-SD increase in "Amino acids-Energy metabolism", "Ketone bodies", "Omega-3 Fatty Acids" and "High-density Lipoproteins" scores were each associated with 7-9% increase in HF risk.

Table 8. Associations between cluster-specific scores and incident heart failure adjusted for interim events.

Cluster-specific score	**Model 2 + interim CAD**		**Model 2 + interim AF**		**Model 2 + interim VHD**		**Model 2 + interim DM**	
	HR	**P-value**	**HR**	**P-value**	**HR**	**P-value**	**HR**	**P-value**
Amino acids-Energy metabolism	1.07	0.01	1.07	0.009	1.07	0.009	1.06	0.012
Ketone bodies	1.08	7.7×10^{-6}	1.07	6.4×10^{-5}	1.07	3.4×10^{-5}	1.08	1.4×10^{-5}
Omega-3 FA	1.07	0.006	1.09	0.001	1.08	0.002	1.08	0.003
Albumin-Amino acids	1.17	3.6×10^{-9}	1.17	1.8×10^{-9}	1.18	8.3×10^{-10}	1.19	1.0×10^{-10}

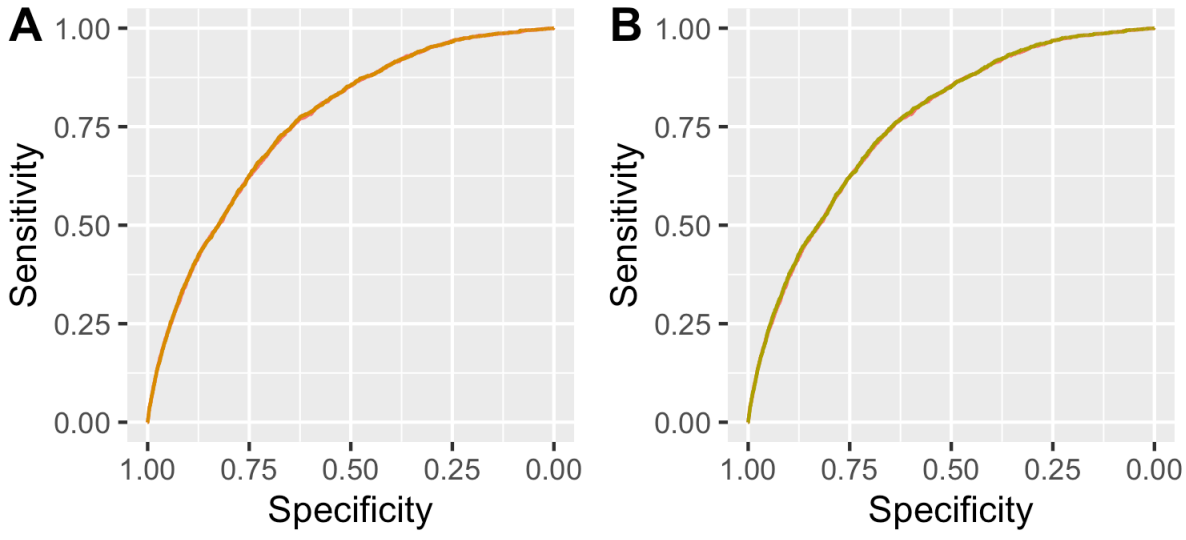
Chylomicrons and XXL VLDL	0.96	0.12	0.99	0.667	0.97	0.257	0.96	0.165
Cholesterol-enriched Lipoproteins	0.98	0.381	1.01	0.708	0.99	0.666	0.99	0.678
High-density Lipoproteins	1.07	0.008	1.11	1.2×10^{-4}	1.09	0.001	1.09	0.001
Triglycerides, Fatty Acids and Other Lipids	1.27	1.7×10^{-19}	1.29	8.4×10^{-23}	1.28	2.8×10^{-22}	1.29	1.8×10^{-22}
Triglyceride-rich Lipoproteins	1.02	0.479	1.04	0.087	1.03	0.27	1.03	0.321
Triglyceride-rich Lipoproteins Remnants	1.14	5.1×10^{-8}	1.15	1.9×10^{-9}	1.14	1.6×10^{-8}	1.15	6.1×10^{-9}

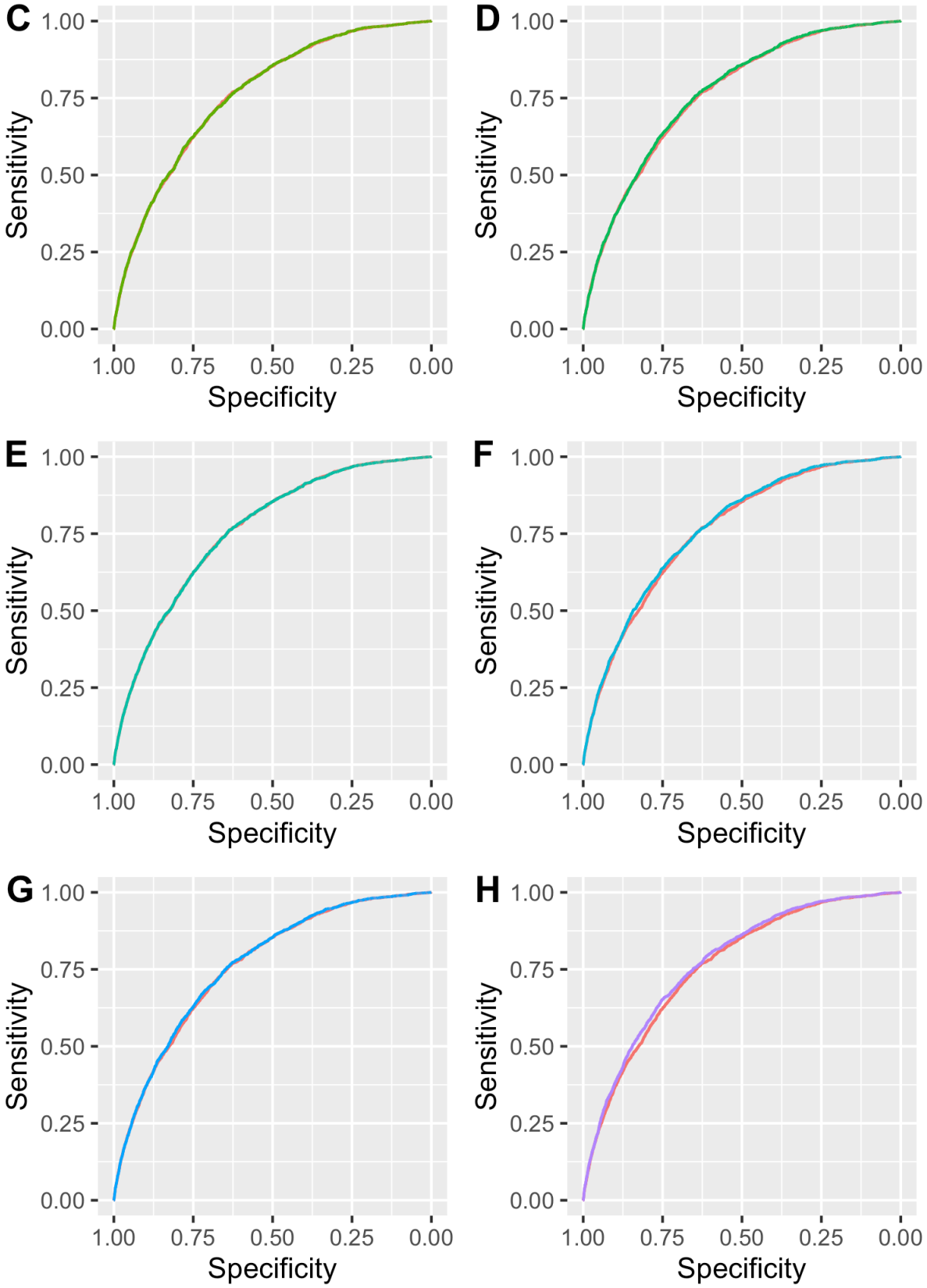
Estimates are derived from a Cox proportional hazards model including one cluster-specific score at a time and further adjusted for age, sex, body mass index, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication, as well as separately for each type of interim event
Abbreviations: CAD – coronary artery disease, AF – atrial fibrillation, VHD – valvular heart disease, DM – diabetes mellitus

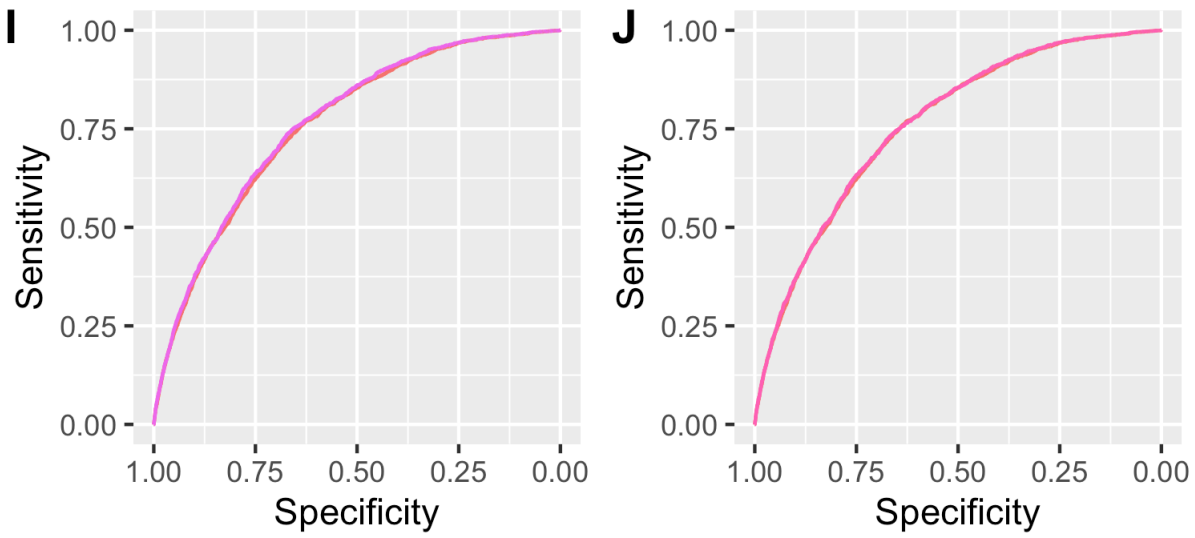
Further adjustment for interim events such as CAD, AF, VHD, and DM did not influence the main findings of the cluster-specific score-HF associations (**Table 8**).

Prediction performance of multi-metabolite measure scores for heart failure

Figure 36. Receiver operator characteristic curves for HF prediction models with and without multi-metabolite measure scores.




C
D
E
F
G
H
Sensitivity
Specificity
1.00
0.75
0.50
0.25
0.00



Red curve in each of figure panels indicates conventional risk factors model including age, sex, age, sex, body mass index, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP, and use of anti-hypertensive medication.
Panel A – second curve indicates conventional model + Amino acids-Energy metabolism score;
Panel B – second curve indicates conventional model + Ketone bodies score;
Panel C – second curve indicates conventional model + Omega-3 Fatty Acids score;
Panel D – second curve indicates conventional model + Albumin-Amino acids score;
Panel E – second curve indicates conventional model + High-density Lipoproteins score;
Panel F – second curve indicates conventional model + Triglycerides, Fatty Acids and Other Lipids score;
Panel G – second curve indicates conventional model + Triglyceride-rich Lipoproteins Remnants score;
Panel H – second curve indicates conventional model + 7 cluster-specific scores;
Panel I – second curve indicates conventional model + Weighted HFMSS;
Panel J – second curve indicates conventional model + Unweighted HFMSS.

Comparison of receiver operator characteristic curves for HF prediction models including multi-metabolite measure scores is presented in **Figure 36**. Respective AUC, NRI and IDI values are summarized in **Table 9**. The addition of the "*Amino acids-Energy metabolism*" score to the conventional risk factors model marginally improved the AUC (0.763 vs. 0.762, P=0.023) and NRI (0.07, P=0.004), but not in IDI (P=0.353). Incorporating the "*Ketone bodies score*" led to an small increase in AUC (0.763 vs. 0.762, P=0.003), NRI (P=0.001), but not IDI (P=0.071). With addition of "*Omega-3 Fatty Acids*" score to the conventional risk factor model, there was no significant change in AUC (P=0.134) nor IDI (P=0.156), but it improved NRI (P=0.001). Incorporating the "*Albumin-Amino acids*" score to the base model showed improvements in AUC (0.765 vs. 0.762, P=1.4 x 10^{-4}), NRI (P=1.2 x 10^{-7}), and IDI (P=0.001). The "*High-density Lipoproteins*" score added to the conventional risk factors did not significantly affect the AUC (P=0.536) and moreover reduced the NRI (-0.05, P=0.049) while affecting IDI (P=0.033). Addition of "Triglycerides, Fatty Acids and Other Lipids score"

resulted in significant improvements in AUC (0.768 vs. 0.762, P=4.8 x 10^{-6}), NRI (P=4.5 x 10^{-15}), and IDI (P=6.1 x 10^-7). The addition of "Triglyceride-rich Lipoproteins Remnants" to conventional risk factors improved AUC (0.764 vs. 0.762, P=0.001), NRI (0.13, P=3.8 x 10^{-7}), and IDI (P=0.015). Inclusion of all seven cluster-specific scores showed the largest increase in AUC (0.772 vs. 0.762, P=2.6 x 10^-10), NRI (0.19, P=2.3 x 10^{-14}), and IDI (P=1.7 x 10^{-11}).

Incorporating weighted HFMSS to the conventional risk factors model also indicated improvements in AUC (0.767 vs. 0.762, P=3.4 x 10^{-8}), NRI (P=7.1 x 10^{-8}), and IDI (P=1.7 x 10^-6). Lastly, the model with added unweighted HFMSS displayed an enhanced AUC (0.764 vs. 0.762, P=0.001), NRI (0.08, P=0.002), and IDI (P=0.001).

Head-to-dead comparison of models incorporating multi-metabolite measure scores yielded that model including all seven cluster-specific score had a superior predictive ability (in context of either AUC, NRI, and IDI) over models including individual cluster-specific scores or HFMSS.

Table 9. Comparison of area under curve, net reclassification index, integrated discriminatory improvement parameters for HF prediction models with and without multi-metabolite measure scores.

Model	AUC (95% CI)	P	NRI (95% CI)	P	IDI (95% CI)	P
Conventional risk factors	0.762 (0.751-0.773)	–	–	–	–	–
CRF + Amino acids-Energy metabolism score	0.763 (0.752-0.774)	0.023	0.07 (0.02-0.12)	0.004	7.7 x 10^{-5} (-8.6 x 10^{-5}-2.4 x 10^{-4})	0.353
CRF + Ketone bodies score	0.763 (0.752-0.774)	0.003	0.08 (0.03-0.12)	0.001	2.1 x 10^{-4} (-1.8 x 10^{-5}-4.4 x 10^{-4})	0.071
CRF + Omega-3 Fatty Acids	0.762 (0.752-0.773)	0.134	0.08 (0.03-0.13)	0.001	1.2 x 10^{-4} (-4.4 x 10^{-5}-2.7 x 10^{-4})	0.156
CRF +Albumin-Amino acids score	0.765 (0.755-0.776)	1.4 x 10^{-4}	0.13 (0.08-0.18)	1.2 x 10^{-7}	5.3 x 10^{-4} (2.0 x 10^{-4}-8.5 x 10^{-4})	0.001
CRF + High-density Lipoproteins score	0.762 (0.751-0.773)	0.536	-0.05 (-0.10-0.00)	0.049	7.4 x 10^{-5} (6.1 x 10^{-6}-1.4 x 10^{-4})	0.033
CRF + Triglycerides, Fatty Acids and Other Lipids score	0.768 (0.757-0.779)	4.8 x 10^{-6}	0.20 (0.15-0.24)	4.5 x 10^{-15}	1.3 x 10^{-3} (7.9 x 10^{-4}-1.8 x 10^{-3})	6.1 x 10^{-7}
CRF + Triglyceride-rich Lipoproteins Remnants	0.764 (0.753-0.775)	0.001	0.13 (0.08-0.18)	3.8 x 10^{-7}	3.6 x 10^{-4} (6.9 x 10^{-5}-6.5 x 10^{-4})	0.015
CRF + 7 cluster-specific scores	0.772 (0.762-0.783)	2.6 x 10^{-10}	0.19 (0.14-0.24)	2.3 x 10^{-14}	2.2 x 10^{-3} (1.5 x 10^{-3}-2.8 x 10^{-3})	1.7 x 10^{-11}
CRF + Weighted HFMSS	0.767 (0.756-0.778)	3.4 x 10^{-8}	0.13 (0.09-0.18)	7.1 x 10^{-8}	1.2 x 10^{-3} (6.7 x 10^{-4}-1.7 x 10^{-3})	1.7 x 10^{-6}
CRF + Unweighted HFMSS	0.764 (0.753-0.775)	0.001	0.08 (0.03-0.13)	0.002	6.2 x 10^{-4} (2.6 x 10^{-4}-9.7 x 10^{-4})	0.001

Metabolite markers in relation to cardiac function in magnetic resonance imaging

Table 10 compares baseline characteristics of overall UK Biobank cohort with participants in the sub-cohort with available data on MRI LV function. From the initial cohort, participants who developed HF, CAD, AF, VHD or DM before the date of MRI assessment were excluded. The final sub-cohort comprised of 7751 participants who underwent the MRI without suspected conditions that could influence LV function. Compared with overall cohort, participants in the sub-cohort were younger, more likely to be of White ethnicity, come from more deprived neighborhood, had higher alcohol intake and simultaneously less likely to use anti-hypertensive medication or currently smoke, as well as had lower BMI.

Table 10. Comparison of participants' baseline characteristics between overall cohort and sub-cohort who underwent heart magnetic resonance imaging assessment.

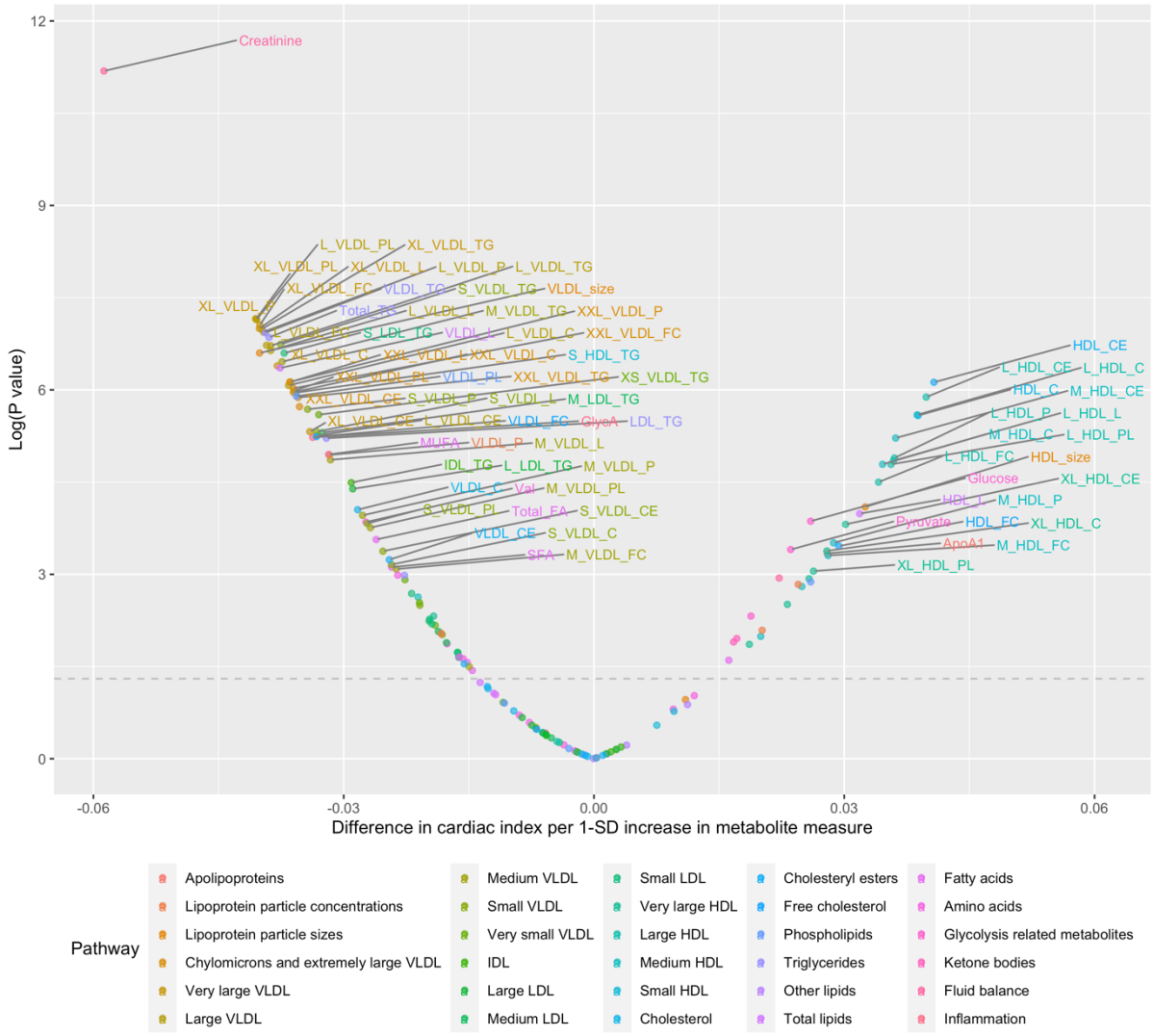
Characteristic	**Overall cohort**	**Subcohort with available heart MRI**
n	93,922	7751
Age, years	55 (8)	54 (7)
Female sex	54,223 (58%)	4,383 (57%)
Body mass index, kg/m2	26.9 (4.6)	26.1 (4.1)
White ethnicity	89,270 (95%)	7,543 (97%)
Townsend Deprivation Index	-1.40 (3.04)	-1.99 (2.70)
High physical activity status	31,635 (41%)	2,660 (40%)
Current smokers	9,813 (10%)	422 (5.4%)
Healthy Diet Score ≥5	19,600 (21%)	1,762 (23%)
Fasting, hrs	3 (2, 4)	3 (2, 4)
eGFR, mL/min/1.73 m^2	87 (77, 98)	87 (78, 97)
Alcohol intake, g/day	13 (1, 28)	15 (4, 29)
HbA1c, mmol/mol	34.7 (32.4, 37.0)	34.2 (31.9, 36.5)
Systolic blood pressure, mmHg	135 (124, 148)	132 (121, 144)
Diastolic blood pressure, mmHg	82 (75, 89)	80 (74, 88)

Characteristic	Overall cohort	Subcohort with available heart MRI
Use of anti-hypertensive medication	10,965 (12%)	566 (7.3%)

Abbreviations: eGFR – estimated glomerular filtration rate, HbA1c – hemoglobin A1c, MRI – magnetic resonance imaging

Metabolome-wide analysis for automatically annotated MRI parameters of left ventricle

Figure 37. Volcano plot summarizing associations between 168 non-ratio metabolite measures with cardiac index.



Estimates are derived from a general linear model including one metabolite measure at a time and further adjusted for age, sex, BMI, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, use of anti-diabetic medication, SBP, DBP, use of anti-hypertensive medication, non-HDL-C, HDL-C, TG, use of lipid-lowering medication, and prevalent CAD/AF/VHD

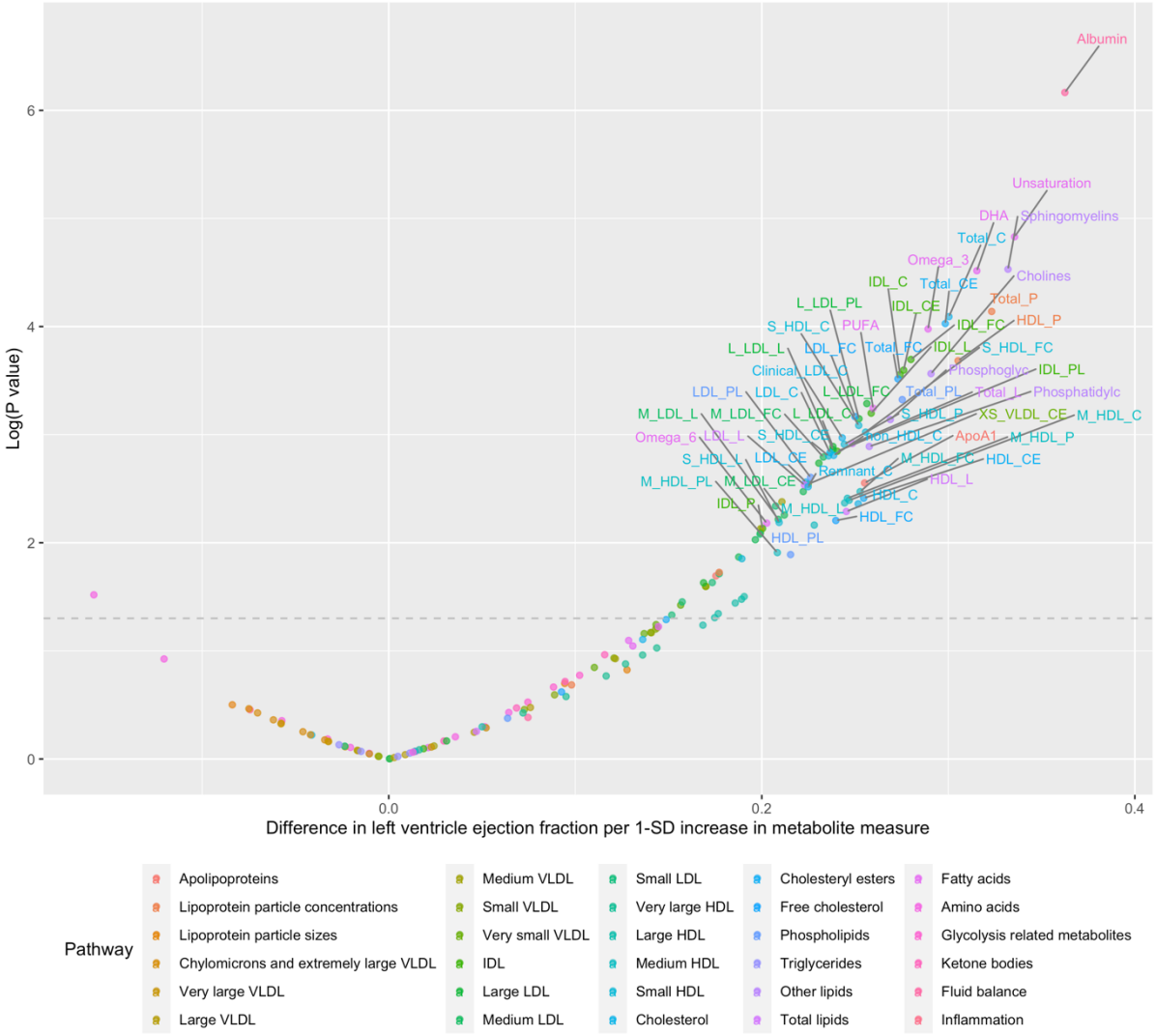
Dashed line represents significance level P=0.05.

Labeled are metabolite measures with strongest positive and negative associations with the endpoint.

Figure 37 summarizes association between 168 non-ratio metabolite measures and CI. Thirty-four metabolite measures showed significant association with higher CI, while 83 showed significant association with lower CI. The largest decrease in CI was associated with higher concentration of creatinine (-0.058 L/min/m^2 per each 1-SD increase in concentration). Inverse associations with CI were observed for 39 measures related to whole size spectrum of VLDL particles, as well as overall VLDL particle count and contents of all lipid fractions (β: -0.041 – -0.023). Furthermore, inverse associations were found for triglyceride content in IDL and LDL size subclasses (β: -0.037 – -0.029). Lower CI was also associated with higher concentration of individual and total BCAAs (β: -0.027 – -0.016), FA measures (with exception for degree of unsaturation with showed a positive association with the measure), different cholesterol fractions in LDL size subclasses and apoB (β: -0.037 – -0.016). In contrast, higher CI was with linked to higher concentration of total and esterified cholesterol in HDL and 24 other measures (β: 0.019 – 0.040) related to large and medium HDL particles (with exception for HDL triglycerides). This was consistent with a positive association observe for higher concentration of apoA1 (β: 0.028) and larger average diameter of HDL (β: 0.033). Higher concentration of glycolysis-related metabolite measures glucose, lactate and pyruvate, ketone bodies, acetone and 3-OHB, as well as albumin were also linked with higher CI (β: 0.017 – 0.026).

Associations between metabolite measures and LVEF is presented in **Figure 38**. Eighty-one out of 168 non-ratio metabolite measures showed significant association with LVEF including 80 associated with higher and one with lower LVEF. Inverse association with this measure was observed for higher concentration of phenylalanine (-0.16% per each 1-SD increase in concentration). The strongest association with LVEF was observed for higher circulating albumin (β: 0.36). Moreover, positive association with LVEF was identified for measures linked with PUFA (β: 0.20 – 0.34; degree of unsaturation, total concentration of omega-3/6 FAs) as well as other lipids (β: 0.26 – 0.33; cholines, SMs, phosphoglycerides and PCs). Higher LVEF was also linked with a higher particle concentration and lipid content in IDL and LDL size subclasses – 31 metabolite measures (β: 0.15 – 0.28). Lastly, higher apoA1 and 24 measures related to large, medium, and small, but not very large HDL particles were linked with a higher LVEF (β: 0.18 – 0.31).

Figure 38. Volcano plot summarizing associations between 168 non-ratio metabolite measures with left ventricle ejection fraction.

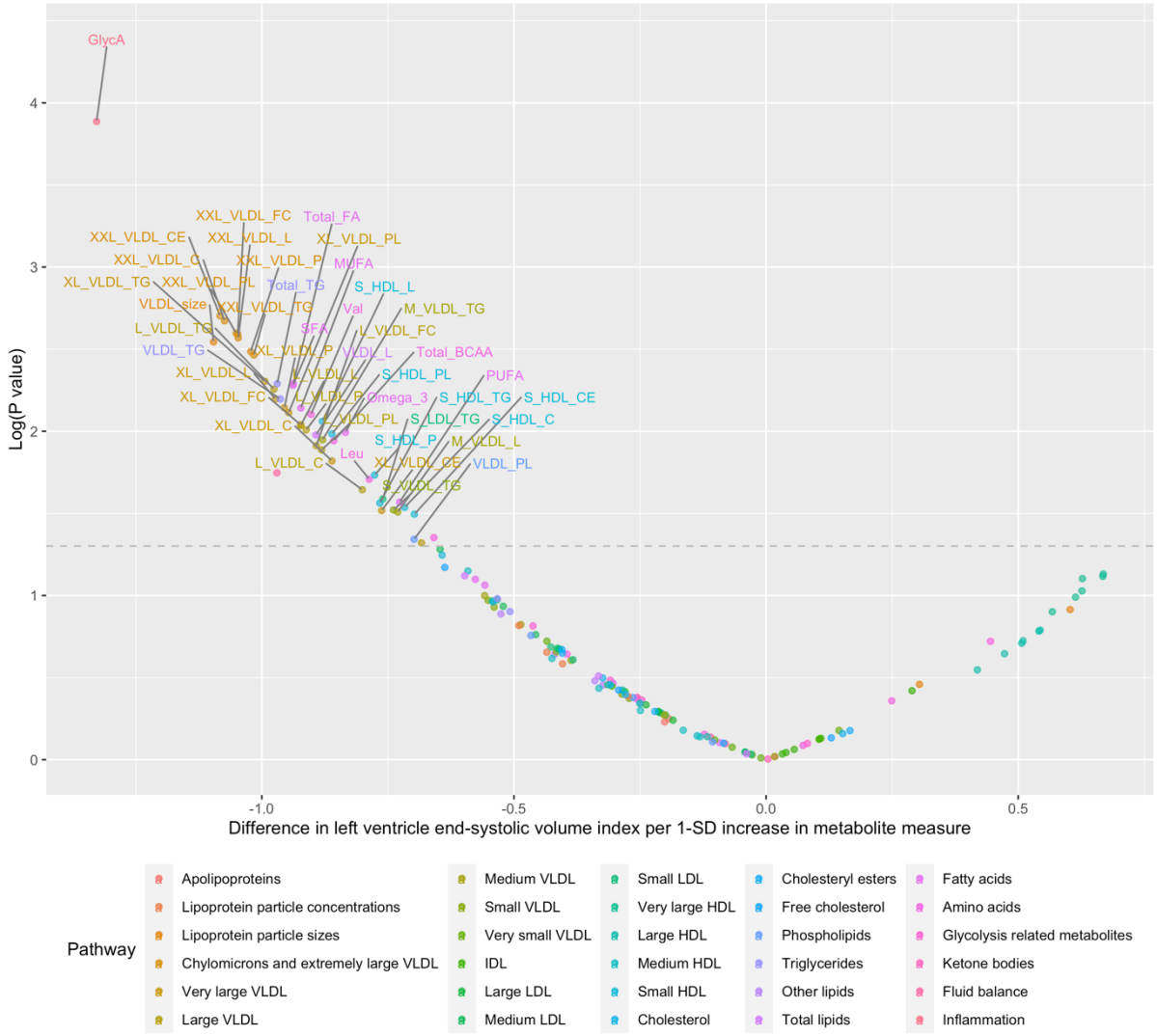


Estimates are derived from a general linear model including one metabolite measure at a time and further adjusted for age, sex, BMI, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, use of anti-diabetic medication, SBP, DBP, use of anti-hypertensive medication, non-HDL-C, HDL-C, TG, use of lipid-lowering medication, and prevalent CAD/AF/VHD
Dashed line represents significance level P=0.05.
Labeled are metabolite measures with strongest positive and negative associations with the endpoint.

Figure 39 summarizes association between 168 non-ratio metabolite measures and LVESVI. Forty-seven metabolite measures were linked with lower LV systolic function. The strongest inverse association was for observed for glycoprotein acetyls (-1.3 mL/m^2 per 1-SD increase in concentration). Twenty-seven of metabolite measures presenting inverse association with LVESVI were linked with VLDL particles. Among them majority belong to chylomicrons, extremely large and large VLDL size subclasses (β: -1.1 – -0.7). Moreover, larger average diameter of VLDL (β: -1.1) was associated with lower LV systolic function. Furthermore,

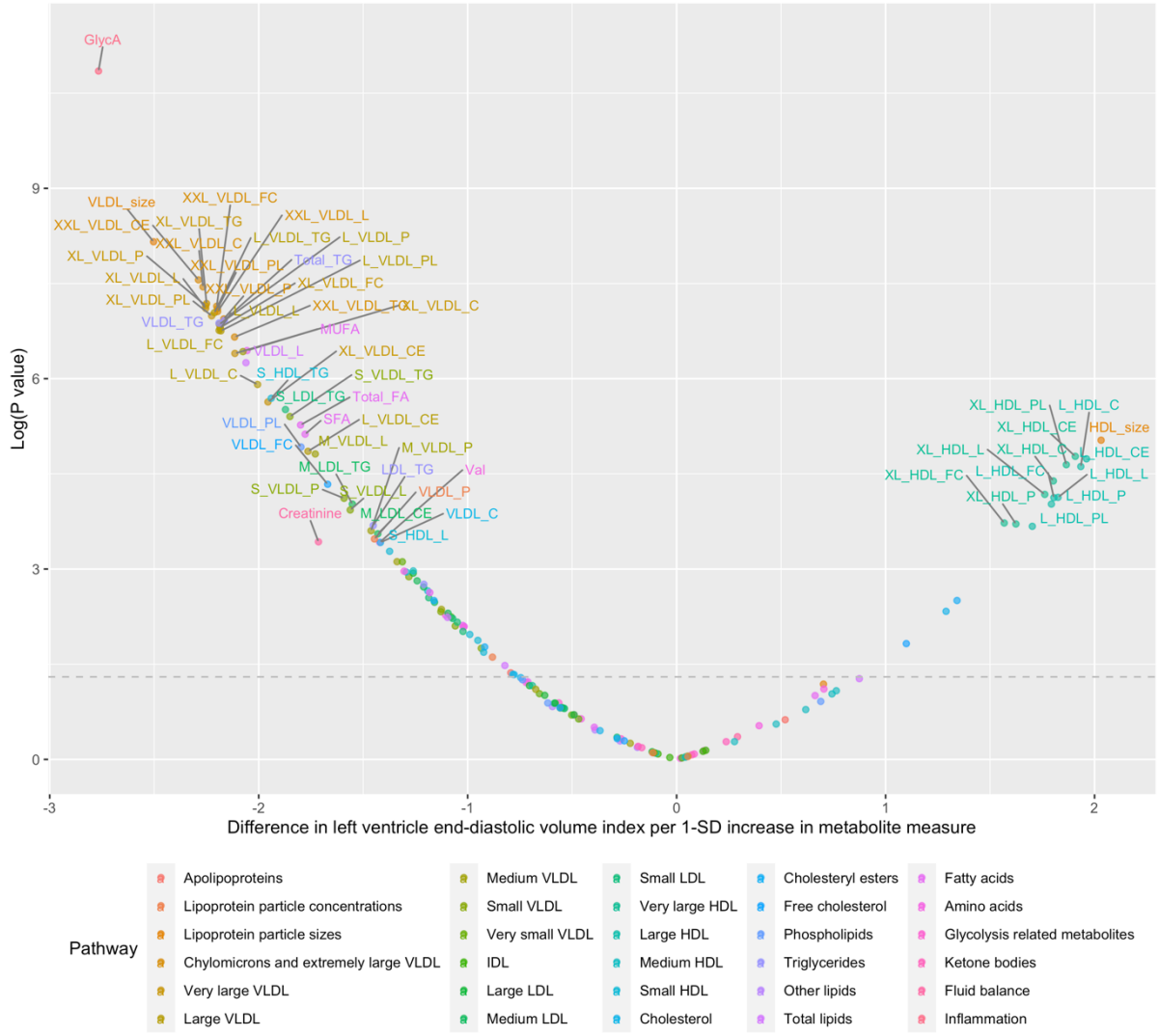
lower LVESVI was found concentration as well as content of all lipid fractions in small HDL particles (β: -0.9 – -0.6). higher concentration of isoleucine, leucine, valine and total BCAAs (β: -0.9 – -0.7) as well as creatinine (β: -1.0) was associated with a lower LVESVI.

Figure 39. Volcano plot summarizing associations between 168 non-ratio metabolite measures with left ventricle end-systolic volume index.



Estimates are derived from a general linear model including one metabolite measure at a time and further adjusted for age, sex, BMI, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, use of anti-diabetic medication, SBP, DBP, use of anti-hypertensive medication, non-HDL-C, HDL-C, TG, use of lipid-lowering medication, and prevalent CAD/AF/VHD
Dashed line represents significance level P=0.05.
Labeled are metabolite measures with strongest positive and negative associations with the endpoint.

Figure 40. Volcano plot summarizing associations between 168 non-ratio metabolite measures with left ventricle end-diastolic volume index.



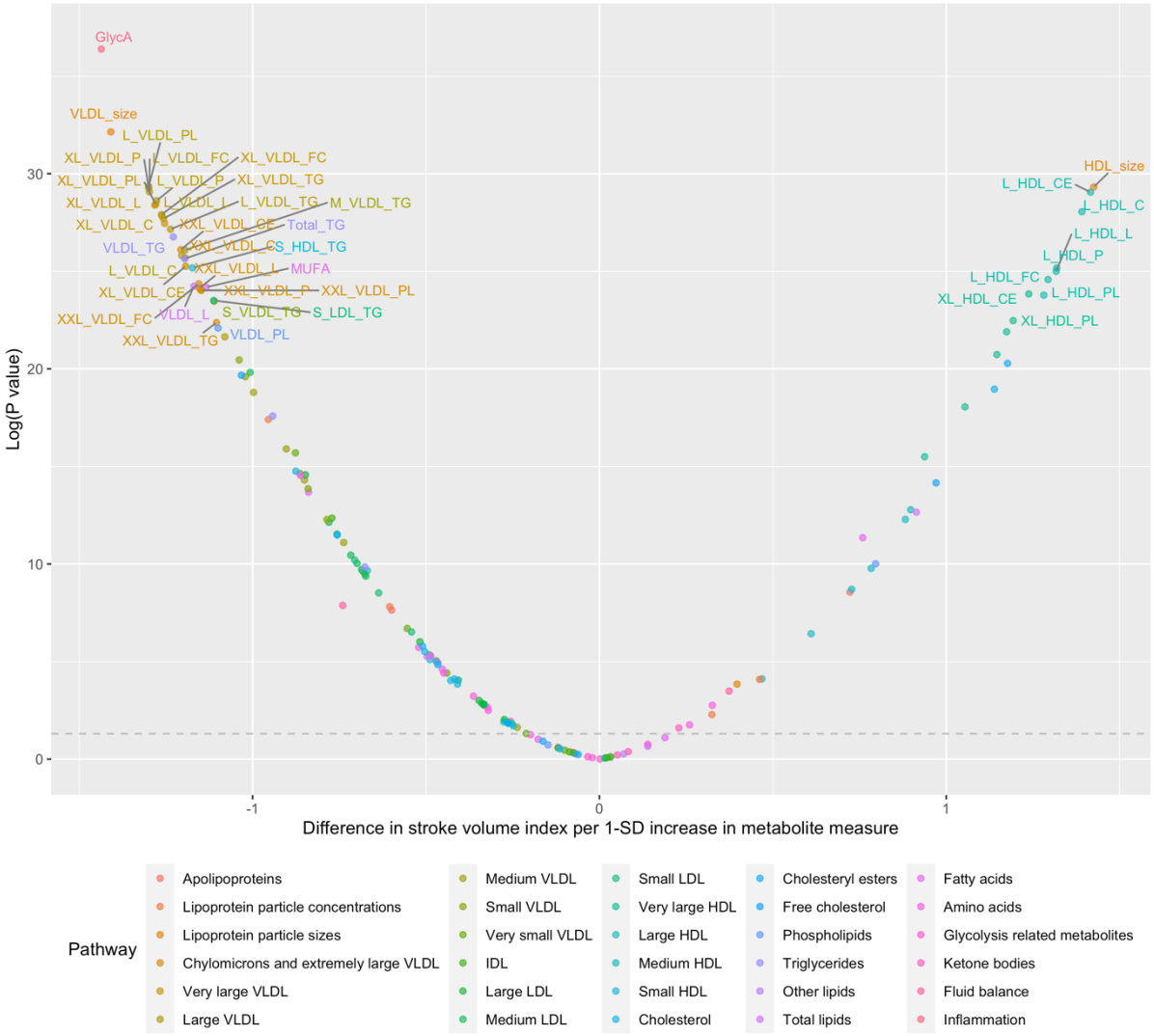
Estimates are derived from a general linear model including one metabolite measure at a time and further adjusted for age, sex, BMI, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, use of anti-diabetic medication, SBP, DBP, use of anti-hypertensive medication, non-HDL-C, HDL-C, TG, use of lipid-lowering medication, and prevalent CAD/AF/VHD
Dashed line represents significance level P=0.05.
Labeled are metabolite measures with strongest positive and negative associations with the endpoint.

Figure 40 presents association between 168 non-ratio metabolite measures and LVEDVI. Sixteen metabolite measures were linked to higher, while 85 with lower LVEDVI. Glycoprotein acetyls were the measure with strongest inverse association with LVEDVI (-1.3 mL/m^2 per 1-SD increase in concentration). Thirty-four of metabolite measures linked LVEDVI belonged to VLDL particles. Among them stronger inverse associations were found for chylomicrons and extremely large, very large and large VLDL size subclasses (β: -2.3 – -

2.0), while weaker for medium, small and very small VLDL (β: -2.1 – -0.9). Moreover, higher concentration of LDL particles, their cholesterol content as well as 13 other measures across LDL size subclasses were associated with a lower LVEDVI. For these measures associations with triglyceride content (β: -1.9 – -1.3) in these lipoprotein subclasses appeared to be stronger than for other measures (β: -1.4 – -1.0). Lower values of LVEDVI were linked higher concentration of isoleucine, leucine, valine and total BCAAs (β: -1.4 – -1.0) as well as creatinine (β: -1.7) was associated with a lower LVEDVI. Among FAs, total, SFA MUFA, and PUFA concentration (including omega-3/6 content) was linked with LVEDVI (β: -2.1 – -1.0). In contrast higher concentration total, esterified and free cholesterol in HDL, as well as twelve different measures belonging to large and very large HDL particles were linked with a higher LVEDVI (all β: ~0.5).

Association of SVI with 168 non-ratio metabolite measures are presented in **Figure 41**. Significant associations with SVI were observed for 148 metabolite measures. Glycoprotein acetyls were the measure with strongest inverse association with SVI (-1.4 mL/m^2 per 1-SD increase in concentration). Among them 49 measures related to whole spectrum of VLDL particles showed an inverse association with this measure (β: -1.3 – -0.2). These associations were more pronounced for larger than for smaller VLDL particles, what was also reflected by inverse trend for average diameter of VLDL (β: -1.4). Furthermore, lower SVI was linked with different measures related to IDL, LDL and small HDL size subclasses as well as overall LDL particle count and cholesterol content. Strong inverse association was identified for total triglycerides concentration. Triglyceride content across VLDL, IDL and LDL classes and subclasses generally tended to have stronger association with SVI then other lipid fractions or particle count of respective lipoprotein subpopulation. In FAs class, all measures with exception for omega-3 FAs and degree of unsaturation were associated with lower SVI (β: -1.1 – -0.4). Also, concentration of isoleucine, leucine, valine, and total BCAAs was associated with lower SVI (β: -0.5 – -0.3). On other hand higher concentration of acetone, glycine, glutamine, and albumin showed a positive association with the cardiac measure (β: 0.2 – 0.4). Moreover, metabolite measures belonging to medium, large, and very large HDL size subclasses (with exception for triglyceride content) were associated with SVI (β: 0.5 – 1.4).

Figure 41. Volcano plot summarizing associations between 168 non-ratio metabolite measures with left ventricle end-systolic volume index.



Estimates are derived from a general linear model including one metabolite measure at a time and further adjusted for age, sex, BMI, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, use of anti-diabetic medication, SBP, DBP, use of anti-hypertensive medication, non-HDL-C, HDL-C, TG, use of lipid-lowering medication, and prevalent CAD/AF/VHD.
Dashed line represents significance level P=0.05.
Labeled are metabolite measures with strongest positive and negative associations with the endpoint.

Multi-metabolite measure scores and automatically annotated MRI parameters of left ventricle

Table 11. Associations between multi-metabolite measure scores and cardiac index in magnetic resonance imaging.

Score	Model 1		Model 2	
	β (95%CI)	P-value	β (95%CI)	P-value
Amino acids-Energy metabolism	0.027 (0.000 - 0.054)	0.047	0.014 (-0.012 - 0.041)	0.297
Ketone bodies	0.003 (-0.023 - 0.029)	0.806	0.000 (-0.026 - 0.026)	0.990
Omega-3 FA	-0.018 (-0.046 - 0.01)	0.208	-0.02 (-0.048 - 0.008)	0.155
Albumin-Amino acids	-0.053 (-0.080 - -0.027)	9.4×10^{-5}	-0.046 (-0.072 - -0.019)	7.4×10^{-4}
Chylomicrons and XXL VLDL	-0.060 (-0.089 - -0.032)	3.9×10^{-5}	-0.065 (-0.094 - -0.037)	7.7×10^{-6}
Cholesterol-enriched Lipoproteins	-0.024 (-0.051 - 0.003)	0.084	-0.030 (-0.057 - -0.002)	0.033
High-density Lipoproteins	-0.066 (-0.094 - -0.038)	4.2×10^{-6}	-0.064 (-0.093 - -0.036)	7.0×10^{-6}
Triglycerides, Fatty Acids and Other Lipids	-0.047 (-0.075 - -0.018)	0.001	-0.049 (-0.078 - -0.021)	6.6×10^{-4}
Triglyceride-rich Lipoproteins	-0.061 (-0.09 - -0.033)	2.7×10^{-5}	-0.063 (-0.091 - -0.034)	1.6×10^{-5}
Triglyceride-rich Lipoproteins Remnants	-0.032 (-0.059 - -0.005)	0.022	-0.036 (-0.063 - -0.009)	0.008
Weighted HFMSS	-0.020 (-0.046 - 0.006)	0.130	-0.006 (-0.032 - 0.019)	0.627
Unweighted HFMSS	0.001 (-0.025 - 0.027)	0.927	0.003 (-0.023 - 0.029)	0.826

Estimates are derived from a general linear model including one cluster-specific score at a time and further adjusted.
Model 1 was adjusted for age, sex, body mass index.
Model 2 included covariates from Model 1 and further self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP and use of anti-hypertensive medication,

Table 11 presents the association between multi-metabolite metabolomic scores and CI. In Model 1, a positive association was observed with “Amino acids-Energy metabolism” (B: 0.027, 95%CI: 0.000-0.054, P=0.047). In contrast, inverse associations were observed for “Albumin-Amino acids” (B: -0.053, 95%CI: -0.080 --0.027, P=9.4×10^{-5}), “Chylomicrons and XXL VLDL” (B: -0.060, 95%CI: -0.089 --0.032, P=3.9×10^{-5}), “High-density Lipoproteins” (B: -0.066, 95%CI: -0.094 - -0.038, P=4.2×10^{-6}), “Triglycerides, Fatty Acids and Other Lipids” (B: -0.047, 95%CI: -0.075 - -0.018, P=0.001), "Triglyceride-rich Lipoproteins" (B: -0.061, 95%CI: -0.09 --0.033, P=2.7×10^{-5}), and "Triglyceride-rich Lipoproteins Remnants" (B: -0.032, 95%CI: -0.059 - -0.005, P=0.022). Following further adjustments in Model 2,

significant associations persisted for all scores with exception for “Amino acids-Energy metabolism”. Moreover, “Cholesterol-enriched lipoproteins” score showed an inverse association which was not previously observed in Model 1(B: -0.030, 95%CI: -0.057 - -0.002, P=0.033).

Table 12. Associations between multi-metabolite measure scores and left ventricle ejection fraction in magnetic resonance imaging.

Score	**Model 1**		**Model 2**	
	β (95%CI)	P-value	β (95%CI)	P-value
Amino acids-Energy metabolism	0.02 (-0.12 - 0.17)	0.778	-0.02 (-0.17 - 0.12)	0.753
Ketone bodies	0.03 (-0.11 - 0.17)	0.704	0.03 (-0.11 - 0.17)	0.68
Omega-3 FA	-0.33 (-0.48 - -0.18)	1.5×10^{-5}	-0.34 (-0.49 - -0.18)	1.5×10^{-5}
Albumin-Amino acids	-0.21 (-0.35 - -0.06)	0.005	-0.20 (-0.34 - -0.05)	0.008
Chylomicrons and XXL VLDL	0.02 (-0.14 - 0.18)	0.799	-0.03 (-0.19 - 0.12)	0.674
Cholesterol-enriched Lipoproteins	0.25 (0.10 - 0.39)	0.001	0.23 (0.08 - 0.37)	0.003
High-density Lipoproteins	-0.17 (-0.32 - -0.02)	0.028	-0.19 (-0.35 - -0.04)	0.014
Triglycerides, Fatty Acids and Other Lipids	-0.20 (-0.35 - -0.04)	0.013	-0.23 (-0.38 - -0.07)	0.004
XL VLDL cluster	-0.02 (-0.17 - 0.14)	0.851	-0.05 (-0.20 - 0.11)	0.555
Other VLDL cluster	-0.11 (-0.26 - 0.04)	0.136	-0.13 (-0.28 - 0.01)	0.073
Weighted HFMSS	-0.25 (-0.39 - -0.11)	4.7×10^{-4}	-0.22 (-0.36 - -0.08)	0.002
Unweighted HFMSS	-0.09 (-0.23 - 0.05)	0.203	-0.08 (-0.22 - 0.06)	0.275

Estimates are derived from a general linear model including one cluster-specific score at a time and further adjusted.
Model 1 was adjusted for age, sex, body mass index.
Model 2 included covariates from Model 1 and further self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP and use of anti-hypertensive medication.

Table 12 illustrates the association between specific metabolomic scores and LVEF in participants with prevalent HF. In Model 1, an inverse association was observed for "Omega-3 FA" (B: -0.33, 95%CI: -0.48 to -0.18, P=1.5 x 10^{-5}), "Albumin-Amino acids" (B: -0.21, 95%CI: -0.35 to -0.06, P=0.005), "High-density Lipoproteins" (B: -0.17, 95%CI: -0.32 to -0.02, P=0.028), "Triglycerides, Fatty Acids and Other Lipids" (B: -0.20, 95%CI: -0.35 to -0.04, P=0.013), and "Weighted HFMSS" (B: -0.25, 95%CI: -0.39 to -0.11, P=4.7 x 10^-4). Conversely, the "Cholesterol-enriched Lipoproteins" score exhibited a positive association with LVEF (B: 0.25, 95%CI: 0.10 to 0.39, P=0.001). Upon adjustments in Model 2, significant inverse associations persisted for "Omega-3 FA", "Albumin-Amino acids", "High-density

Lipoproteins", "Triglycerides, Fatty Acids and Other Lipids", and "Weighted HFMSS", while the positive association remained for the "Cholesterol-enriched Lipoproteins" score.

Table 13. Associations between multi-metabolite measure scores and left ventricle end-systolic volume index in magnetic resonance imaging.

Score	Model 1		Model 2	
	β (95%CI)	P-value	β (95%CI)	P-value
Amino acids-Energy metabolism	-0.2 (-0.8 - 0.5)	0.611	-0.2 (-0.8 - 0.4)	0.546
Ketone bodies	-0.2 (-0.8 - 0.5)	0.616	-0.1 (-0.7 - 0.5)	0.725
Omega-3 FA	0.1 (-0.6 - 0.7)	0.872	0.1 (-0.6 - 0.8)	0.786
Albumin-Amino acids	-0.2 (-0.8 - 0.4)	0.527	-0.2 (-0.8 - 0.5)	0.575
Chylomicrons and XXL VLDL	-1.2 (-1.9 - -0.5)	4.2×10^{-4}	-1.1 (-1.8 - -0.4)	0.002
Cholesterol-enriched Lipoproteins	-0.4 (-1.0 - 0.3)	0.283	-0.2 (-0.8 - 0.5)	0.558
High-density Lipoproteins	-0.3 (-1.0 - 0.3)	0.332	-0.3 (-0.9 - 0.4)	0.436
Triglycerides, Fatty Acids and Other Lipids	-0.6 (-1.3 - 0.0)	0.064	-0.5 (-1.2 - 0.2)	0.135
XL VLDL cluster	-0.8 (-1.5 - -0.2)	0.016	-0.7 (-1.4 - 0.0)	0.051
Other VLDL cluster	-0.1 (-0.8 - 0.5)	0.665	-0.1 (-0.7 - 0.6)	0.823
Weighted HFMSS	0.1 (-0.5 - 0.7)	0.817	0.1 (-0.5 - 0.8)	0.675
Unweighted HFMSS	-0.2 (-0.8 - 0.5)	0.612	-0.1 (-0.7 - 0.5)	0.795

Estimates are derived from a general linear model including one cluster-specific score at a time and further adjusted.
Model 1 was adjusted for age, sex, body mass index.
Model 2 included covariates from Model 1 and further self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP and use of anti-hypertensive medication.

Table 13 presents the relationship between specific metabolomic scores and LVESVI among participants with prevalent HF. In Model 1, a pronounced inverse association was noted for "Chylomicrons and XXL VLDL" (B: -1.2, 95%CI: -1.9 to -0.5, P=0.0004) and "XL VLDL" scores (B: -0.8, 95%CI: -1.5 to -0.2, P=0.016). Upon adjustments in Model 2, the inverse association for "Chylomicrons and XXL VLDL" persisted (B: -1.1, 95%CI: -1.8 to -0.4, P=0.002), but the association for "XL VLDL" was attenuated (P=0.051).

Table 14. Associations between multi-metabolite measure scores and left ventricle end-diastolic volume index in magnetic resonance imaging.

Score	Model 1		Model 2	
	β (95%CI)	P-value	β (95%CI)	P-value
Amino acids-Energy metabolism	-0.2 (-1.0 - 0.5)	0.514	-0.4 (-1.1 - 0.4)	0.312
Ketone bodies	-0.3 (-1.0 - 0.5)	0.462	-0.2 (-0.9 - 0.5)	0.617
Omega-3 FA	-0.7 (-1.5 - 0.0)	0.065	-0.7 (-1.4 - 0.1)	0.098
Albumin-Amino acids	-0.7 (-1.4 - 0.1)	0.072	-0.6 (-1.4 - 0.1)	0.105
Chylomicrons and XXL VLDL	-2.5 (-3.3 - -1.7)	1.1×10^{-9}	-2.3 (-3.1 - -1.5)	3.2×10^{-8}
Cholesterol-enriched Lipoproteins	-0.9 (-1.6 - -0.1)	0.023	-0.6 (-1.4 - 0.1)	0.098
High-density Lipoproteins	-1.4 (-2.2 - -0.6)	6.0×10^{-4}	-1.3 (-2.1 - -0.5)	0.002
Triglycerides, Fatty Acids and Other Lipids	-1.8 (-2.6 - -1.1)	4.7×10^{-6}	-1.7 (-2.5 - -0.9)	3.3×10^{-5}
XL VLDL cluster	-2.0 (-2.8 - -1.2)	1.4×10^{-6}	-1.7 (-2.5 - -0.9)	2.5×10^{-5}
Other VLDL cluster	-0.8 (-1.6 - -0.1)	0.028	-0.7 (-1.5 - 0.0)	0.059
Weighted HFMSS	-0.3 (-1.0 - 0.4)	0.441	-0.1 (-0.9 - 0.6)	0.724
Unweighted HFMSS	-0.6 (-1.3 - 0.1)	0.099	-0.4 (-1.2 - 0.3)	0.240

Estimates are derived from a general linear model including one cluster-specific score at a time and further adjusted.
Model 1 was adjusted for age, sex, body mass index.
Model 2 included covariates from Model 1 and further self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP and use of anti-hypertensive medication.

Associations between multi-metabolite measure scores and LVEDVI is presented in **Table 14**. In Model 1, all scores except for "Amino acids-Energy metabolism" and "Ketone bodies" cluster-specific scores were associated with lower value of this measure (B: -2.5. The strongest inverse associations were observed for "Chylomicrons and XXL VLDL" (B: -2.5, 95%CI: -3.3 to -1.7, P=1.1×10^{-9}), "XL VLDL cluster" (B: -2.0, 95%CI: -2.8 to -1.2, P=1.4×10^{-6}), "Triglycerides, Fatty Acids and Other Lipids" (B: -1.8, 95%CI: -2.6 to -1.1, P=4.7×10^{-6}). Following adjustment for covariates in Model 2, the associations for "Cholesterol-enriched Lipoproteins" (P=0.098) and "Other VLDL cluster" (P=0.059) scores were no longer significant.

Table 15. Associations between multi-metabolite measure scores and stroke volume index in magnetic resonance imaging.

Score	Model 1		Model 2	
	β (95%CI)	P-value	β (95%CI)	P-value
Amino acids-Energy metabolism	-0.1 (-0.3 - 0.1)	0.437	0.2 (-0.4 - 0.0)	0.072
Ketone bodies	-0.1 (-0.3 - 0.1)	0.256	0.1 (-0.3 - 0.1)	0.448
Omega-3 FA	-0.8 (-1.0 - -0.6)	8.6×10^{-13}	-0.8 (-1.0 - -0.5)	4.2×10^{-12}
Albumin-Amino acids	-0.5 (-0.7 - -0.3)	5.7×10^{-6}	-0.4 (-0.6 - -0.2)	2.8×10^{-5}
Chylomicrons and XXL VLDL	-1.3 (-1.5 - -1.0)	1.0×10^{-28}	-1.2 (-1.4 - -1.0)	1.3×10^{-26}
Cholesterol-enriched Lipoproteins	-0.5 (-0.7 - -0.3)	1.1×10^{-6}	-0.4 (-0.7 - -0.2)	2.7×10^{-5}
High-density Lipoproteins	-1.0 (-1.3 - -0.8)	7.6×10^{-21}	-1.0 (-1.2 - -0.8)	6.1×10^{-20}
Triglycerides, Fatty Acids and Other Lipids	-1.2 (-1.4 - -1.0)	3.7×10^{-27}	-1.2 (-1.4 - -1.0)	5.5×10^{-26}
XL VLDL cluster	-1.1 (-1.4 - -0.9)	1.1×10^{-23}	-1.0 (-1.3 - -0.8)	8.1×10^{-21}
Other VLDL cluster	-0.7 (-0.9 - -0.5)	5.4×10^{-11}	-0.7 (-0.9 - -0.4)	5.4×10^{-10}
Weighted HFMSS	-0.4 (-0.6 - -0.2)	4.4×10^{-4}	-0.3 (-0.5 - -0.1)	0.008
Unweighted HFMSS	-0.5 (-0.7 - -0.3)	1.1×10^{-5}	-0.4 (-0.6 - -0.2)	4.4×10^{-4}

Estimates are derived from a general linear model including one cluster-specific score at a time and further adjusted.
Model 1 was adjusted for age, sex, body mass index.
Model 2 included covariates from Model 1 and further self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, SBP, DBP and use of anti-hypertensive medication.

Table 15 presents associations between multi-metabolite measure scores and SVI. In Model 1, all scores with exception for "Amino acids-Energy metabolism" and "Ketone bodies" cluster-specific scores were associated with lower SVI. The strongest magnitude of associations was observed for "Chylomicrons and XXL VLDL" (B: -1.3, 95%CI: -1.5 to -1.0, P=1.0×10^{-28}), "Triglycerides, Fatty Acids and Other Lipids" (B: -1.2, 95%CI: -1.4 to -0.9, P=3.7×10^{-27}), "XL VLDL cluster" (B: -1.1, 95%CI: -1.4 to -0.9, P=1.1×10^{-23}) and "High-density Lipoproteins" (B: -1.0, 95%CI: -1.3 to -0.8, P=7.6×10^{-21}). Following adjustment in Model 2, all associations from Model 1 remained statistically significant.

Metabolite markers in relation to survival among heart failure patients

In the analysis of metabolite measures associated with survival among patients with HF, a subcohort of 690 UK Biobank participants who had been diagnosed with HF prior to their enrollment in the cohort was included. Their baseline characteristics is presented in **Table 16**. The median time from HF diagnosis to recruitment was 4.2 years (IQR: 1.8 – 7.3). The average age of participants in the subcohort was 61 years. Of these participants, 27% were female, and 94% reported their ethnicity as White. With regards to medication, 16% of the subcohort used anti-diabetic drugs, 89% used anti-hypertensive drugs, and 69% used lipid-lowering medication. Approximately two-thirds (62%) of the participants had a history of CAD while one-third (31%) had a history of AF. Moreover, 18% of the participants with prevalent HF had been previously diagnosed with VHD.

Compared with the overall subcohort, participants who died during follow-up were older, more likely to be male, had a higher BMI, were more likely to be current smokers, were more likely to use anti-hypertensive, anti-diabetic, and lipid-lowering medication, were more likely to have had prevalent events predisposing them to HF, and had a shorter time since HF diagnosis.

Table 16. Comparison of participants' baseline characteristics between overall subcohort of prevalent HF and individuals who died during follow-up.

Characteristic	Subcohort with HF diagnosis before recruitment	All-cause deaths during follow-up	Cardiovascular deaths during follow-up
n	690	238	119
Age, years	61 (7)	63 (6)	62 (6)
Female sex	186 (27%)	51 (21%)	20 (17%)
Body mass index, kg/m2	30.2 (5.6)	30.9 (5.9)	31.0 (5.9)
White ethnicity	647 (94%)	220 (92%)	108 (91%)
Townsend Deprivation Index	-0.4 (3.5)	0.0 (3.7)	0.1 (3.6)
High physical activity status	166 (31%)	47 (27%)	20 (23%)
Current smokers	82 (12%)	39 (16%)	19 (16%)
Fasting, hrs	4 (3, 5)	4 (3, 5)	4 (3, 5)
eGFR, mL/min/1.73 m^2	75 (64, 90)	70 (53, 85)	70 (52, 78)
Alcohol intake, g/day	8 (0, 28)	7 (0, 28)	7 (0, 32)
HbA1c, mmol/mol	38 (35, 42)	41 (37, 49)	40 (37, 49)
Use of anti-diabetic medication	111 (16%)	66 (28%)	31 (26%)

Systolic blood pressure, mmHg	132 (120, 146)	132 (121, 146)	132 (119, 145)
Diastolic blood pressure, mmHg	78 (70, 84)	76 (69, 83)	74 (67, 83)
Use of anti-hypertensive medication	614 (89%)	220 (92%)	110 (92%)
Use of lipid-lowering medication	477 (69%)	175 (74%)	86 (72%)
non-HDL-C, mmol/L	2.51 (2.06, 3.06)	2.35 (1.95, 2.81)	2.35 (1.89, 2.83)
HDL-C, mmol/L	1.05 (0.90, 1.22)	0.98 (0.85, 1.16)	0.98 (0.85, 1.14)
Triglycerides, mmol/L	1.23 (0.92, 1.64)	1.18 (0.92, 1.62)	1.20 (0.93, 1.63)
Prevalent CAD	421 (61%)	176 (74%)	95 (80%)
Prevalent AF	219 (32%)	99 (42%)	48 (40%)
Prevalent VHD	127 (18%)	48 (20%)	27 (23%)
Time since HF diagnosis, years	4.2 (1.8, 7.3)	3.8 (1.9, 7.3)	4.4 (2.4, 7.3)

Values are mean (standard deviation) or median (interquartile range) for continuous variables, and number of participants (%) for categorical variables.
Abbreviations: AF - atrial fibrillation, CAD - coronary artery disease, eGFR - estimated glomerular filtration rate, HbA1c - hemoglobin A1c, HDL-C - high-density lipoprotein cholesterol, HF - heart failure, non-HDL-C - non-high-density lipoprotein cholesterol, VHD - valvular heart disease.

During follow-up, 238 participants died from any cause, including 119 participants who died from a cardiovascular cause. The median time since recruitment to death from any cause was 11.4 years (IQR: 9.1 – 12.9) and 11.8 years (IQR: 11.0 - 12.7) for cardiovascular deaths. The cumulative incidence of deaths from any cause and from cardiovascular causes is presented in **Figures 42** and **43**.

Figure 42. Kaplan-Meier curve of the cumulative all-cause mortality among participants with heart failure at baseline.

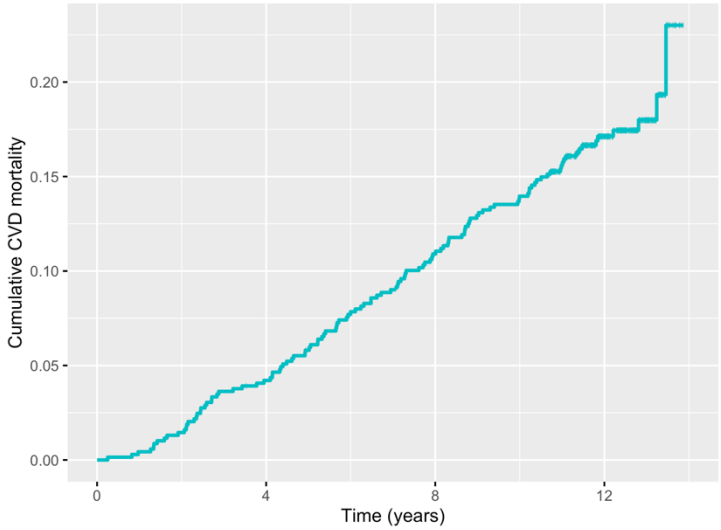
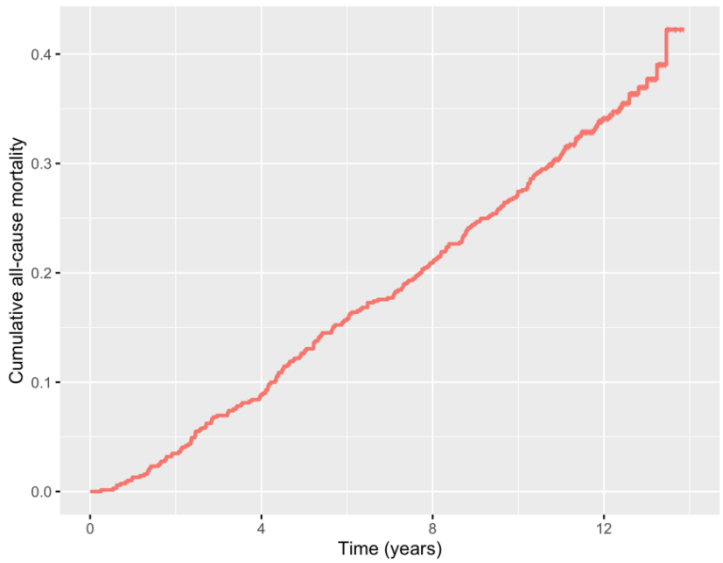


Figure 43. Kaplan-Meier curve of the cumulative cardiovascular mortality among participants with heart failure at baseline.

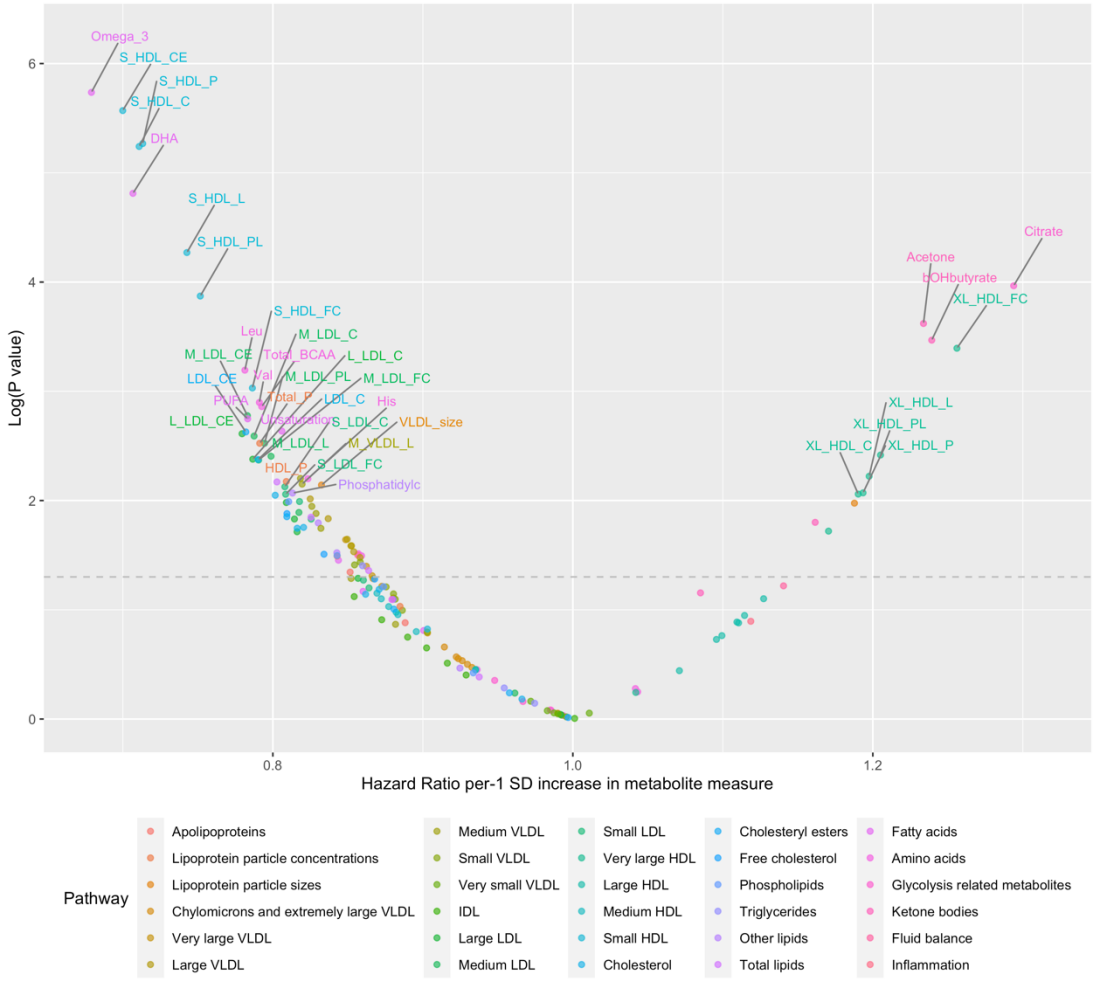


Metabolome-wide analysis for all-cause and cardiovascular mortality in heart failure patients

Associations between 168 non-ratio metabolite measures and all-cause mortality in participants with baseline HF is presented in **Figure 44**. 83 of 168 metabolite measures were significantly associated with the endpoint. Inverse associations with risk were found for 72 metabolite measures with largest magnitude of association found for PUFA, small HDL particles, BCAAs and LDL particles classes. 1-SD increase in concertation of omega-3 FAs, DHA and overall PUFA was respectively associated with 32%, 29% and 22% lower risk of all-cause death. Higher concentration of HDL particles or their lipid content (total lipids, total cholesterol, cholesteryl esters, free cholesterol and phospholipids) were associated with 22-31% lower all-cause mortality. Different cholesterol fractions (total cholesterol, cholesteryl esters and free cholesterol) in LDL size subclasses as well as total LDL cholesterol were linked to 18-22% lower risk of all-cause death. Total concentration of BCAA (HR: 0.79, 95%CI: 0.65-0.94, P=0.001) as well as individually isoleucine (HR: 0.86, 95%CI: 0.72-1.00, P=0.032), leucine (HR: 0.78, 95%CI: 0.64-0.92, P=0.001), and valine (HR: 0.79, 95%CI: 0.65-0.93, P=0.001) were associated with lower all-cause mortality. Inverse associations were further observed for VLDL size (HR: 0.83, 95%CI: 0.7-0.97, P=0.007) as well as metabolite measures related to medium (17-18% lower risk), large (14-18% lower risk) and very large VLDL (13-15% lower risk). Moreover, inverse associations with all-cause mortality were found for albumin, alanine and histidine.

Eleven metabolite measures were significantly associated with higher risk of all-cause mortality. The strongest association was found for citrate. Higher concentration of ketone bodies, acetone (HR: 1.24, 95%CI: 1.12-1.36, P=2.4 x 10^{-4}), acetoacetate (HR: 1.16, 95%CI: 1.04-1.28, P=0.016), and 3-OHB (HR: 1.24, 95%CI: 1.12-1.36, P=3.4 x 10^{-4}) were associated with higher risk of all-cause death. Higher concentration of XL HDL particles and their lipid content (total lipids, total cholesterol, cholesteryl esters, free cholesterol, and phospholipids) were linked with 17-26% higher all-cause mortality. Moreover, a higher risk of all-cause death was associated with larger HDL particle diameter (HR: 1.19, 95%CI: 1.06-1.32, P=0.011).

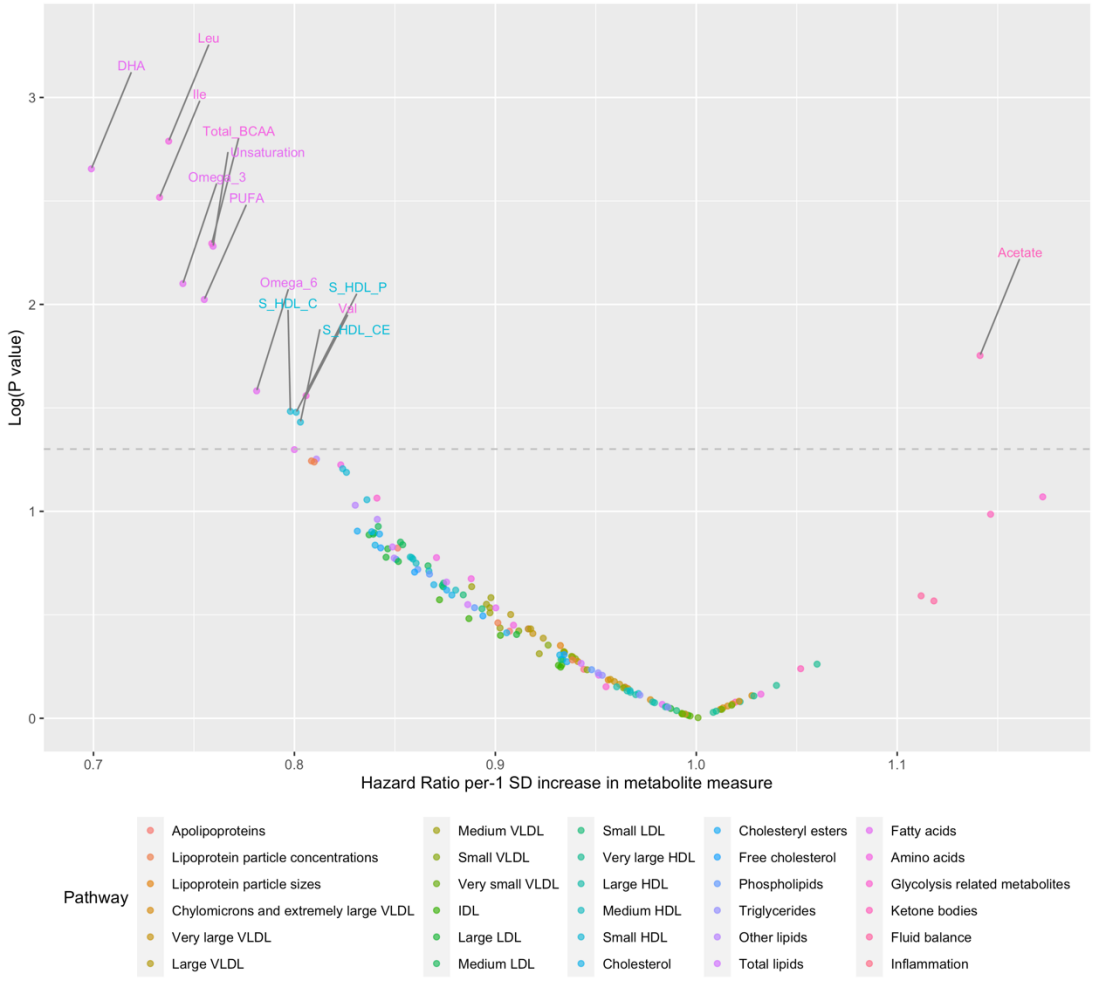
Figure 44. Volcano plot summarizing associations between non-ratio 168 metabolite measures with all-cause mortality among participants with prevalent HF.



Estimates are derived from a Cox proportional hazards model including one metabolite measure at a time and further adjusted for age, sex, BMI, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, use of anti-diabetic medication, SBP, DBP, use of anti-hypertensive medication, non-HDL-C, HDL-C, TG, use of lipid-lowering medication, and prevalent CAD/AF/VHD
Dashed line represents significance level P=0.05.
Labeled are metabolite measures with strongest positive and negative associations with the endpoint.

Figure 45 summarizes associations between 168 non-ratio metabolite measures and cardiovascular mortality in participants with baseline HF. Higher concentrations of DHA, total omega-3 FAs, total omega-6 FAs and total PUFA were associated with 16-32% lower cardiovascular mortality. Similarly, isoleucine, leucine, valine, and total concentration of BCAAs were linked to 16-22% lower risk of cardiovascular death. Moreover, inverse association with cardiovascular mortality was observed for concentration of small HDL particles and their total and esterified cholesterol content (23-31% lower risk). Higher risk of cardiovascular death was found only for higher concentration of acetate (HR: 1.09, 95%CI: 1.02-1.25, P=0.02).

Figure 45. Volcano plot summarizing associations between non-ratio 168 metabolite measures with cardiovascular mortality among participants with prevalent HF.



Estimates are derived from a Cox proportional hazards model including one metabolite measure at a time and further adjusted for age, sex, BMI, self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, use of anti-diabetic medication, SBP, DBP, use of anti-hypertensive medication, non-HDL-C, HDL-C, TG, use of lipid-lowering medication, and prevalent CAD/AF/VHD
Dashed line represents significance level P=0.05.
Labeled are metabolite measures with strongest positive and negative associations with the endpoint

Multi-metabolite measure scores for all-cause and cardiovascular mortality in heart failure patients

Associations between cluster-specific scores previously derived for incident HF risk and all-cause mortality in participants with prevalent HF are presented in **Table 17**. In Model 1 adjusted for age, sex and BMI, higher "Ketone bodies" (HR: 1.19, 95%CI: 1.06-1.33, P=0.002), "Omega-3 Fatty Acids" (HR: 1.30, 95%CI: 1.14-1.48, P=9.2 x 10^{-5}), "Albumin-Amino acids" (HR: 1.16, 95%CI: 1.02-1.33, P=0.028), "Triglycerides, Fatty Acids and Other Lipids" (HR: 1.80, 95%CI: 1.58-2.04, P=2.1 x 10^{-19}), "Other VLDL" scores (HR: 1.31, 95%CI: 1.16-1.48, P=1.2 x 10^{-5}), weighted (HR: 1.42, 95%CI: 1.27-1.59, P=5.7 x 10^{-10}) and unweighted HFMSS (HR: 1.26, 95%CI: 1.12-1.41, P=7.2 x 10^{-5}) were associated with higher risk of all-cause death, while "Cholesterol-enriched Lipoproteins" with lower risk of all-cause death. Following further adjustment, the positive association all-cause mortality was no longer significant for "Albumin -Amino acids" (P=0.099) and "Cholesterol-enriched Lipoproteins" scores (P=0.158). In Model 2, the strongest associations with all-cause mortality were observed for "Triglycerides, Fatty Acids and Other Lipids" (HR: 1.54, 95%CI: 1.34-1.78, P=1.4 x 10^{-9}) and weighted HFMSS (HR: 1.39, 95%CI: 1.23-1.57, P=1.9 x 10^{-7}).

Table 17. Associations between multi-metabolite measure scores and all-cause mortality in participants with prevalent heart failure.

Score	Model 1		Model 2	
	HR (95%CI)	P-value	HR (95%CI)	P-value
Amino acids-Energy metabolism	1.11 (0.98 - 1.26)	0.114	1.06 (0.93 - 1.22)	0.377
Ketone bodies	1.19 (1.06 - 1.33)	0.002	1.16 (1.03 - 1.31)	0.015
Omega-3 FA	1.30 (1.14 - 1.48)	9.2 x 10^{-5}	1.24 (1.08 - 1.43)	0.002
Albumin-Amino acids	1.16 (1.02 - 1.33)	0.028	1.13 (0.98 - 1.30)	0.099
Chylomicrons and XXL VLDL	0.98 (0.86 - 1.12)	0.764	0.90 (0.78 - 1.04)	0.168
Cholesterol-enriched Lipoproteins	0.87 (0.76 – 1.00)	0.046	0.89 (0.77 - 1.04)	0.158
High-density Lipoproteins	1.15 (1.00 - 1.32)	0.055	1.11 (0.96 - 1.29)	0.15
Triglycerides, Fatty Acids and Other Lipids	1.80 (1.58 - 2.04)	2.1 x 10^{-19}	1.54 (1.34 - 1.78)	1.4 x 10^{-9}
XL VLDL	1.06 (0.93 - 1.21)	0.376	1.02 (0.89 - 1.17)	0.777
Other VLDL	1.31 (1.16 - 1.48)	1.2 x 10^{-5}	1.22 (1.06 - 1.40)	0.004
Weighted HFMSS	1.42 (1.27 - 1.59)	5.7 x 10^{-10}	1.39 (1.23 - 1.57)	1.9 x 10^{-7}
Unweighted HFMSS	1.26 (1.12 - 1.41)	7.2 x 10^{-5}	1.29 (1.14 - 1.46)	7.1 x 10^{-5}

Estimates are derived from a Cox proportional hazards model including one cluster-specific score at a time and further adjusted.
Model 1 was adjusted for age, sex, body mass index.

Model 2 included covariates from Model 1 and further self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, use of anti-diabetic medication, SBP, DBP, use of anti-hypertensive medication, non-HDL-C, HDL-C, TG, use of lipid-lowering medication, and prevalent CAD/AF/VHD

Table 18 summarizes association between cluster-specific scores and cardiovascular mortality in participants with prevalent HF. Higher values of "Omega-3 Fatty acids" (HR: 1.34, 95%CI: 1.11-1.61, P=0.002), "Triglycerides, Fatty Acids and Other Lipids" (HR: 1.49, 95%CI: 1.26-1.77, P=3.3 x 10^{-6}), "Other VLDL" (HR: 1.18, 95%CI: 1.01-1.37, P=0.033) scores as well as weighted (HR: 1.27, 95%CI: 1.08-1.49, P=0.004) and unweighted HFMSS (HR: 1.24, 95%CI: 1.06-1.46, P=0.008) were associated with higher cardiovascular mortality. Following further adjustment, the significant associations with cardiovascular mortality remained only for "Omega-3 Fatty acids" (HR: 1.32, 95%CI: 1.09-1.60, P=0.005) and "Triglycerides, Fatty Acids and Other Lipids" (HR: 1.27, 95%CI: 1.04-1.53, P=0.016) scores.

Table 18. Associations between multi-metabolite measure scores and cardiovascular mortality in participants with prevalent heart failure.

Score	Model 1		Model 2	
	HR (95%CI)	P-value	HR (95%CI)	P-value
Amino acids-Energy metabolism	1.04 (0.88 - 1.24)	0.620	0.99 (0.82 - 1.21)	0.937
Ketone bodies	1.13 (0.97 - 1.33)	0.124	1.02 (0.85 - 1.22)	0.837
Omega-3 FA	1.34 (1.11 - 1.61)	0.002	1.32 (1.09 - 1.60)	0.005
Albumin-Amino acids	0.98 (0.82 - 1.19)	0.864	0.94 (0.77 - 1.14)	0.514
Chylomicrons and XXL VLDL	1.03 (0.86 - 1.24)	0.758	0.98 (0.81 - 1.18)	0.822
Cholesterol-enriched Lipoproteins	0.90 (0.74 - 1.09)	0.280	0.92 (0.73 - 1.14)	0.434
High-density Lipoproteins	1.09 (0.90 - 1.32)	0.377	1.12 (0.91 - 1.37)	0.297
Triglycerides, Fatty Acids and Other Lipids	1.49 (1.26 - 1.77)	3.3 x 10^{-6}	1.27 (1.04 - 1.53)	0.016
XL VLDL	1.13 (0.95 - 1.34)	0.108	1.11 (0.92 - 1.34)	0.274
Other VLDL	1.18 (1.01 - 1.37)	0.033	1.15 (0.95 - 1.37)	0.144
Weighted HFMSS	1.27 (1.08 - 1.49)	0.004	1.15 (0.96 - 1.36)	0.121
Unweighted HFMSS	1.24 (1.06 - 1.46)	0.008	1.18 (0.99 - 1.41)	0.063

Estimates are derived from a Cox proportional hazards model including one cluster-specific score at a time and further adjusted.
Model 1 was adjusted for age, sex, body mass index.
Model 2 included covariates from Model 1 and further self-reported ethnicity, Townsend Deprivation Index, fasting time, smoking status, alcohol intake, eGFR, HbA1c, use of anti-diabetic medication, SBP, DBP, use of anti-hypertensive medication, non-HDL-C, HDL-C, TG, use of lipid-lowering medication, and prevalent CAD/AF/VHD

5. Discussion

Summary of study findings

This study sought to identify metabolite signatures of incident HF in a general UK population comprising middle-aged to elderly adults. When analyzing each metabolite measure individually, 90 metabolite measures were associated with incident HF in the fully adjusted model and after correction for multiple testing. Among metabolite measures associated with higher HF risk there were measures related to, ketone bodies (3-OHB, acetate, acetoacetate, acetone), particle diameter (average diameter for HDL particles), very large HDL (concentration of particles, phospholipids, cholesteryl esters), large HDL (concentration of particles, free cholesterol) and inflammation (glycoprotein acetyls). Metabolite measures linked to a lower HF risk covered lipoprotein particle sizes (average diameter for VLDL particles), FAs (degree of unsaturation, omega-3 FAs, DHA, PUFA), amino acids (alanine, histidine, valine), ketone bodies (3-OHB, acetate, acetoacetate, acetone), fluid balance (albumin), very large VLDL (concentration particles, triglycerides, total lipids), large VLDL (concentration particles, phospholipids, free cholesterol), medium VLDL (triglycerides, total lipids), medium LDL (phospholipids, cholesteryl esters, free cholesterol), small LDL (cholesteryl esters, cholesterol in), small HDL (concentration of particles, phospholipids, cholesteryl esters), cholesterol (LDL cholesterol, cholesterol in medium LDL), lipoprotein particle concentrations (concentration of HDL particles, total concentration of lipoprotein particles), triglycerides (triglycerides in VLDL), cholesteryl esters (cholesteryl esters in LDL).

Subgroup, sensitivity, and mediation analyses revealed that of 90, 23 specific metabolite measures had an independent association with incident HF. These metabolite measures encompassed amino acids (alanine, histidine), ketone bodies (acetate, acetoacetate, acetone, 3-OHB), fluid balance markers (albumin), and FAs (DHA, omega-3 FAs), along with various subtypes of HDL and LDL lipoproteins. Although metabolite measure ratios were initially considered, they were later excluded due to information redundancy. An HFMSS was subsequently developed, drawing on 10 metabolite measures. These included amino acids (alanine, histidine), ketone bodies (acetate, acetoacetate, acetone, 3-OHB), albumin, concentrations of small and large HDL particles, and DHA. Following full covariate adjustment, a 10-point increase in the HFMSS was associated with a 51% (95%CI: 34-70%) and 69% (95%CI: 52-88%) higher HF risk, depending on the score version. When comparing the top to bottom quartiles across both versions of the HFMSS, the HF risk was 88% (95%CI:

61-120%) and 112% (95%CI: 82-147%) higher, respectively. Therefore, the HFMSS serves as a robust metric for assessing joint metabolic risk for incident HF.

Using GGMs and the walk-trap algorithm, we identified ten distinct clusters based on metabolite measures: “Amino Acid-Energy Metabolism” (four metabolite measures positively correlated with tyrosine, glucose, and lactate, and negatively correlated with alanine), “Ketone Bodies” (a single metabolite measure, acetoacetate), “Omega-3 Fatty Acids” (a single metabolite measure indicating degree of unsaturation), “Chylomicrons and Extremely Large VLDL” (a single metabolite measure, cholesteryl esters), “Cholesterol-enriched Lipoproteins’ (14 metabolite measures showing negative correlation with some LDL and VLDL measures but positive with others), “Albumin-Amino Acids” (three metabolite measures, all negatively correlated with various amino acids), “High-density Lipoproteins” (12 metabolite measures positively correlated with specific HDL fractions and negatively with HDL diameter), “Triglycerides, Fatty Acids and Other Lipids’ (16 metabolite measures positively correlated with specific FAs and triglycerides, and negatively with others), “Triglyceride-rich Lipoproteins” (five metabolite measures positively correlated with VLDL particle content but negatively with certain VLDL triglycerides), and “Triglyceride-rich Lipoprotein Remnants” (11 metabolite measures positively correlated with specific VLDL measures and negatively with medium VLDL metrics). The strongest risk associations were for “Triglycerides, Fatty Acids and Other Lipids” (29% increase in risk), “Albumin-Amino Acids” (19% increase in risk), and “Triglyceride-rich Lipoproteins Remnants” (15% increase in risk). A one SD increase in scores for “Amino acids-Energy metabolism”, “Ketone bodies”, “Omega-3 Fatty Acids”, and “High-density Lipoproteins” corresponded to a 7-9% increase in HF risk. The inclusion of these metabolite measure-based cluster scores and HFMSS into conventional risk factors improved prediction of incident HF.

Further, as a secondary aim, this study evaluated the associations between metabolite measures and LV function in MRI. Higher CI was positively associated with 34 metabolites including glycolysis-related compounds like glucose, lactate, and pyruvate, as well as larger average HDL particles and their lipid contents. Conversely, 83 metabolites were negatively associated with higher CI. Notably, higher concentrations of creatinine, BCAAs, and lipid fractions in VLDL, IDL, and LDL particles were associated with lower CI. Regarding LVEF, the study found that 82 metabolite measures had a significant association with LVEF. Eighty of these were positively associated, while only one, phenylalanine, was inversely associated. The

strongest positive association was with circulating albumin. Other positively associated measures included PUFA such as degree of unsaturation and total concentration of omega-3/6 FAs, as well as various lipids like cholines, SMs, phosphoglycerides, and PCs. In addition, 31 metabolite measures related to IDL and LDL particle concentration and lipid content were positively associated with higher LVEF. ApoA1 and 24 measures related to large, medium, and small HDL particles also showed positive associations with higher LVEF. For LVESVI, 47 metabolites were linked with lower systolic function, with the strongest inverse association being with glycoprotein acetyls. A majority of the 27 metabolites that had an inverse association belonged to VLDL particles, specifically chylomicrons and larger VLDL subclasses. Smaller HDL particles, BCAAs, and creatinine were also inversely related to LVESVI. For LVEDVI, 16 metabolites were associated with higher, while 85 were linked to lower values. Glycoprotein acetyls again showed the strongest inverse association. Stronger negative associations were identified in 34 metabolites belonging to chylomicrons and larger VLDL subclasses, while weaker associations were found for smaller VLDL particles. Regarding SVI, the study found significant associations with 148 out of 168 non-ratio metabolite measures. The strongest inverse association was observed with glycoprotein acetyls. A total of 49 measures related to VLDL particles exhibited negative associations, with larger VLDL particles showing more pronounced effects than smaller ones. This was supported by the inverse trend noted for the average diameter of VLDL. Lower SVI was also associated with various measures related to IDL, LDL, and small HDL size subclasses, as well as total triglyceride concentration.

Lastly, this study aimed to evaluate the associations between metabolite measures and survival in individuals with prevalent HF. Increased concentrations of certain metabolites like PUFA, small HDL particles, BCAAs, and LDL particles were associated with lower mortality risks ranging from 18-32%. On the other hand, eleven metabolite measures including citrate and higher concentrations of ketone bodies were associated with higher risks of all-cause mortality, ranging from 17-26%. Specifically, higher levels of XL HDL particles and their lipid content, as well as a larger HDL particle diameter, were linked to increased all-cause mortality risk. Regarding CVD mortality in individuals with prevalent HF, higher concentrations of DHA, omega-3 and omega-6 FAs, and total PUFA were associated with a 16-32% lower risk of cardiovascular death, consistent with their protective effect on all-cause mortality. A similar pattern of lower risk was observed for BCAAs and small HDL particles, with a 16-31% reduction in cardiovascular mortality.

Potential mechanisms

Aromatic amino acids

In single marker models, both measured aromatic amino acids phenylalanine and tyrosine were not significantly associated with incident HF. In contrast, tyrosine was significant contributor to “Energy metabolism-Amino acids” cluster score which was linked with a higher HF risk in this study. Moreover, a higher concentration of phenylalanine was the only metabolite measure linked with lower LVEF in this study. This association was not replicated for tyrosine, nor for “Energy metabolism-Amino acids” cluster score. For HF survival, none of amino acids nor the score was linked with all-cause or CVD mortality. These results are partially inconsistent with previous studies. Higher phenylalanine concentration was positively associated both with prevalent and incident HF, as well as worse prognosis in HF survivors (Chen, Wang, et al., 2020; Delles et al., 2018).

Phenylalanine is hydroxylated to tyrosine via the enzyme phenylalanine hydroxylase, and subsequently, tyrosine serves as a critical precursor for the biosynthesis of catecholamines, thyroid hormones, and melanin. Impaired phenylalanine catabolism can potentially explain its role in HF development. Elevated phenylalanine levels appear to induce a state of cellular 'aging' in cardiomyocytes by upregulating p21, a well-known cell cycle inhibitor. This impedes the normal regeneration capacity of the cardiac tissue and sets the stage for a cascade of senescence-related changes, leading ultimately to functional decline characteristic of HF. Concurrently, phenylalanine's oxidative metabolites contribute to a pro-oxidative cellular environment. This oxidative stress further deteriorates cardiac function by impairing mitochondrial function, damaging DNA, and exacerbating cellular senescence. Importantly, these pathological changes are not static; they can be reversed. Both the administration of tetrahydrobiopterin, which enhances hepatic catabolism of phenylalanine, and dietary phenylalanine restriction have been shown to reverse these detrimental cardiac outcomes in a mouse model (Czibik et al., 2021).

On the other hand, elevated phenylalanine concentrations were associated with markers of subclinical inflammation, which is an independent contributor to HF pathogenesis (Chen, Wang, et al., 2020). Moreover, elevated phenylalanine is more indicative for prevalent CAD or DM, which were excluded from this study. This could eventually lead to weakening the association between aromatic amino acids and incident HF and explaining inability to detect it.

Branched-chain amino acids

Among BCAA, a significant inverse association with incident HF was found for valine, but not for leucine, isoleucine, or their total concentration. Estimate for valine was no longer significant after adjusting for interim AF and VHD. Moreover, none of BCAAs was identified in derived cluster scores. In terms of MRI function, higher concentration of all measures were linked with poorer LV function in healthy participants. It is intriguing that all three individual amin acids were significantly inversely associated with risk of all-cause mortality and were one the strongest predictors for CVD mortality in individuals with prevalent HF.

Higher concentration of BCAA was associated with a higher HF risk, but not consistently in all previous prospective studies (Delles et al., 2018; Lim et al., 2020). The link of BCAA with higher HF risk would not be surprising given strong evidence supporting their positive association with many cardiometabolic conditions (Morze et al., 2022). The mediation of transcriptional regulation by KPF-15 is crucial for the coordinated diminution of BCAA breakdown. Studies at both the genetic and cellular strata propose that deficiencies in BCAA catabolism, leading to an accumulation of branch chain keto acids byproducts, induce cardiac oxidative stress and systemic metabolic changes. These factors are notably influential in the advancement of HF (Sun et al., 2016). However, some studies in mouse models of HF while confirming downregulation of genes responsible for BCAA catabolism and increase in cardiac end-metabolites levels did not detect changes in plasma concentration (Sansbury et al., 2014; Wang et al., 2016). Another study did not confirm effects of promoting BCAA oxidation in heart with a pharmacological agent. Instead, whole body activation of BCAA oxidation led to cardiac protection and decreased blood pressure (Murashige et al., 2022). Compared with skeletal muscles, brown adipose tissue and liver, cardiac cells are a minor contributor to BCAA oxidation. Moreover, catabolism is not the only factor influencing the net circulating concentration of BCAA. Both proteolysis, intake, supplementation and to minor extend gut microbiota are additional important factors influencing the pool of available BCAA. Therefore, it might a case that in terms of longitudinal HF risk, plasma BCAA concentration is not necessarily an indicator of dysfunction in cardiac BCAA utilization.

Instead, decreased BCAA concentration can be observed with increasing age linked with a higher prevalence of sarcopenia, frailty and malnutrition, which all are associated with worse prognosis. This could be one of the reasons explaining the inverse associations between BCAA and survival in individuals with prevalent HF in this study. Such an explanation finds

confirmation in previous studies. Analysis of Ludwigshafen Risk in Cardiovascular Health Study showed that while BCAA were directly associated with numerous cardiometabolic risk factors in baseline, there was a strong inverse association with mortality in participants with intermediate-to-high cardiovascular risk (Moissl et al., 2023). Recently, another analysis showed that a joint metabolic score including low isoleucine, leucine, and valine as markers of malnutrition was associated with worse survival in a community cohort of individuals with HF. This association was independent from prognosis score based on established clinical factors (Conners et al., 2023).

Glycine

Higher glycine concentration was found to be nominally associated with lower incident HF in this study. Moreover, glycine was a significant negative contributor to "Albumin-Amino Acids" cluster score which showed second strongest association with incident HF. Individual glycine levels were not associated with prognosis in individuals with HF, but a significant association was observed for "Albumin-Amino Acids" cluster score and all-cause mortality. In relation to LV function, increase in glycine concertation was linked to a higher SVI.

Prospective studies previously linked circulating glycine concentration with lower risk of CAD and type 2 DM (Ding et al., 2015; Morze et al., 2022). Moreover, causal inverse association of glycine on CVD risk was supported by mendelian randomization study suggesting that protective effect of glycine can be attributed to lowering blood pressure (Wittemans et al., 2019). In contrast, insulin resistance was identified as one of the factors accounting for reduced glycine levels. This could also serve as one of the explanations for the inverse association with HF, even though this association remained present after adjusting for glycated hemoglobin and BMI in this study. Glycine generally shows pleotropic antioxidant, anti-inflammatory, cytoprotective, and immunomodulatory effects in peripheral tissues. Except for intermediate mechanisms and association with risk factors, there are possible direct mechanism supporting protective role of glycine in HF. In a murine model, glycine administration caused reverse remodeling in LV hypertrophy caused by transverse aortic constriction and administration of angiotensin II. The mechanism underlying this effect is hypothesized to involve the activation of glycine receptors, suppression of extracellular signal-regulated kinase phosphorylation, and reduced synthesis of TGF-β and endothelin-1 in cardiomyocytes (Lu et al., 2017).

Histidine

This study reported robust inverse association between histidine and incident HF, which was largely unexplained by contribution to interim predisposing events. Additionally, histidine notably contributed negatively to the HFMSS and "Albumin-Amino Acids" cluster score, which both exhibited the strong association with incident HF. Moreover, higher concentration of histidine and higher value of both scores were associated with better survival in individuals with prevalent HF. No link was identified between histidine and LV function in MRI.

These findings align with a prior retrospective study that identified low histidine levels as an effective marker for distinguishing between different stages of HF, as well as a significant predictor of increased mortality in patients with stage C and stage D HF (Wang et al., 2018). In prospective cohort studies, histidine concentration was inversely correlated with incident HF; however, the association lost its significance upon adjustment for covariates, likely due to limited statistical power (Delles et al., 2018; Jauhiainen et al., 2021).

Mechanisms for protective effect of histidine are unclear. Histidine is an essential amino acid, responsible for proton buffering, metal ion sequestration, neutralization of reactive oxygen and nitrogen species, and modulation of the histaminergic system. One Chinese case-control study observed that loss-of-function polymorphism in a gene coding histidine decarboxylase, which coverts histidine to histamine, was associated with lower HF prevalence (He et al., 2015). Combined histidine and β-alanine improved functional but not echocardiographic parameters in a rat HF model (Stefani et al., 2020). Moreover, histidine is a significant amino acid contributor to several proteins including hemoglobin, and histidine-rich calcium-binding protein. The latter, responsible for regulating the calcium ion concentration in the sarcolemma, plays a significant role in the development of cardiomyopathies and arrhythmias. Considering the abundant reserve of histidine in muscles and hemoglobin, its pool tends to be long-lasting. However, it may be speculatively diminished in HF, a condition associated with increased proteolysis, anemia, and the utilization of histidine for the formation of glutamate that enters the TCA cycle (Wang et al., 2018).

Other amino acids

Higher alanine concentrations were significantly and robustly associated with a lower HF risk. Alanine negatively impacted both the HFMSS and the “Amino acids-Energy metabolism” scores, both of which were associated with an elevated HF risk. Neither alanine nor scores were

linked with LV function or survival in HF. Glutamine was not associated with any of the outcomes, with exception for positive association with SVI.

The results for alanine are somewhat counterintuitive, especially considering its positive link with CAD and type 2 DM (Morze et al., 2022; Ruiz-Canela et al., 2017). The inverse relationship remained consistent even after adjustments were made for the incidence of both conditions prior to a HF diagnosis. Earlier studies examining the relationship between alanine and HF have yielded mixed outcomes. One prospective cohort study identified a positive correlation with HF risk, but this relationship was weakened after adjusting for BMI (Jauhiainen et al., 2021). The majority of studies did not support a link between alanine and HF risk or prevalence. In a single case-control study, alanine concentrations were significantly lower in HF cases and exhibited an inverse relationship with higher NYHA classes (Aquilani et al., 2017). The mechanisms underlying alanine's role in HF remain unclear. Given the ubiquity of alanine as a component amino acid in various proteins, reduced alanine concentration could possibly be a result of its utilization in gluconeogenesis. However, if this were the case, the association with alanine would likely be more pronounced in HF survival rather than incident HF.

Glycolysis

None of the glycolysis-related metabolite measures—namely, lactate, pyruvate, glucose, nor citrate—were associated with incident HF in this study. Among them, pyruvate showed a strong positive association with mortality in individuals with prevalent HF. Regarding LV function, glucose, lactate, and pyruvate showed a positive association with CI.

Previous investigation in ARIC showed that lactate levels were associated with a higher incident HF in general population (Matsushita et al., 2013). In patients with acute HF, elevated blood lactate levels on admission were prevalent in 43% of the cohort and were associated with higher markers of organ dysfunction, including high-sensitivity troponin I, aspartate aminotransferase, and endothelin-1. These elevated lactate levels predicted poor outcomes, including a higher one-year mortality rate, even after adjusting for other well-established prognostic variables (Zymliński et al., 2018). Pyruvate concentration was linked with higher HF risk in a Finnish study, however the model included only adjustment for age, BMI, use of statins and DM (Jauhiainen et al., 2021). In Cardiovascular Health Study, circulating citrate nor pyruvate was not associated with incident HF (Niezen et al., 2023). This lack of association was confirmed by observation form metabolomic nested case-control study in PREDIMED

trial (Bulló et al., 2021). Consistently with our study higher citrate concentration was associated with worse survival in patients with HF (Stryeck et al., 2019).

Previous research indicated that disruption in lactate-pyruvate axis might play a significant role in HF. In a healthy cardiac metabolic state, pyruvate primarily enters the mitochondria via the MPC to participate in the TCA cycle, which provides the energy currency ATP for the cell. Loss of MPC activity led to an imbalance in the pyruvate-lactate axis, promoting the conversion of pyruvate to lactate and its subsequent export via the MCT4 lactate exporter. This was shown to be sufficient to cause cellular cardiac remodeling, hypertrophic growth and development of HF (Fernandez-Caggiano & Eaton, 2021). Inhibiting the MCT4 lactate exporter had a preventative effect on cardiac hypertrophy. This inhibition caused an intracellular buildup of lactate, which then reverses the direction of LDH activity towards pyruvate production. Accumulation of pyruvate then drives its entry into the mitochondria, enhancing mitochondrial metabolism. One notable outcome was a reduction in reactive oxygen species, likely due to the free radical scavenging properties of lactate and pyruvate, thereby improving mitochondrial health. Moreover, the imbalance in the lactate-pyruvate axis also had implications for other metabolic pathways. For example, it affected the compartmentalization of aspartate and glutamate, which are key metabolites in the TCA cycle and in cellular biosynthesis (Cluntun et al., 2021). Although the disruption in the lactate-pyruvate axis presents an elegant mechanism for metabolism-induced cardiac dysfunction, a failure in this axis does not necessarily manifest as an increase in blood lactate and pyruvate levels. Additionally, this study excluded participants with high baseline cardiovascular risk, which could potentially result in the magnitude of the association being too small to detect.

Ketone bodies

All ketone bodies acetone, acetate, acetoacetate and 3-OHB were consistently associated with a higher HF risk. Moreover, higher concentrations of acetone, acetoacetate and 3-OHB were associated with higher all-cause mortality, while acetate with higher CVD mortality. Regarding LV function in healthy participants acetone and 3-OHB were linked with a higher CI, while acetone with a greater SVI.

These findings are fully consistent with recent prospective cohort studies. In analysis of data from Multi-ethnic Study of Atherosclerosis (MESA), 10-fold increase in baseline concentration of ketone bodies was associated with 68% higher incident of HF. The association was consistent for individual ketone bodies 3-OHB and acetoacetate. Moreover, this

association was consistent after exclusion of participants with a very high concentrations, which could indicate keto-acidosis (95% of participants had concentration <0.6 mmol/L) (Shemesh et al., 2023). Similarly, among 1758 participants from Cardiovascular Health Study, total ketone bodies were associated in a dose-response manner with incident HF and all-cause mortality (Niezen et al., 2023).

Despite the robustness of these associations, the potential mechanisms explaining them are not fully understood. In patients with HFrEF, there is a notable increase in myocardial ketone oxidation rates. This upregulation is supported by evidence of higher levels of ketogenic derivatives, specifically 3-hydroxybutyryl-CoA, and increased expression of ketone oxidative enzymes such as BDH1 in myocardial tissue. Animal models with induced HF through transverse aortic constriction and distal coronary ligation also exhibit similar upregulation in ketone transport proteins. This phenomenon appears to be a compensatory mechanism aimed at sustaining energy production in the face of reduced utilization of other substrates. However, the elevated rate of ketone body oxidation doesn't result in a reduced pool of circulating substrates; instead, it signifies a chronic increase in energy demand (Lopaschuk & Dyck, 2023).

The concept of augmenting ketone oxidation has been advanced as a therapeutic avenue for treating HF. This objective can be achieved through various means: direct administration of ketone bodies, implementation of a ketogenic diet, or the use of SGLT-2 inhibitors. Acute injection of 3-OHB into patients with HFrEF led to positive hemodynamic outcomes. Interestingly, these benefits did not go hand-in-hand with increased cardiac efficiency. This implies that the positive effects of 3-OHB may not come from the heart using a more efficient fuel, but rather from an increased supply of ketonic energy to the heart (Nielsen et al., 2019). A study in murine and rat models of HF used two different ketone ester supplementation strategies to investigate their effects on cardiac dysfunction. The study showed that both types of ketone esters were effective in attenuating LV dysfunction and remodeling, although they did not affect cardiac fibrosis post-myocardial infarction (Yurista et al., 2021). One study examining the impact of carbohydrate restriction showed that a month-long ketogenic regimen mitigated cardiac hypertrophy and hindered the progression of HF in wild-type mice subjected to transverse aortic constriction, as compared to those on a control high-carbohydrate diet (Nakamura et al., 2021). Another recent investigation unveiled two key insights: firstly, cardiac endothelial cells have the ability to metabolize ketone bodies, which fosters cell proliferation, migration, and vessel sprouting; and secondly, a ketogenic diet can enhance endothelial cell

proliferation and counteract the pathological rarefication of blood vessels in a mouse model of cardiac hypertrophy (Weis et al., 2022). Empagliflozin positively impacted cardiac performance and led to structural improvements in non-diabetic rats suffering from LV dysfunction after myocardial infarction. This was correlated with significant advancements in cardiac metabolism and ATP production. Empagliflozin also raised circulating ketone levels and enhanced myocardial expression of both the ketone body transporter and crucial ketogenic enzymes, implying increased myocardial utilization of ketone bodies. These metabolic alterations were collectively associated with an increase in cardiac ATP production (Yurista et al., 2019). Treatment of patients with canagliflozin led to a roughly 2-fold increase in serum ketone body levels in Japanese patients with T2DM. The median percent change from baseline was 62% for acetoacetate and 78% for 3-OHB (Polidori et al., 2018).

Fatty acids

In the study, metabolite measures related to PUFA, notably omega-3 FAs, were robustly associated with lower incident HF. Inverse associations were consistently observed for total concentration of PUFA total concentration of omega-3 FAs, overall degree of unsaturation, and DHA. These associations were not found for omega-6 FAs and the specific example of linoleic acid. Regarding survival in individuals with prevalent HF, similar trends were observed – higher concentrations of the aforementioned metabolite measures were the strongest predictors of lower all-cause and CVD mortality. Additionally, DHA and degree of unsaturation negatively contributed to HFMSS and "Omega-3 Fatty Acids" cluster score, both associated with higher HF risk and poor prognosis in individuals with prevalent HF. Higher concentrations of omega-3 FA-related metabolites were linked with a better LV function in MRI, notably with higher LVEF and SVI. The individual concentrations of total FAs, as well as saturated and PUFA, were largely not associated with any study outcomes. However, the latter two, along with omega-6 FAs, were significant contributors to the "Triglycerides, Fatty Acids and Other Lipids" cluster, which was linked to both a higher incidence of HF and worse survival outcomes in individuals with HF.

Association between FAs and HF outcomes can be considered in at least three different dimensions: dietary intake, circulating FA levels and supplementation with particular FAs. Replacement of intake from saturated and trans FAs with mono- and PUFA is a cornerstone of dietary prevention of CVD. While there is a considerable body of evidence showing benefit of dietary fat replacement on lowering CVD incidence, studies directly linking dietary fat intake

and HF incidence are sparse. In the Physicians' Health Study, dietary marine omega-3 FAs showed a trend toward a lower HF risk, and fish (a rich source of this nutrient) intake greater than once per month was associated with a reduced HF risk (Wilk et al., 2012). In the PREDIMED trial, following a Mediterranean diet with either high monounsaturated fat or polyunsaturated fat intake was not associated with a lower HF incidence compared to the low-fat diet group; however, the study could be underpowered to detect this association. (Papadaki et al., 2017).

Circulating levels of specific plasma omega-3 FAs were differentially associated with the HF risk in US male physicians. While plasma α-linolenic acid and docosapentaenoic acid were inversely associated with HF risk in a nonlinear fashion, eicosapentaenoic acid (EPA) and DHA showed no such association (Wilk et al., 2012). MESA study found that higher levels of plasma phospholipid, EPA, DHA, and docosapentaenoic acid were equally associated with a reduced HF risk, including both HFrEF and HFpEF (Block et al., 2019). Trends toward lower congestive HF risk were also observed for docosapentaenoic acid and total long-chain omega-3 FAs, while DHA showed no significant association (Mozaffarian et al., 2011). Regarding circulating saturated and omega-6 FAs, results from the Physicians' Health Study revealed no significant association with incident HF (Matsumoto et al., 2013; Petrone et al., 2013). Recently, analysis of lipidomic data in PREDIMED and EPIC highlighted the importance of lipid compounds containing saturated palmitic oleic fatty acyl chain for HF prediction (Wittenbecher et al., 2021). In context of PUFA, there was evidence of a higher HF risk associated with increased levels of the oleic acid, as well as concentrations of long-chain PUFA (Imamura et al., 2013; Steffen et al., 2018).

In the GISSI-HF trial, administering low-dose omega-3 PUFAs (0.84 g/day) alongside conventional treatment lowered the risk of all-cause mortality, as well as hospitalization due to HF (Heydari et al., 2016). In the OMEGA-REMODEL trial, elevated omega-3 PUFA dosage (3.4 g/day) for 6 months following MI led to reductions in infarct size, enhancements in ventricular systolic performance, and decreased non-infarct myocardial scarring (Heydari et al., 2016). In contrast, the VITAL trial revealed that low-dose omega-3 PUFAs (0.84 g/day) were ineffective in modifying incidence a composite of cardiovascular death, nonfatal MI, and stroke over a 5-year duration, although a secondary endpoint indicated a lower MI risk (Heydari et al., 2016). This diverges from the REDUCE-IT trial findings, where a 25% risk reduction in composite cardiovascular events, including death, nonfatal MI, nonfatal stroke, coronary

revascularization, and unstable angina, was observed with high-dose icosapent ethyl (EPA precursor) at 4 g/day. However, REDUCE-IT noted no change in the HF incidence, showing 4.1% with placebo compared to 4.3% with EPA (Bhatt et al., 2019). Similarly, supplementation with high-dose omega-3 FAs in STRENGTH trial did not lead to lower HF incidence compared to corn oil (Nicholls et al., 2020). Despite sparse findings, effectiveness of omega-3 FAs supplementation in HF seems to depend on baseline concentration, as well as provided dose. In contrast, a recent meta-analysis of 12 trials indicated that omega-3 FAs reduced rehospitalization rate in HF patients (Barbarawi et al., 2021).

Except for reduction in risk of predisposing events such as CAD, omega-3 FAs show potential direct mechanisms of action in HF. Omega-3 FAs positively influence LV function and remodeling. In patients with nonischemic dilated cardiomyopathy and minimal symptoms, treatment with 2 g of n-3 PUFAs for 12 months significantly improved LV systolic function, functional capacity, and reduced HF hospitalization rates compared to placebo (Nodari et al., 2011). Another study found that angiotensin II induces miR-21 expression, leading to suppression of RECK and promoting cardiac fibroblast migration, effects that are mitigated by omega-3 FAs, specifically DHA. DHA was more effective than EPA in reversing angiotensin II-induced miR-21 expression and its downstream effects, offering potential therapeutic benefits against cardiac fibrosis (Siddesha et al., 2014). A cell culture model found that DHA attenuated tumor necrosis factor-α-induced expression of genes related to senescence and endothelial cell dysfunction, and reversed the suppression of endothelial NO synthase, suggesting its regulatory role in endothelial health (Yamagata et al., 2016). Lastly, omega-3 FAs show anti-inflammatory potential by changed FA composition in cell membrane phospholipids, destabilization of lipid rafts, suppression of the pro-inflammatory transcription factor nuclear factor kappa B, leading to decreased expression of inflammatory genes, activation of the anti-inflammatory transcription factor NR1C3 and engagement with the G protein-coupled receptor GPR120 (Calder, 2013).

Inflammation

The Nightingale platform included a measure of one inflammation-related metabolite – glycoprotein acetyls. A higher concentration of this metabolite is associated with an elevated HF risk and worse LV function (lower CI, LVESVI, LVEDVI, and SVI). However, no correlation was found with all-cause or CVD mortality among individuals with HF.

Historically, inflammation has been one of the earliest mechanisms investigated in the pathophysiology of HF. Various inflammation biomarkers have consistently been reported to correlate with higher risk and poorer survival (Reina-Couto et al., 2021; S. Zhao et al., 2021). In this study, the significance of the association for glycoprotein acetyls diminished after adjusting for preceding CAD, suggesting that the inflammation signal in HF is primarily explained by this interim event. This is somewhat counterintuitive, as CAD is more commonly linked with HFrEF, which generally shows lower levels of inflammation biomarkers compared to HFpEF.

Glycoprotein acetyls serve as specific inflammation-sensitive signals observable in the 1H-NMR spectra of both serum and plasma in clinical settings. This NMR signal mainly stems from the N-acetyl methyl group protons, which are associated with mobile N-acetyl glucosamine residues present on the glycan sections of acute-phase proteins like α1-acid glycoprotein, haptoglobin, α1-antitrypsin, and α1-antichymotrypsin. Quantified in mmol/L, glycoprotein acetyls levels correlate with markers such as IL-6, tumor necrosis factor-α, fibrinogen, C-reactive protein, serum amyloid A, and lipoprotein-associated phospholipase A2 (Connelly et al., 2017).

Empirically, glycoprotein acetyls have shown greater temporal stability than C-reactive protein. This stability, coupled with their reduced sensitivity to acute infections – an average increase of approximately 5% in subjects reporting acute illness within the preceding three weeks, compared to a roughly 120% increase for C-reactive protein, makes them more reliable markers of subclinical inflammation, which plays a pivotal role in HF pathophysiology (Chiesa et al., 2022).

Fluid balance

In this study, a higher concentration of albumin showed a robust inverse association with the HF risk. Albumin was a significant contributor of HFMSS and "Albumin-Amino-acids" cluster score both associated with incident HF. Furthermore, among individuals already diagnosed with prevalent HF, elevated levels of albumin were correlated with improved survival outcomes. Additionally, albumin was positively associated with CI, LVEF, and SVI, as determined by MRI assessments of the LV.

Consistently with these findings, hypoalbuminemia has been for a long time recognized as a risk factor for both incident HF in general population, as well as poor outcomes in patients with

developed HF (Gopal et al., 2010; Liu et al., 2012; Zhuang et al., 2020). Albumin is a protein synthesized in the liver, and its plasma concentration is primarily determined by various factors such as the rate of albumin synthesis, exogenous loss of albumin, and dilutional effects. Approximately one-quarter to one-third of patients with HF have hypoalbuminemia at the time of diagnosis. Hypoalbuminemia can either be a secondary outcome of other pathophysiological processes contributing to HF, or it can independently influence the onset of HF through its biological activity. Net decrease in circulating albumin concentration in patients with HF can be attributed to various factors including liver dysfunction, hemodilution, cardiac cachexia affecting BMI, and renal failure (Liu et al., 2012). Albumin is a very highly abundant plasma protein and plays a crucial role in maintaining fluid balance across capillaries primarily through its osmotic effect and sodium ion attraction. It also safeguards capillary membrane integrity – loss of albumin increases capillary permeability, causing edema and sodium retention. Moreover, albumin can exhibit anticoagulant and antiplatelet activities by inhibiting histone-induced aggregation and enhancing PGI2 bioavailability. It can further act by binding to platelet agonists like thromboxane A2. Albumin also exhibits anticoagulant actions via fibrinogen interaction and is suggested to have heparin-like activity. Additionally, its synthesis is negatively impacted by pro-inflammatory cytokines, yet it acts as an anti-inflammatory agent by inhibiting endothelial activation. Lastly, albumin offers significant antioxidant benefits, mitigating the effects of reactive-oxygen and nitrogen species in the plasma (Zoanni et al., 2023).

Triglyceride-rich lipoproteins

In this study, metabolite measures related to triglyceride-rich lipoproteins exhibited a general trend towards a decreased HF risk. At the lipoprotein class level, the triglyceride content in VLDL particles (their primary lipid fraction) was inversely associated with the HF incidence. This was observed without any notable correlation with particle count or cholesterol levels. A larger average diameter for VLDL particles correlated with a lower HF risk. In a consistent manner, larger VLDL size subclasses (ranging from chylomicrons and very large VLDL to small VLDL and their respective lipid contents) were associated with a reduced HF risk. In contrast, very small VLDL subclasses indicated a trend towards a higher HF risk. However, after correcting for multiple testing, only specific metabolite measures in very large, large, and medium VLDL subclasses – namely, concentration, triglycerides, and total lipids = showed significant inverse associations. Subgroup and sensitivity analyses revealed that these aforementioned measures were no longer significant after adjusting for interim AF or VHD.

Interestingly, network analysis did not generally identify triglyceride content or particle concentration within these size subclasses as significant contributors to cluster-specific scores. An exception was found in the triglyceride content of large VLDL particles, which negatively contributed to the 'Triglyceride-rich Lipoprotein' cluster score, subsequently being not associated with incident HF after full covariate adjustment. The same pattern of associations was observed for VLDL survival, showing a generally inverse relationship between VLDL class and size subclass measures.

Previous studies in the context of triglyceride-rich lipoproteins and HF have primarily focused on approximations derived from conventional lipid profiles, such as total triglycerides and remnant cholesterol concentrations. However, both of these are imperfect determinants of the actual count of triglyceride-rich lipoproteins. In 5688 participants aged 45 to 84 from the MESA, who had no clinical CVD and were not on lipid-lowering medications, triglyceride levels were not associated with a higher HF risk in those without DM. However, in individuals with baseline DM, both continuous triglyceride concentration and hypertriglyceridemia (>150 mg/dL) were linked to a significantly higher HF risk, with twofold and 60% increases, respectively (Ebong et al., 2013). Analysis of 113,554 individuals from the Copenhagen General Population Study and the Copenhagen City Heart Study, including those with lipid-lowering therapy and CAD, revealed a stepwise increase in HF risk with higher non-fasting triglyceride levels. The risk increase for a 1 mmol/L (88 mg/dL) increase in triglycerides was 19%, consistent across conventional risk factors for HF (Varbo & Nordestgaard, 2018). In a real-world retrospective study involving statin-treated patients aged ≥45 with high cardiovascular risk factors like DM and/or atherosclerotic CVD, elevated triglyceride levels (TG ≥150 mg/dL) were associated with a 19% higher rate of new HF diagnoses, and the sub-cohort with TG levels between 200-499 mg/dL had a 24% higher rate, each compared to a comparator cohort (Toth et al., 2019). A secondary analysis of the ARIC study involving 12,595 participants without coronary heart disease, showed that elevated levels of remnant cholesterol were found to be progressively associated with a higher HF risk during a median follow-up of 22.5 years (Liu et al., 2022).

Despite the mixed results of previous studies on triglyceride-rich lipoproteins and HF, one would expect a prospective study to show a positive association, which eventually may become non-significant after adjusting for causes of HF. In other words, a positive association with triglyceride-rich lipoproteins could be explained by contribution to events predisposing to HF,

for example by impacting atherosclerosis in CAD or by co-occurrence in hypertension as a component of metabolic syndrome and other consequences of obesity. This study showed a trend towards lower HF risk for all metabolite measures with exception for small VLDL, which eventually became attenuated after adjustment for interim events – in this case after accounting for interim AF. There are several possible explanations for this phenomenon. First, this study excluded high-baseline risk individuals, such as participants with DM and prevalent CVD events. This could potentially weaken the association for the rest of the study population, as most participants at risk due to elevated triglyceride-rich lipoproteins may have been removed from the analysis. As previously indicated, associations between hypertriglyceridemia and HF, while observed in participants with DM, are not present in individuals without HF (Ebong et al., 2013). Moreover, since the study did not collect data on follow-up treatment, it is possible that participants diagnosed with hypertriglyceridemia after recruitment received appropriate treatment, indirectly lowering their HF risk. Similarly, the observed association with survival in individuals with prevalent HF could be explained by a higher probability of receiving treatment during follow-up. Therefore, reverse causation bias cannot be excluded. An intriguing finding was the attenuation of significant associations for VLDL-related metabolite measures after adjustment for interim AF and VHD. Moreover, interim AF and VHD mediated large portion of associations between triglyceride-rich lipoprotein measures and HF. A recent prospective cohort study involving 65,136 individuals aged 45 to 60 years without prevalent CVD from the Swedish Apolipoprotein-Related Mortality Risk cohort showed an increased HF risk for increasing concentrations of triglycerides (Ding et al., 2022). This contrasts with our study, where triglycerides were associated with a lower interim AF risk. Interim AF and VHD could serve as a proxy for the diagnosis of other comorbidities during follow-up, and the introduction of treatment may attenuate the association between TG and HF. While a definitive advantage of this study is the direct profiling of lipoprotein-rich lipoproteins, rather than relying on triglycerides or remnant cholesterol as a proxy, further studies are required to elucidate the role of triglyceride-rich lipoproteins in HF.

Cholesterol-enriched lipoproteins

Metabolite measures associated with cholesterol-enriched lipoproteins generally correlated with a lower HF risk. An inverse association was observed for LDL cholesterol but not for the count of LDL particles. While other lipid fractions within LDL trended toward lower HF risk, they either did not reach nominal statistical significance or did not survive multiple testing correction. The pattern of associations was consistent across large, medium, and small LDL

size subclasses. Conversely, no null associations were observed for measures related to IDL particles. Associations significantly linked with incident HF were attenuated after adjustment for interim AF and VHD. The “Cholesterol-enriched lipoproteins “cluster score representing joint exposure to higher circulating cholesterol-enriched particles was not associated with incident HF in either the basic or comprehensively adjusted models. Regarding survival among individuals with prevalent HF, all LDL class and subclass-related metabolite measures were linked with higher all-cause mortality but not CVD mortality.

Unlike for triglyceride-rich lipoproteins, previous studies show more consistent findings regarding cholesterol-enriched lipoproteins. A recent analysis of four large, U.S.-based prospective cohort studies: ARIC, Coronary Artery Risk Development in Young Adults Study, FHS Offspring Cohort, and MESA showed that cumulative LDL cholesterol, its time-weighted average and slope during young adulthood and middle age showed no association with incident HF (Zhang et al., 2021). Similarly, LDL cholesterol was not associated with incident HF in 113,554 individuals from the Copenhagen General Population Study and the Copenhagen City Heart Study (Varbo & Nordestgaard, 2018).

The expected null association was not observed in this study; instead, many measures showed an inverse association. This, similar to the case with triglyceride-rich lipoproteins, can be partially explained by potential reverse causation. Participants with higher circulating cholesterol-enriched lipoproteins were more likely to receive treatment, thus lowering their risk—most importantly for CAD, which is the most prevalent HF cause. For participants with prevalent HF, these inverse associations have also been observed in previous studies and are referred to as the "cholesterol paradox." An extreme example of this is advanced HF, where a catabolic and cachexic state may lead to lower circulating cholesterol concentrations (and its fractions including LDL), thereby serving as an indicator of illness severity and poor prognosis (Velavan et al., 2007).

High-density lipoproteins

In this study, metabolite measures associated with the HDL particle class showed no association with incident HF. However, the relationship varied substantially depending on the size of the HDL particle, as indicated by a positive association between average diameter for HDL and HF risk. Specifically, metabolite measures related to large and very large HDL showed a positive association with HF risk, while those related to medium and small HDL displayed an inverse association. This pattern of associations was also observed for all-cause

and, to a lesser extent, for CVD mortality. HDL size subclasses demonstrated strong but opposite associations with prognosis in individuals with prevalent HF, with respectively smaller and larger HDL serving as the strongest predictors for better and worse survival in this group.

These findings are highly in line with previous investigations. In a case-control study using the CATHGEN biorepository, patients with HFrEF, HFpEF, and no HF underwent NMR-based lipoprotein profiling. The study found higher concentrations of large HDL particles in HFrEF compared to controls and HFpEF, while smaller concentrations of small HDL particles were noted in HFpEF and HFrEF compared to controls. Notably, these three groups did not differ in terms of HDL cholesterol. Total and small HDL particles were inversely associated with all-cause mortality across all groups. On the other hand, greater mean HDL size was linked with an increased risk of all-cause mortality for all groups, but larger HDL particles were linked with a higher risk only in HFpEF and the control group, not in HFrEF (Hunter et al., 2019). Similar pattern of associations was observed in PROPSER trial, however was no longer evident following covariate adjustment (Delles et al., 2018).

The primary mechanism of small HDL particle involvement in HF pathogenesis is their enhanced protective capacities over large and very large HDL particle. These include greater reverse cholesterol transport facilitated by high surface lipid fluidity and low lipid content, potent antioxidant abilities to remove oxidized lipids from LDL, anti-inflammatory effects via ABCA1 and Factor H, and endothelial cell protection and vasodilation due to enrichment in sphingosine-1 phosphate and apoA1. These attributes collectively contribute to improved endothelial cell function, vasorelaxation, and potentially reduced systemic inflammation, offering key protective effects on the failing heart and systemic vasculature (Hunter et al., 2019).

Clinical application and future perspectives

The overarching aim of this research agenda is to advance long-term strategies in personalized, precision medicine for the management of heart failure. The availability of extensive omics datasets - including metabolomics, genomics, transcriptomics, proteomics, epigenomics, and metagenomics - has radically transformed the field of biology, giving rise to the discipline of systems biology. This approach aims to construct comprehensive models of intricate biological interactions by synthesizing insights from various interdisciplinary domains in a holistic

manner, departing from the traditional reductionist perspective (Alyass et al. 2015). Extending this "systems" approach to epidemiology, researchers strive to unravel the complexity of biological information that underpins the impact of exposures on outcomes.

In recent years, there has been a marked increase in available therapeutic options for HF. Unlike in diseases such as cancer, new treatments for HF have often been first identified not in human studies, but in investigational models, such as animal studies (Feldman et al., 2017). Given that HF is a clinical syndrome with typically complex underlying etiologies, this approach provides more controlled conditions for studying disease pathogenesis, especially considering the limited availability of heart biopsies in human research. The advent of systems epidemiology, where deep clinical phenotyping is combined with multi-omic data, has enabled researchers to overcome the limitations inherent in previous reductionist clinical studies, facilitating hypothesis generation.

While earlier genome-wide association studies led to the discovery of new loci associated with specific pathological processes driving HF and resulted in the creation of multiple polygenic risk scores (Han et al. 2023, Lee et al. 2023), the application of other omic techniques has been limited. As stated in the introduction section of this thesis, human observational studies utilizing metabolomics in HF have been constrained by various methodological factors, including small sample sizes and questions regarding the validity and reproducibility of analytical platforms. Unlike CAD and type 2 DM, which have been the focus of more extensive research, the list of metabolites studied in the context of HF seldom overlapped between studies. Moreover, in many instances, a significant metabolic trait identified in one study could not be replicated in another.

Nevertheless, a few studies have utilized untargeted techniques to identify novel metabolomic traits associated with HF. The limited availability of exploratory data from these untargeted assays does not negate the value of targeted assays, which frequently cover several hundred annotated metabolites and are supported by robust external standards for quantification. In many cases, targeted metabolomic assays also include metabolites previously studied in HF in a hypothesis-driven manner based on results from pre-clinical studies. Although some well-studied metabolite classes were previously thought to have conclusive associations with significant alterations in HF, the present study challenges these assumptions by showing either opposite or null associations. A case in point are BCAA, which were strongly positively associated with HF outcomes and other cardiometabolic traits in prior research, yet displayed

a null association with incident HF and an inverse association with mortality in prevalent HF in the current study.

For other metabolite groups, despite existing mechanistic evidence and results from smaller studies, associations have not been tested in large, population-based cohorts. Ketone bodies serve as an example; they have only recently been examined in the context of incident HF in other large-scale studies. In summary, there is a relative dearth of evidence concerning metabolomics and HF risk compared to other chronic diseases. Thus, future analyses in other observational studies using various analytical techniques are warranted not only to corroborate current findings but also to identify novel biological signals.

Given that incident HF is a heterogeneous endpoint due to varying clinical manifestations and underlying etiologies, investigating its associations with imaging markers for the LV and RV provides deeper insights. Specifically, it helps elucidate how particular groups of metabolites may contribute to HF development and to specific disease phenotypes. This study is the first to explore prospective associations with LV parameters using MRI. This approach holds promise not just because systolic and diastolic dysfunctions are important predictors of patient prognosis, but also because impaired systolic and diastolic functions can serve as surrogate endpoints for approximating a pre-HF stage. In future investigations, it would be warranted to extend the current analyses to additional LV parameters, as well as those related to the RV. Another promising avenue for gaining a deeper understanding of the relationship between metabolites and HF involves stratifying phenotypes based on predisposing etiologies. This approach has been infrequently employed in previous epidemiological studies focused on HF. However, it offers the opportunity to disentangle the associations between specific metabolites and predisposing events from those between metabolites and HF. From a mechanistic standpoint, this is crucial; while predisposing events may serve as mediators for HF, their associated metabolomic patterns are not necessarily unique indicators for the onset of HF. Therefore, an approach that utilizes mediation by predisposing events warrants further validation in external cohort studies.

Beyond exploratory objectives, this study aimed to assess the utility of incorporating multi-metabolite measurement scores into classical risk factor models to enhance HF prediction in a general population. Although the study was not designed to develop a fully validated predictive model, the inclusion of metabolite measures did result in modest improvements in the AUC, as well as in net reclassification and discrimination indices. This suggests that metabolite

measures could potentially serve as additional biomarkers for HF onset. A recent systematic review identified 39 risk prediction models for HF within the general community (Nadarajah et al. 2023). Only a few of these models incorporated circulating biomarkers, particularly those focused on HF, such as natriuretic peptides. These models were deemed to have a high risk of bias. Consequently, a logical next step in this line of research would be to examine how these newly identified biomarkers, namely metabolites, could contribute to the development of a validated risk prediction model for HF. Particular attention could be paid to specific risk groups that have recently gained attention in the context of HF screening, such as patients with type 2 DM.

Lastly, given the escalating interest in polygenic risk scores and proteomic signatures specific to HF, the integration of these strategies with metabolomics and other omics technologies holds substantial promise for future multi-omic human studies in HF. This comprehensive approach has the potential to significantly advance the field in several key areas. Not only could it yield new molecular targets for therapeutic intervention, but it could also facilitate the exploration of data-driven disease subtypes. By doing so, this would enable a more nuanced understanding of the heterogeneity inherent in HF cases, thereby allowing for more personalized treatment regimens. Moreover, the holistic view provided by multi-omic analyses could lead to the identification of previously unrecognized interactions between different biological layers, offering a richer, more complete picture of HF pathophysiology. This, in turn, could provide clinicians with enhanced diagnostic and prognostic tools, thereby improving patient outcomes through targeted and timely interventions.

Strengths and limitations

The strength of the study lies predominantly in its groundbreaking approach as the first comprehensive prospective analysis of metabolomics in the context of HF. Utilization of the UK Biobank data, comprising over 98,000 participants with in-depth phenotyping information, enabled sufficiently-powered statistical analyses that included essential covariates, addressing a gap in previous studies. Implementation of a comprehensive statistical protocol was achieved, which encompassed rigorous multistep subgroup, sensitivity, and mediation analyses, as well as innovative multivariable methods like metabolite measure networks constructed using GGMs. Further robustness was added by the adoption of the Nightingale NMR platform, which has been externally validated and is widely employed in over 30 different cohorts globally. The inclusion of cardiac MRI data facilitated the investigation of indirect mechanisms linking

metabolite measures to cardiac function. Additionally, the study benefitted from a sizable sub-cohort of prevalent HF patients, providing an opportunity to examine the significance of metabolite measures for HF survival.

However, the present study faces several limitations that must be acknowledged. Firstly, although the prospective cohort design allows us to imply possible causality by addressing the temporality of associations, it falls short of confirming it. This limitation is especially acute due to the presence of residual confounding, particularly time-variant confounders like the introduction of treatment during the follow-up period. These factors could skew the observed relationships, potentially resulting in spurious findings.

Secondly, our study's reliance on the NMR platform to measure markers has its own set of constraints. The majority of the markers in this platform are related to lipoproteins and, to some extent, carry redundant information. Furthermore, the number of metabolite measures analyzed was comparatively smaller, which hampered our ability to make comprehensive inferences about certain biochemical pathways. For instance, our study failed to observe any associations for sphingolipids, a finding that diverges from previous lipidomic studies which have demonstrated varied outcomes depending on the fatty acyl residues in a sphingolipid particle. The use of MS-based methods would add depth to this metabolite analysis; however, the destructibility of samples limits its application in repositories like the UK Biobank.

Thirdly, the generalizability of the findings is constrained by the demographic composition of the cohort, which predominantly comprised White individuals. The analyses in subgroups with other ethnicities had limited statistical power. This is particularly noteworthy given the notable differences in the baseline risk of incident HF among Asian and Black individuals compared to the majority White cohort.

Fourthly, the study suffered from a lack of biological replicates, with only a small number of participants having repeated measurements. This absence did not allow us to account for the inherent biological variability in metabolite measures, which could influence conclusions for some classes of metabolite measures.

Fifthly, HF was defined using statistical codes based on hospital registries, a method while superior to older epidemiological definitions such as the Framingham criteria, comes with its own set of limitations. The way HF was diagnosed and coded could vary depending on the place of diagnosis and over time as new clinical guidelines were introduced. The absence of

centrally adjudicated outcomes, unlike in clinical trials, limited the ability to stratify HF according to LVEF into categories like HFrEF, HFmrEF, and HFpEF.

Sixthly, NT-pro-BNP nor BNP measurements were not available for the study. This omission is particularly problematic as natriuretic peptides as a gold standard in biomarker research in HF, and its absence hampers the comparative predictive ability of our study.

Finally, the study could potentially suffer from selection bias, especially in the survival HF sub-cohort. This bias may manifest due to the "healthy volunteer" phenomenon, which would result in the recruitment of more mild cases of patients with HF. Individuals with advanced HF, who are unable to undergo assessment at the UK Biobank assessment center, are thus underrepresented, limiting the generalizability of our findings to the broader HF patient population.

Conclusions

The primary aim of this study was to identify pathways associated with the development and progression of HF, independently from classical risk factors. Our exploration of metabolite measures in a primary prevention subcohort of UK Biobank participants revealed specific associations. Elevated levels of omega-3 FAs, small HDL particles, albumin, and selected amino acids such as alanine and histidine were inversely correlated with HF risk. Conversely, metabolite measures belonging to ketone bodies, very large HDL particles, and glycoprotein acetyls showed a direct association with HF. These findings remained consistent across subgroups, sensitivity analyses, and mediation analyses, suggesting that the up- or down-regulation of pathways related to these metabolite measures may play an independent role in the development of HF and is not merely a consequence of other predisposing etiologies.

When these metabolite measures were later aggregated into a multi-marker HF metabolic susceptibility score, or accounted for in data-driven cluster scores that considered correlations with other metabolic pathways, they revealed robust associations with HF. An additional aim was to explore the relationship between metabolite measures and LV function as assessed by MRI. For several metabolite measures, significant observations extended to both favorable and detrimental patterns of LV function, providing potential mechanistic evidence for their involvement in HF pathogenesis.

Furthermore, this study sought to link metabolite measures with survival rates among participants who had prevalent HF at baseline. The significant associations were replicated, particularly for ketone bodies, omega-3 FAs, and both small and large HDL particles. The evidence supports a significant interlink with HF on three levels: association with preclinical LV dysfunction as approximated by MRI, association with incident HF, and survival outcomes in patients with HF. This multi-level validation, corroborated by existing preclinical studies, substantiates the role of these metabolite measures in the development of HF. Further research is needed to confirm their validity as potential biomarkers, which could be beneficial in both the prevention and management of HF.

References

Aimo, A., Januzzi, J. L., Jr., Vergaro, G., Ripoli, A., Latini, R., Masson, S., Magnoli, M., Anand, I. S., Cohn, J. N., Tavazzi, L., Tognoni, G., Gravning, J., Ueland, T., Nymo, S. H., Brunner-La Rocca, H. P., Bayes-Genis, A., Lupón, J., de Boer, R. A., Yoshihisa, A., Takeishi, Y., Egstrup, M., Gustafsson, I., Gaggin, H. K., Eggers, K. M., Huber, K., Tentzeris, I., Tang, W. H. W., Grodin, J., Passino, C., & Emdin, M. (2018). Prognostic Value of High-Sensitivity Troponin T in Chronic Heart Failure: An Individual Patient Data Meta-Analysis. *Circulation, 137*(3), 286-297. https://doi.org/10.1161/circulationaha.117.031560

Alyass, A., Turcotte, M., & Meyre, D. (2015). From big data analysis to personalized medicine for all: challenges and opportunities. BMC medical genomics, 8, 33. https://doi.org/10.1186/s12920-015-0108-y

Aquilani, R., La Rovere, M. T., Corbellini, D., Pasini, E., Verri, M., Barbieri, A., Condino, A. M., & Boschi, F. (2017). Plasma Amino Acid Abnormalities in Chronic Heart Failure. Mechanisms, Potential Risks and Targets in Human Myocardium Metabolism. *Nutrients, 9*(11). https://doi.org/10.3390/nu9111251

Aragam, K. G., Chaffin, M., Levinson, R. T., McDermott, G., Choi, S. H., Shoemaker, M. B., Haas, M. E., Weng, L. C., Lindsay, M. E., Smith, J. G., Newton-Cheh, C., Roden, D. M., London, B., Wells, Q. S., Ellinor, P. T., Kathiresan, S., & Lubitz, S. A. (2019). Phenotypic Refinement of Heart Failure in a National Biobank Facilitates Genetic Discovery. *Circulation, 139*(4), 489-501. https://doi.org/10.1161/circulationaha.118.035774

Bachtiger, P., Petri, C. F., Scott, F. E., Ri Park, S., Kelshiker, M. A., Sahemey, H. K., Dumea, B., Alquero, R., Padam, P. S., Hatrick, I. R., Ali, A., Ribeiro, M., Cheung, W.-S., Bual, N., Rana, B., Shun-Shin, M., Kramer, D. B., Fragoyannis, A., Keene, D., Plymen, C. M., & Peters, N. S. (2022). Point-of-care screening for heart failure with reduced ejection fraction using artificial intelligence during ECG-enabled stethoscope examination in London, UK: a prospective, observational, multicentre study. *The Lancet Digital Health, 4*(2), e117-e125. https://doi.org/10.1016/S2589-7500(21)00256-9

Balasubramanian, R., Paynter, N. P., Giulianini, F., Manson, J. E., Zhao, Y., Chen, J. C., Vitolins, M. Z., Albert, C. A., Clish, C., & Rexrode, K. M. (2020). Metabolomic profiles associated with all-cause mortality in the Women's Health Initiative. *International Journal of Epidemiology, 49*(1), 289-300. https://doi.org/10.1093/ije/dyz211

Barbarawi, M., Lakshman, H., Barbarawi, O., Alabdouh, A., Al Kasasbeh, M., Djousse, L., & Manson, J. E. (2021). Omega-3 supplementation and heart failure: A meta-analysis of 12 trials including 81,364 participants. *Contemporary Clinical Trials, 107*, 106458. https://doi.org/10.1016/j.cct.2021.106458

Bertero, E., & Maack, C. (2018). Metabolic remodelling in heart failure. *Nature Reviews Cardiology, 15*(8), 457-470. https://doi.org/10.1038/s41569-018-0044-6

Bhatt, D. L., Steg, P. G., Miller, M., Brinton, E. A., Jacobson, T. A., Ketchum, S. B., Doyle, R. T., Jr., Juliano, R. A., Jiao, L., Granowitz, C., Tardif, J. C., & Ballantyne, C. M. (2019). Cardiovascular Risk Reduction with Icosapent Ethyl for Hypertriglyceridemia. *New England Journal of Medicine, 380*(1), 11-22. https://doi.org/10.1056/NEJMoa1812792

Biomarkers Definitions Working Group. (2001). Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clinical Pharmacology and Therapeutics, 69*(3), 89-95. https://doi.org/10.1067/mcp.2001.113989

Blecker, S., Paul, M., Taksler, G., Ogedegbe, G., & Katz, S. (2013). Heart failure–associated hospitalizations in the United States. *Journal of the American College of Cardiology, 61*(12), 1259-1267. https://doi.org/10.1016/j.jacc.2012.12.038

Block, R. C., Liu, L., Herrington, D. M., Huang, S., Tsai, M. Y., O'Connell, T. D., & Shearer, G. C. (2019). Predicting Risk for Incident Heart Failure With Omega-3 Fatty Acids: From MESA. *JACC: Heart Failure, 7*(8), 651-661. https://doi.org/10.1016/j.jchf.2019.03.008

Blumer, V., Mentz, R. J., Sun, J.-L., Butler, J., Metra, M., Voors, A. A., Hernandez, A. F., O'Connor, C. M., & Greene, S. J. (2021). Prognostic Role of Prior Heart Failure Hospitalization Among Patients Hospitalized for Worsening Chronic Heart Failure. *Circulation: Heart Failure, 14*(4), e007871. https://doi.org/10.1161/CIRCHEARTFAILURE.120.007871

Boorsma, E. M., ter Maaten, J. M., Damman, K., Dinh, W., Gustafsson, F., Goldsmith, S., Burkhoff, D., Zannad, F., Udelson, J. E., & Voors, A. A. (2020). Congestion in heart failure: a contemporary look at physiology, diagnosis and treatment. *Nature Reviews Cardiology, 17*(10), 641-655. https://doi.org/10.1038/s41569-020-0379-7

Bosch, L., Assmann, P., de Grauw, W. J. C., Schalk, B. W. M., & Biermans, M. C. J. (2019). Heart failure in primary care: prevalence related to age and comorbidity. *Primary Health Care Research & Development, 20*, e79. https://doi.org/10.1017/s1463423618000889

Bošnjak, I., Selthofer-Relatić, K., & Včev, A. (2015). Prognostic value of galectin-3 in patients with heart failure. *Disease Markers, 2015*, 690205. https://doi.org/10.1155/2015/690205

Bozkurt, B., Coats, A. J. S., Tsutsui, H., Abdelhamid, C. M., Adamopoulos, S., Albert, N., Anker, S. D., Atherton, J., Böhm, M., Butler, J., Drazner, M. H., Michael Felker, G., Filippatos, G., Fiuzat, M., Fonarow, G. C., Gomez-Mesa, J.-E., Heidenreich, P., Imamura, T., Jankowska, E. A., Januzzi, J., Khazanie, P., Kinugawa, K., Lam, C. S. P., Matsue, Y., Metra, M., Ohtani, T., Francesco Piepoli, M., Ponikowski, P., Rosano, G. M. C., Sakata, Y., Seferović, P., Starling, R. C., Teerlink, J. R., Vardeny, O., Yamamoto, K., Yancy, C., Zhang, J., & Zieroth, S. (2021). Universal definition and classification of heart failure: a report of the Heart Failure Society of America, Heart Failure Association of the European Society of Cardiology, Japanese Heart Failure Society and Writing Committee of the Universal Definition of Heart Failure. *European Journal of Heart Failure, 23*(3), 352-380. https://doi.org/10.1002/ejhf.2115

Bulló, M., Papandreou, C., García-Gavilán, J., Ruiz-Canela, M., Li, J., Guasch-Ferré, M., Toledo, E., Clish, C., Corella, D., Estruch, R., Ros, E., Fitó, M., Lee, C. H., Pierce, K., Razquin, C., Arós, F., Serra-Majem, L., Liang, L., Martínez-González, M. A., Hu, F. B., & Salas-Salvadó, J. (2021). Tricarboxylic acid cycle related-metabolites and risk of atrial fibrillation and heart failure. *Metabolism: Clinical and Experimental, 125*, 154915. https://doi.org/10.1016/j.metabol.2021.154915

Calder, P. C. (2013). Omega-3 polyunsaturated fatty acids and inflammatory processes: nutrition or pharmacology? *British Journal of Clinical Pharmacology, 75*(3), 645-662. https://doi.org/10.1111/j.1365-2125.2012.04374.x

Canepa, M., Kapelios, C. J., Benson, L., Savarese, G., & Lund, L. H. (2022). Temporal Trends of Heart Failure Hospitalizations in Cardiology Versus Noncardiology Wards According to Ejection Fraction: 16-Year Data From the SwedeHF Registry. *Circulation: Heart Failure, 15*(8), e009462. https://doi.org/10.1161/circheartfailure.121.009462

Caraballo, C., Desai, N. R., Mulder, H., Alhanti, B., Wilson, F. P., Fiuzat, M., Felker, G. M., Piña, I. L., O'Connor, C. M., Lindenfeld, J., Januzzi, J. L., Cohen, L. S., & Ahmad, T. (2019). Clinical Implications of the New York Heart Association Classification. *Journal of the American Heart Association, 8*(23), e014240. https://doi.org/10.1161/jaha.119.014240

Chang, P., Lei, S., Zhang, X., Zhang, J., Wang, X., Wu, J., Wang, J., Geng, J., Chen, B., & Yu, J. (2020). Metabonomics Analysis of Myocardial Metabolic Dysfunction in Patients with Cardiac Natriuretic Peptide Resistance. *Cardiology Research and Practice, 2020*. https://doi.org/10.1155/2020/1416945

Chen, W. S., Liu, M. H., Cheng, M. L., & Wang, C. H. (2020). Decreases in circulating concentrations of short-chain acylcarnitines are associated with systolic function improvement after decompensated heart failure. *International Heart Journal, 61*(5), 1014-1021. https://doi.org/10.1536/ihj.20-053

Chen, W. S., Wang, C. H., Cheng, C. W., Liu, M. H., Chu, C. M., Wu, H. P., Huang, P. C., Lin, Y. T., Ko, T., Chen, W. H., Wang, H. J., Lee, S. C., & Liang, C. Y. (2020). Elevated plasma phenylalanine predicts mortality in critical patients with heart failure. *ESC Heart Faillure, 7*(5), 2884-2893. https://doi.org/10.1002/ehf2.12896

Chen, Z., Lu, D., Qi, B., Wu, Y., Xia, Y., Chen, A., Li, S., Tang, H., Qian, J., & Ge, J. (2023). Quantitative Profiling of Serum Carnitines Facilitates the Etiology Diagnosis and Prognosis Prediction in Heart Failure. *Molecules, 28*(14), 5345. https://doi.org/10.3390/molecules28145345

Cheng, M. L., Wang, C. H., Shiao, M. S., Liu, M. H., Huang, Y. Y., Huang, C. Y., Mao, C. T., Lin, J. F., Ho, H. Y., & Yang, N. I. (2015). Metabolic disturbances identified in plasma are associated with outcomes in patients with heart failure: Diagnostic and prognostic value of metabolomics. *Journal of the American College of Cardiology, 65*(15), 1509-1520. https://doi.org/10.1016/j.jacc.2015.02.018

Chiesa, S. T., Charakida, M., Georgiopoulos, G., Roberts, J. D., Stafford, S. J., Park, C., Mykkänen, J., Kähönen, M., Lehtimäki, T., Ala-Korpela, M., Raitakari, O., Pietiäinen, M., Pussinen, P., Muthurangu, V., Hughes, A. D., Sattar, N., Timpson, N. J., & Deanfield, J. E. (2022). Glycoprotein Acetyls: A Novel Inflammatory Biomarker of Early Cardiovascular Risk in the Young. *Journal of the American Heart Association, 11*(4), e024380. https://doi.org/10.1161/jaha.121.024380

Chioncel, O., Lainscak, M., Seferovic, P. M., Anker, S. D., Crespo-Leiro, M. G., Harjola, V.-P., Parissis, J., Laroche, C., Piepoli, M. F., Fonseca, C., Mebazaa, A., Lund, L., Ambrosio, G. A., Coats, A. J., Ferrari, R., Ruschitzka, F., Maggioni, A. P., & Filippatos, G. (2017). Epidemiology and one-year outcomes in patients with chronic heart failure and preserved, mid-range and reduced ejection fraction: an analysis of the ESC Heart Failure Long-Term Registry. *European Journal of Heart Failure, 19*(12), 1574-1585. https://doi.org/10.1002/ejhf.813

Clark, K. A. A., Reinhardt, S. W., Chouairi, F., Miller, P. E., Kay, B., Fuery, M., Guha, A., Ahmad, T., & Desai, N. R. (2022). Trends in Heart Failure Hospitalizations in the US from 2008 to 2018. *Journal of Cardiac Failure, 28*(2), 171-180. https://doi.org/10.1016/j.cardfail.2021.08.020

Cluntun, A. A., Badolia, R., Lettlova, S., Parnell, K. M., Shankar, T. S., Diakos, N. A., Olson, K. A., Taleb, I., Tatum, S. M., Berg, J. A., Cunningham, C. N., Van Ry, T., Bott, A. J., Krokidi, A. T., Fogarty, S., Skedros, S., Swiatek, W. I., Yu, X., Luo, B., Merx, S., Navankasattusas, S., Cox, J. E., Ducker, G. S., Holland, W. L., McKellar, S. H., Rutter, J., & Drakos, S. G. (2021). The pyruvate-lactate axis modulates cardiac hypertrophy and heart failure. *Cell Metabolism, 33*(3), 629-648.e610. https://doi.org/10.1016/j.cmet.2020.12.003

Connelly, M. A., Otvos, J. D., Shalaurova, I., Playford, M. P., & Mehta, N. N. (2017). GlycA, a novel biomarker of systemic inflammation and cardiovascular disease risk. *Journal of Translational Medicine, 15*(1), 219. https://doi.org/10.1186/s12967-017-1321-6

Conners, K. M., Shearer, J. J., Joo, J., Park, H., Manemann, S. M., Remaley, A. T., Otvos, J. D., Connelly, M. A., Sampson, M., Bielinski, S. J., Wolska, A., Turecamo, S., & Roger, V. L. (2023). The Metabolic Vulnerability Index: A Novel Marker for Mortality Prediction in Heart Failure. *JACC: Heart Failure*. https://doi.org/10.1016/j.jchf.2023.06.013

Conrad, N., Judge, A., Canoy, D., Tran, J., Pinho-Gomes, A. C., Millett, E. R. C., Salimi-Khorshidi, G., Cleland, J. G., McMurray, J. J. V., & Rahimi, K. (2019). Temporal Trends and Patterns in Mortality After Incident Heart Failure: A Longitudinal Analysis of 86 000 Individuals. *JAMA Cardiology, 4*(11), 1102-1111. https://doi.org/10.1001/jamacardio.2019.3593

Cui, X., Ye, L., Li, J., Jin, L., Wang, W., Li, S., Bao, M., Wu, S., Li, L., Geng, B., Zhou, X., Zhang, J., & Cai, J. (2018). Metagenomic and metabolomic analyses unveil dysbiosis of gut microbiota in chronic heart failure patients. *Scientific Reports, 8*(1), 635. https://doi.org/10.1038/s41598-017-18756-2

Cunningham, J. W., Vaduganathan, M., Claggett, B. L., Zile, M. R., Anand, I. S., Packer, M., Zannad, F., Lam, C. S. P., Janssens, S., Jhund, P. S., Kober, L., Rouleau, J., Shah, S. J., Chopra, V. K., Shi, V. C., Lefkowitz, M. P., Prescott, M. F., Pfeffer, M. A., McMurray, J. J. V., & Solomon, S. D. (2020). Effects of Sacubitril/Valsartan on N-Terminal Pro-B-Type Natriuretic Peptide in Heart Failure With Preserved Ejection Fraction. *JACC: Heart Failure, 8*(5), 372-381. https://doi.org/10.1016/j.jchf.2020.03.002

Cuthbert, J. J., Gopal, J., Crundall-Goode, A., & Clark, A. L. (2019). Are there patients missing from community heart failure registers? An audit of clinical practice.

European Journal of Preventive Cardiology, 26(3), 291-298. https://doi.org/10.1177/2047487318810839

Czibik, G., Mezdari, Z., Altintas, D. M., Bréhat, J., Pini, M., d'Humières, T., Delmont, T., Radu, C., Breau, M., Liang, H., Martel, C., Abatan, A., Sarwar, R., Marion, O., Naushad, S., Zhang, Y., Halfaoui, M., Suffee, N., Morin, D., Adnot, S., Hatem, S., Yavari, A., Sawaki, D., & Derumeaux, G. (2021). Dysregulated Phenylalanine Catabolism Plays a Key Role in the Trajectory of Cardiac Aging. *Circulation, 144*(7), 559-574. https://doi.org/10.1161/CIRCULATIONAHA.121.054204

Daubert, M. A., Adams, K., Yow, E., Barnhart, H. X., Douglas, P. S., Rimmer, S., Norris, C., Cooper, L., Leifer, E., Desvigne-Nickens, P., Anstrom, K., Fiuzat, M., Ezekowitz, J., Mark, D. B., O'Connor, C. M., Januzzi, J., & Felker, G. M. (2019). NT-proBNP Goal Achievement Is Associated With Significant Reverse Remodeling and Improved Clinical Outcomes in HFrEF. *JACC: Heart Failure, 7*(2), 158-168. https://doi.org/10.1016/j.jchf.2018.10.014

Deidda, M., Piras, C., Dessalvi, C. C., Locci, E., Barberini, L., Torri, F., Ascedu, F., Atzori, L., & Mercuro, G. (2015). Metabolomic approach to profile functional and metabolic changes in heart failure. *Journal of Translational Medicine, 13*, 297. https://doi.org/10.1186/s12967-015-0661-3

Delles, C., Rankin, N. J., Boachie, C., McConnachie, A., Ford, I., Kangas, A., Soininen, P., Trompet, S., Mooijaart, S. P., Jukema, J. W., Zannad, F., Ala-Korpela, M., Salomaa, V., Havulinna, A. S., Welsh, P., Würtz, P., & Sattar, N. (2018). Nuclear magnetic resonance-based metabolomics identifies phenylalanine as a novel predictor of incident heart failure hospitalisation: results from PROSPER and FINRISK 1997. *European Journal of Heart Failure, 20*(4), 663-673. https://doi.org/10.1002/ejhf.1076

Deng, L. Y., Ji, R., Givens, R., Templeton, D. L., Wu, C., & Schulze, P. C. (2014). Increased circulating levels of total and very long-chain ceramides as markers of abnormal lipid metabolism and lipotoxicity in patients with chronic stable heart failure. *Circulation, 130*.

Dharmarajan, K., Hsieh, A. F., Kulkarni, V. T., Lin, Z., Ross, J. S., Horwitz, L. I., Kim, N., Suter, L. G., Lin, H., Normand, S. L., & Krumholz, H. M. (2015). Trajectories of risk after hospitalization for heart failure, acute myocardial infarction, or pneumonia: retrospective cohort study. *BMJ, 350*, h411. https://doi.org/10.1136/bmj.h411

Ding, M., Wennberg, A., Gigante, B., Walldius, G., Hammar, N., & Modig, K. (2022). Lipid levels in midlife and risk of atrial fibrillation over 3 decades—Experience from the Swedish AMORIS cohort: A cohort study. *PLoS Medicine, 19*(8), e1004044. https://doi.org/10.1371/journal.pmed.1004044

Ding, Y., Svingen, G. F., Pedersen, E. R., Gregory, J. F., Ueland, P. M., Tell, G. S., & Nygård, O. K. (2015). Plasma Glycine and Risk of Acute Myocardial Infarction in Patients With Suspected Stable Angina Pectoris. *Journal of the American Heart Association, 5*(1). https://doi.org/10.1161/jaha.115.002621

Du, Z., Shen, A., Huang, Y., Su, L., Lai, W., Wang, P., Xie, Z., Xie, Z., Zeng, Q., Ren, H., & Xu, D. (2014). 1H-NMR-Based Metabolic Analysis of Human Serum Reveals Novel Markers of Myocardial Energy Expenditure in Heart Failure Patients. *PloS One, 9*(2), e88102. https://doi.org/10.1371/journal.pone.0088102

Dunn, W. B., Broadhurst, D. I., Deepak, S. M., Buch, M. H., McDowell, G., Spasic, I., Ellis, D. I., Brooks, N., Kell, D. B., & Neyses, L. (2007). Serum metabolomics reveals many novel metabolic markers of heart failure, including pseudouridine and 2-oxoglutarate. *Metabolomics, 3*(4), 413-426. https://doi.org/10.1007/s11306-007-0063-5

Ebong, I. A., Goff, D. C., Rodriguez, C. J., Chen, H., Sibley, C. T., & Bertoni, A. G. (2013). Association of Lipids With Incident Heart Failure Among Adults With and Without Diabetes Mellitus. *Circulation: Heart Failure, 6*(3), 371-378. https://doi.org/10.1161/CIRCHEARTFAILURE.112.000093

Feldman, A. M., Kontos, C. D., McClung, J. M., Gerhard, G. S., Khalili, K., & Cheung, J. Y. (2017). Precision Medicine for Heart Failure: Lessons From Oncology. *Circulation: Heart Failure*, 10(6), e004202. https://doi.org/10.1161/CIRCHEARTFAILURE.117.004202

Felker, G. M., Anstrom, K. J., Adams, K. F., Ezekowitz, J. A., Fiuzat, M., Houston-Miller, N., Januzzi, J. L., Jr, Mark, D. B., Piña, I. L., Passmore, G., Whellan, D. J., Yang, H., Cooper, L. S., Leifer, E. S., Desvigne-Nickens, P., & O'Connor, C. M. (2017). Effect of Natriuretic Peptide–Guided Therapy on Hospitalization or Cardiovascular Mortality in High-Risk Patients With Heart Failure and Reduced Ejection Fraction: A Randomized Clinical Trial. *JAMA, 318*(8), 713-720. https://doi.org/10.1001/jama.2017.10565

Felker, G. M., Hasselblad, V., Hernandez, A. F., & O'Connor, C. M. (2009). Biomarker-guided therapy in chronic heart failure: a meta-analysis of randomized controlled trials. *American Heart Journal, 158*(3), 422-430. https://doi.org/10.1016/j.ahj.2009.06.018

Fernandez-Caggiano, M., & Eaton, P. (2021). Heart failure-emerging roles for the mitochondrial pyruvate carrier. *Cell Death and Differentiation, 28*(4), 1149-1158. https://doi.org/10.1038/s41418-020-00729-0

Friedewald, W. T., Levy, R. I., & Fredrickson, D. S. (1972). Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clinical Chemistry, 18*(6), 499-502.

Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., Collins, R., & Allen, N. E. (2017). Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *American Journal of Epidemiology, 186*(9), 1026-1034. https://doi.org/10.1093/aje/kwx246

Fry, D., Almond, R., Moffat, S., Gordon, M., & Singh, P. (2019). *Companion Document to Accompany Serum Biomarker Data*. Companion Document to Accompany Serum Biomarker Data

Goni, L., Razquin, C., Toledo, E., Guasch-Ferré, M., Clish, C. B., Babio, N., Wittenbecher, C., Atzeni, A., Li, J., Liang, L., Dennis, C., Alonso-Gómez, Á., Fitó, M.,

Corella, D., Gómez-Gracia, E., Estruch, R., Fiol, M., Lapetra, J., Serra-Majem, L., Ros, E., Arós, F., Salas-Salvadó, J., Hu, F. B., Martínez-González, M. A., & Ruiz-Canela, M. (2022). Arginine catabolism metabolites and atrial fibrillation or heart failure risk: 2 case-control studies within the Prevención con Dieta Mediterránea (PREDIMED) trial. *American Journal of Clinical Nutrition, 116*(3), 653-662. https://doi.org/10.1093/ajcn/nqac139

Gopal, D. M., Kalogeropoulos, A. P., Georgiopoulou, V. V., Tang, W. W., Methvin, A., Smith, A. L., Bauer, D. C., Newman, A. B., Kim, L., Harris, T. B., Kritchevsky, S. B., & Butler, J. (2010). Serum albumin concentration and heart failure risk The Health, Aging, and Body Composition Study. *American Heart Journal, 160*(2), 279-285. https://doi.org/10.1016/j.ahj.2010.05.022

Greene, S. J., Fonarow, G. C., Vaduganathan, M., Khan, S. S., Butler, J., & Gheorghiade, M. (2015). The vulnerable phase after hospitalization for heart failure. *Nature Reviews Cardiology, 12*(4), 220-229. https://doi.org/10.1038/nrcardio.2015.14

Guasch-Ferré, M., Hruby, A., Toledo, E., Clish, C. B., Martínez-González, M. A., Salas-Salvadó, J., & Hu, F. B. (2016). Metabolomics in Prediabetes and Diabetes: A Systematic Review and Meta-analysis. *Diabetes Care, 39*(5), 833-846. https://doi.org/10.2337/dc15-2251

Han, Y., Lu, J., Chen, B., Li, X., Dai, H., Zhang, L., Yan, X., Liu, J., Zhang, H., Fu, X., Yu, Q., Ren, J., Cui, H., Gao, Y., & Li, J. (2023). A novel polygenic risk score improves prognostic prediction of heart failure with preserved ejection fraction in the Chinese Han population. *European Journal of Preventive Cardiology*, 30(13), 1382–1390. https://doi.org/10.1093/eurjpc/zwad209

He, G.-H., Cai, W.-K., Meng, J.-R., Ma, X., Zhang, F., Lu, J., & Xu, G.-L. (2015). Relation of Polymorphism of the Histidine Decarboxylase Gene to Chronic Heart Failure in Han Chinese. *The American Journal of Cardiology, 115*(11), 1555-1562. https://doi.org/10.1016/j.amjcard.2015.02.062

Heidenreich, P. A., Bozkurt, B., Aguilar, D., Allen, L. A., Byun, J. J., Colvin, M. M., Deswal, A., Drazner, M. H., Dunlay, S. M., Evers, L. R., Fang, J. C., Fedson, S. E., Fonarow, G. C., Hayek, S. S., Hernandez, A. F., Khazanie, P., Kittleson, M. M., Lee, C. S., Link, M. S., Milano, C. A., Nnacheta, L. C., Sandhu, A. T., Stevenson, L. W., Vardeny, O., Vest, A. R., & Yancy, C. W. (2022). 2022 AHA/ACC/HFSA Guideline for the Management of Heart Failure: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation, 145*(18), e895-e1032. https://doi.org/10.1161/CIR.0000000000001063

Heydari, B., Abdullah, S., Pottala, J. V., Shah, R., Abbasi, S., Mandry, D., Francis, S. A., Lumish, H., Ghoshhajra, B. B., Hoffmann, U., Appelbaum, E., Feng, J. H., Blankstein, R., Steigner, M., McConnell, J. P., Harris, W., Antman, E. M., Jerosch-Herold, M., & Kwong, R. Y. (2016). Effect of Omega-3 Acid Ethyl Esters on Left Ventricular Remodeling After Acute Myocardial Infarction: The OMEGA-REMODEL Randomized Clinical Trial. *Circulation, 134*(5), 378-391. https://doi.org/10.1161/circulationaha.115.019949

Huelsmann, M., Neuhold, S., Resl, M., Strunk, G., Brath, H., Francesconi, C., Adlbrecht, C., Prager, R., Luger, A., Pacher, R., & Clodi, M. (2013). PONTIAC (NT-proBNP

Selected PreventiOn of cardiac eveNts in a populaTion of dIabetic patients without A history of Cardiac disease). *Journal of the American College of Cardiology, 62*(15), 1365-1372. https://doi.org/10.1016/j.jacc.2013.05.069

Hunter, W. G., Kelly, J. P., McGarrah, R. W., 3rd, Khouri, M. G., Craig, D., Haynes, C., Ilkayeva, O., Stevens, R. D., Bain, J. R., Muehlbauer, M. J., Newgard, C. B., Felker, G. M., Hernandez, A. F., Velazquez, E. J., Kraus, W. E., & Shah, S. H. (2016). Metabolomic Profiling Identifies Novel Circulating Biomarkers of Mitochondrial Dysfunction Differentially Elevated in Heart Failure With Preserved Versus Reduced Ejection Fraction: Evidence for Shared Metabolic Impairments in Clinical Heart Failure. *Journal of the American Heart Association, 5*(8). https://doi.org/10.1161/jaha.115.003190

Hunter, W. G., McGarrah, R. W., 3rd, Kelly, J. P., Khouri, M. G., Craig, D. M., Haynes, C., Felker, G. M., Hernandez, A. F., Velazquez, E. J., Kraus, W. E., & Shah, S. H. (2019). High-Density Lipoprotein Particle Subfractions in Heart Failure With Preserved or Reduced Ejection Fraction. *Journal of the American College of Cardiology, 73*(2), 177-186. https://doi.org/10.1016/j.jacc.2018.10.059

Ibrahim, N. E., & Januzzi, J. L. (2018). Established and Emerging Roles of Biomarkers in Heart Failure. *Circulation Research, 123*(5), 614-629. https://doi.org/10.1161/CIRCRESAHA.118.312706

Imamura, F., Lemaitre, R. N., King, I. B., Song, X., Steffen, L. M., Folsom, A. R., Siscovick, D. S., & Mozaffarian, D. (2013). Long-chain monounsaturated Fatty acids and incidence of congestive heart failure in 2 prospective cohorts. *Circulation, 127*(14), 1512-1521, 1521e1511-1518. https://doi.org/10.1161/circulationaha.112.001197

Inouye, M., Kettunen, J., Soininen, P., Silander, K., Ripatti, S., Kumpula, L. S., Hämäläinen, E., Jousilahti, P., Kangas, A. J., Männistö, S., Savolainen, M. J., Jula, A., Leiviskä, J., Palotie, A., Salomaa, V., Perola, M., Ala-Korpela, M., & Peltonen, L. (2010). Metabonomic, transcriptomic, and genomic variation of a population cohort. *Molecular Systems Biology, 6*, 441. https://doi.org/10.1038/msb.2010.93

Januzzi, J. L., Jr., Butler, J., Zannad, F., Filippatos, G., Ferreira, J. P., Pocock, S. J., Sattar, N., Verma, S., Vedin, O., Iwata, T., Brueckmann, M., Packer, M., & Anker, S. D. (2022). Prognostic Implications of N-Terminal Pro-B-Type Natriuretic Peptide and High-Sensitivity Cardiac Troponin T in EMPEROR-Preserved. *JACC: Heart Failure, 10*(7), 512-524. https://doi.org/10.1016/j.jchf.2022.05.004

Januzzi, J. L., Jr., Filippatos, G., Nieminen, M., & Gheorghiade, M. (2012). Troponin elevation in patients with heart failure: on behalf of the third Universal Definition of Myocardial Infarction Global Task Force: Heart Failure Section. *European Heart Journal, 33*(18), 2265-2271. https://doi.org/10.1093/eurheartj/ehs191

Jauhiainen, R., Jauhiainen, M., Vangipurapu, J., Kuulasmaa, T., Ala-Korpela, M., Laakso, M., & Kuusisto, J. (2021). Novel biomarkers associated with incident heart failure in 10 106 Finnish men. *ESC Heart Faillure, 8*(1), 605-614. https://doi.org/10.1002/ehf2.13132

Jones, N. R., Roalfe, A. K., Adoki, I., Hobbs, F. D. R., & Taylor, C. J. (2019). Survival of patients with chronic heart failure in the community: a systematic review and

meta-analysis. *European Journal of Heart Failure, 21*(11), 1306-1325. https://doi.org/10.1002/ejhf.1594

Joseph, P., Dokainish, H., McCready, T., Budaj, A., Roy, A., Ertl, G., Gomez-Mesa, J. E., Leong, D., Ezekowitz, J., Hage, C., Lanas, F., Maggioni, A. P., Sliwa, K., Zhu, J., Rouleau, J., Balasubramanian, K., & Yusuf, S. (2020). A multinational registry to study the characteristics and outcomes of heart failure patients: The global congestive heart failure (G-CHF) registry. *American Heart Journal, 227*, 56-63. https://doi.org/10.1016/j.ahj.2020.06.002

Julkunen, H., Cichońska, A., Tiainen, M., Koskela, H., Nybo, K., Mäkelä, V., Nokso-Koivisto, J., Kristiansson, K., Perola, M., Salomaa, V., Jousilahti, P., Lundqvist, A., Kangas, A. J., Soininen, P., Barrett, J. C., & Würtz, P. (2023). Atlas of plasma NMR biomarkers for health and disease in 118,461 individuals from the UK Biobank. *Nature Communications, 14*(1), 604. https://doi.org/10.1038/s41467-023-36231-7

Lahoz, R., Fagan, A., McSharry, M., Proudfoot, C., Corda, S., & Studer, R. (2020). Recurrent heart failure hospitalizations are associated with increased cardiovascular mortality in patients with heart failure in Clinical Practice Research Datalink. *ESC Heart Faillure, 7*(4), 1688-1699. https://doi.org/10.1002/ehf2.12727

Lawson, C. A., Zaccardi, F., Squire, I., Ling, S., Davies, M. J., Lam, C. S. P., Mamas, M. A., Khunti, K., & Kadam, U. T. (2019). 20-year trends in cause-specific heart failure outcomes by sex, socioeconomic status, and place of diagnosis: a population-based study. *The Lancet Public Health, 4*(8), e406-e420. https://doi.org/10.1016/S2468-2667(19)30108-2

Ledwidge, M., Gallagher, J., Conlon, C., Tallon, E., O'Connell, E., Dawkins, I., Watson, C., O'Hanlon, R., Bermingham, M., Patle, A., Badabhagni, M. R., Murtagh, G., Voon, V., Tilson, L., Barry, M., McDonald, L., Maurer, B., & McDonald, K. (2013). Natriuretic Peptide–Based Screening and Collaborative Care for Heart Failure: The STOP-HF Randomized Trial. *JAMA, 310*(1), 66-74. https://doi.org/10.1001/jama.2013.7588

Lee, D. S., Gona, P., Albano, I., Larson, M. G., Benjamin, E. J., Levy, D., Kannel, W. B., & Vasan, R. S. (2011). A systematic assessment of causes of death after heart failure onset in the community: impact of age at death, time period, and left ventricular systolic dysfunction. *Circulation: Heart Failure, 4*(1), 36-43. https://doi.org/10.1161/circheartfailure.110.957480

Lee, D. S. M., DePaolo, J. S., Aragam, K. G., Biddinger, K., Conery, M., Dilitikas, O., Hoffman-Andrews, L., Judy, R. L., Khan, A., Kulo, I., Puckelwartz, M. J., Reza, N., Satterfield, B. A., Singhal, P., Arany, Z. P., Cappola, T. P., Carruth, E., Day, S. M., Do, R., Haggarty, C. M., Joseph, J., McNally, E., Nadkarni, G., Owens, A. T., Rader, D. J., Ritchie, M. D., Sun, Y., Voight, B. F., Levin, M. G., & Damrauer, S. M. (2023). Common- and rare-variant genetic architecture of heart failure across the allele frequency spectrum. *medRxiv*. https://doi.org/10.1101/2023.07.16.23292724

Levey, A. S., Coresh, J., Greene, T., Stevens, L. A., Zhang, Y. L., Hendriksen, S., Kusek, J. W., & Van Lente, F. (2006). Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate.

Annals of Internal Medicine, 145(4), 247-254. https://doi.org/10.7326/0003-4819-145-4-200608150-00004

Levy, D., Kenchaiah, S., Larson, M. G., Benjamin, E. J., Kupka, M. J., Ho, K. K. L., Murabito, J. M., & Vasan, R. S. (2002). Long-Term Trends in the Incidence of and Survival with Heart Failure. *New England Journal of Medicine, 347*(18), 1397-1402. https://doi.org/10.1056/NEJMoa020265

Lim, L. L., Lau, E. S. H., Fung, E., Lee, H. M., Ma, R. C. W., Tam, C. H. T., Wong, W. K. K., Ng, A. C. W., Chow, E., Luk, A. O. Y., Jenkins, A., Chan, J. C. N., & Kong, A. P. S. (2020). Circulating branched-chain amino acids and incident heart failure in type 2 diabetes: The Hong Kong Diabetes Register. *Diabetes/Metabolism Research and Reviews, 36*(3), e3253. https://doi.org/10.1002/dmrr.3253

Lindahl, A., Forshed, J., & Nordström, A. (2016). Overlap in serum metabolic profiles between non-related diseases: Implications for LC-MS metabolomics biomarker discovery. *Biochemical and Biophysical Research Communications, 478*(3), 1472-1477. https://doi.org/10.1016/j.bbrc.2016.08.155

Liu, C., Li, R., Liu, Y., Li, Z., Sun, Y., Yin, P., & Huang, R. (2020). Characteristics of Blood Metabolic Profile in Coronary Heart Disease, Dilated Cardiomyopathy and Valvular Heart Disease Induced Heart Failure. *Frontiers in Cardiovascular Medicine, 7*, 622236. https://doi.org/10.3389/fcvm.2020.622236

Liu, H., Zhang, J., Li, Z., Liu, J., Lian, S., & Le, J. (2022). Association between remnant cholesterol and heart failure: A prospective cohort study. *Frontiers in Cardiovascular Medicine, 9*, 938647. https://doi.org/10.3389/fcvm.2022.938647

Liu, M., Chan, C. P., Yan, B. P., Zhang, Q., Lam, Y. Y., Li, R. J., Sanderson, J. E., Coats, A. J., Sun, J. P., Yip, G. W., & Yu, C. M. (2012). Albumin levels predict survival in patients with heart failure and preserved ejection fraction. *European Journal of Heart Failure, 14*(1), 39-44. https://doi.org/10.1093/eurjhf/hfr154

Lopaschuk, G. D., & Dyck, J. R. B. (2023). Ketones and the cardiovascular system. *Nature Cardiovascular Research, 2*(5), 425-437. https://doi.org/10.1038/s44161-023-00259-1

Lopaschuk, G. D., Karwi, Q. G., Tian, R., Wende, A. R., & Abel, E. D. (2021). Cardiac Energy Metabolism in Heart Failure. *Circulation Research, 128*(10), 1487-1513. https://doi.org/10.1161/CIRCRESAHA.121.318241

Lu, Y., Zhu, X., Li, J., Fang, R., Wang, Z., Zhang, J., Li, K., Li, X., Bai, H., Yang, Q., Ben, J., Zhang, H., & Chen, Q. (2017). Glycine prevents pressure overload induced cardiac hypertrophy mediated by glycine receptor. *Biochemical Pharmacology, 123*, 40-51. https://doi.org/10.1016/j.bcp.2016.11.008

Marcinkiewicz-Siemion, M., Ciborowski, M., Ptaszynska-Kopczynska, K., Szpakowicz, A., Lisowska, A., Jasiewicz, M., Waszkiewicz, E., Kretowski, A., Musial, W. J., & Kaminski, K. A. (2018). LC–MS-based serum fingerprinting reveals significant

dysregulation of phospholipids in chronic heart failure. *Journal of Pharmaceutical and Biomedical Analysis, 154*, 354-363. https://doi.org/10.1016/j.jpba.2018.03.027

Matsumoto, C., Hanson, N. Q., Tsai, M. Y., Glynn, R. J., Gaziano, J. M., & Djoussé, L. (2013). Plasma phospholipid saturated fatty acids and heart failure risk in the Physicians' Health Study. *Clinical Nutrition, 32*(5), 819-823. https://doi.org/10.1016/j.clnu.2013.02.006

Matsushita, K., Williams, E. K., Mongraw-Chaffin, M. L., Coresh, J., Schmidt, M. I., Brancati, F. L., Hoogeveen, R. C., Ballantyne, C. M., & Young, J. H. (2013). The association of plasma lactate with incident cardiovascular outcomes: the ARIC Study. *American Journal of Epidemiology, 178*(3), 401-409. https://doi.org/10.1093/aje/kwt002

McDonagh, T. A., Metra, M., Adamo, M., Gardner, R. S., Baumbach, A., Böhm, M., Burri, H., Butler, J., Čelutkienė, J., Chioncel, O., Cleland, J. G. F., Coats, A. J. S., Crespo-Leiro, M. G., Farmakis, D., Gilard, M., Heymans, S., Hoes, A. W., Jaarsma, T., Jankowska, E. A., Lainscak, M., Lam, C. S. P., Lyon, A. R., McMurray, J. J. V., Mebazaa, A., Mindham, R., Muneretto, C., Francesco Piepoli, M., Price, S., Rosano, G. M. C., Ruschitzka, F., Kathrine Skibelund, A., & ESC Scientific Document Group. (2021). 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: Developed by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) With the special contribution of the Heart Failure Association (HFA) of the ESC. *European Heart Journal, 42*(36), 3599-3726. https://doi.org/10.1093/eurheartj/ehab368

McDonagh, T. A., Metra, M., Adamo, M., Gardner, R. S., Baumbach, A., Böhm, M., Burri, H., Butler, J., Čelutkienė, J., Chioncel, O., Cleland, J. G. F., Crespo-Leiro, M. G., Farmakis, D., Gilard, M., Heymans, S., Hoes, A. W., Jaarsma, T., Jankowska, E. A., Lainscak, M., Lam, C. S. P., Lyon, A. R., McMurray, J. J. V., Mebazaa, A., Mindham, R., Muneretto, C., Francesco Piepoli, M., Price, S., Rosano, G. M. C., Ruschitzka, F., Skibelund, A. K., & ESC Scientific Document Group. (2023). 2023 Focused Update of the 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: Developed by the task force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) With the special contribution of the Heart Failure Association (HFA) of the ESC. *European Heart Journal*. https://doi.org/10.1093/eurheartj/ehad195

McGarrah, R. W., Crown, S. B., Zhang, G.-F., Shah, S. H., & Newgard, C. B. (2018). Cardiovascular Metabolomics. *Circulation Research, 122*(9), 1238-1258. https://doi.org/10.1161/CIRCRESAHA.117.311002

Meijers, W. C., Bayes-Genis, A., Mebazaa, A., Bauersachs, J., Cleland, J. G. F., Coats, A. J. S., Januzzi, J. L., Maisel, A. S., McDonald, K., Mueller, T., Richards, A. M., Seferovic, P., Mueller, C., & de Boer, R. A. (2021). Circulating heart failure biomarkers beyond natriuretic peptides: review from the Biomarker Study Group of the Heart Failure Association (HFA), European Society of Cardiology (ESC). *European Journal of Heart Failure, 23*(10), 1610-1632. https://doi.org/10.1002/ejhf.2346

Moissl, A. P., Lorkowski, S., Meinitzer, A., Pilz, S., Scharnagl, H., Delgado, G. E., Kleber, M. E., Krämer, B. K., Pieske, B., Grübler, M. R., Brussee, H., von Lewinski, D., Toplak, H., Fahrleitner-Pammer, A., März, W., & Tomaschitz, A. (2023). Association of

branched-chain amino acids with mortality-the Ludwigshafen Risk and Cardiovascular Health (LURIC) study. *iScience, 26*(4), 106459. https://doi.org/10.1016/j.isci.2023.106459

Morze, J., Wittenbecher, C., Schwingshackl, L., Danielewicz, A., Rynkiewicz, A., Hu, F. B., & Guasch-Ferré, M. (2022). Metabolomics and Type 2 Diabetes Risk: An Updated Systematic Review and Meta-analysis of Prospective Cohort Studies. *Diabetes Care, 45*(4), 1013-1024. https://doi.org/10.2337/dc21-1705

Mozaffarian, D., Lemaitre, R. N., King, I. B., Song, X., Spiegelman, D., Sacks, F. M., Rimm, E. B., & Siscovick, D. S. (2011). Circulating long-chain ω-3 fatty acids and incidence of congestive heart failure in older adults: the cardiovascular health study: a cohort study. *Annals of Internal Medicine, 155*(3), 160-170. https://doi.org/10.7326/0003-4819-155-3-201108020-00006

Mueller-Hennessen, M., Düngen, H. D., Lutz, M., Trippel, T. D., Kreuter, M., Sigl, J., Müller, O. J., Tahirovic, E., Witt, H., Ternes, P., Carvalho, S., Peter, E., Rein, D., Schatz, P., Herth, F., Giannitsis, E., Weis, T., Frey, N., & Katus, H. A. (2017). A novel lipid biomarker panel for the detection of heart failure with reduced ejection fraction. *Clinical Chemistry, 63*(1), 267-277. https://doi.org/10.1373/clinchem.2016.257279

Mueller-Hennessen, M., Sigl, J., Fuhrmann, J. C., Witt, H., Reszka, R., Schmitz, O., Kastler, J., Fischer, J. J., Müller, O. J., Giannitsis, E., Weis, T., Frey, N., & Katus, H. A. (2017). Metabolic profiles in heart failure due to non-ischemic cardiomyopathy at rest and under exercise. *ESC Heart Faillure, 4*(2), 178-189. https://doi.org/10.1002/ehf2.12133

Murashige, D., Jung, J. W., Neinast, M. D., Levin, M. G., Chu, Q., Lambert, J. P., Garbincius, J. F., Kim, B., Hoshino, A., Marti-Pamies, I., McDaid, K. S., Shewale, S. V., Flam, E., Yang, S., Roberts, E., Li, L., Morley, M. P., Bedi, K. C., Hyman, M. C., Frankel, D. S., Margulies, K. B., Assoian, R. K., Elrod, J. W., Jang, C., Rabinowitz, J. D., & Arany, Z. (2022). Extra-cardiac BCAA catabolism lowers blood pressure and protects from heart failure. *Cell Metabolism, 34*(11), 1749-1764.e1747. https://doi.org/10.1016/j.cmet.2022.09.008

Nadarajah, R., Younsi, T., Romer, E., Raveendra, K., Nakao, Y. M., Nakao, K., Shuweidhi, F., Hogg, D. C., Arbel, R., Zahger, D., Iakobishvili, Z., Fonarow, G. C., Petrie, M. C., Wu, J., & Gale, C. P. (2023, Jul 5). Prediction models for heart failure in the community: A systematic review and meta-analysis. Eur J Heart Fail. https://doi.org/10.1002/ejhf.2970

Nakamura, M., Odanovic, N., Nakada, Y., Dohi, S., Zhai, P., Ivessa, A., Yang, Z., Abdellatif, M., & Sadoshima, J. (2021). Dietary carbohydrates restriction inhibits the development of cardiac hypertrophy and heart failure. *Cardiovascular Research, 117*(11), 2365-2376. https://doi.org/10.1093/cvr/cvaa298

Nicholls, S. J., Lincoff, A. M., Garcia, M., Bash, D., Ballantyne, C. M., Barter, P. J., Davidson, M. H., Kastelein, J. J. P., Koenig, W., McGuire, D. K., Mozaffarian, D., Ridker, P. M., Ray, K. K., Katona, B. G., Himmelmann, A., Loss, L. E., Rensfeldt, M., Lundström, T., Agrawal, R., Menon, V., Wolski, K., & Nissen, S. E. (2020). Effect of High-Dose Omega-3 Fatty Acids vs Corn Oil on Major Adverse Cardiovascular Events in Patients at High

Cardiovascular Risk: The STRENGTH Randomized Clinical Trial. *JAMA, 324*(22), 2268-2280. https://doi.org/10.1001/jama.2020.22258

Nielsen, R., Møller, N., Gormsen, L. C., Tolbod, L. P., Hansson, N. H., Sorensen, J., Harms, H. J., Frøkiær, J., Eiskjaer, H., Jespersen, N. R., Mellemkjaer, S., Lassen, T. R., Pryds, K., Bøtker, H. E., & Wiggers, H. (2019). Cardiovascular Effects of Treatment With the Ketone Body 3-Hydroxybutyrate in Chronic Heart Failure Patients. *Circulation, 139*(18), 2129-2141. https://doi.org/10.1161/circulationaha.118.036459

Niezen, S., Connelly, M. A., Hirsch, C., Kizer, J. R., Benitez, M. E., Minchenberg, S., Perez-Matos, M. C., Jiang, Z. G., & Mukamal, K. J. (2023). Elevated Plasma Levels of Ketone Bodies Are Associated With All‐Cause Mortality and Incidence of Heart Failure in Older Adults: The CHS. *Journal of the American Heart Association, 12*(17), e029960. https://doi.org/10.1161/JAHA.123.029960

Nodari, S., Triggiani, M., Campia, U., Manerba, A., Milesi, G., Cesana, B. M., Gheorghiade, M., & Dei Cas, L. (2011). Effects of n-3 Polyunsaturated Fatty Acids on Left Ventricular Function and Functional Capacity in Patients With Dilated Cardiomyopathy. *Journal of the American College of Cardiology, 57*(7), 870-879. https://doi.org/10.1016/j.jacc.2010.11.017

Obokata, M., Reddy, Y. N. V., Melenovsky, V., Kane, G. C., Olson, T. P., Jarolim, P., & Borlaug, B. A. (2018). Myocardial Injury and Cardiac Reserve in Patients With Heart Failure and Preserved Ejection Fraction. *Journal of the American College of Cardiology, 72*(1), 29-40. https://doi.org/10.1016/j.jacc.2018.04.039

OECD European Union. (2022). *Health at a Glance: Europe 2022*. https://doi.org/10.1787/507433b0-en

Ohkuma, T., Jun, M., Woodward, M., Zoungas, S., Cooper, M. E., Grobbee, D. E., Hamet, P., Mancia, G., Williams, B., Welsh, P., Sattar, N., Shaw, J. E., Rahimi, K., & Chalmers, J. (2017). Cardiac Stress and Inflammatory Markers as Predictors of Heart Failure in Patients With Type 2 Diabetes: The ADVANCE Trial. *Diabetes Care, 40*(9), 1203-1209. https://doi.org/10.2337/dc17-0509

Otaki, Y., Watanabe, T., & Kubota, I. (2017). Heart-type fatty acid-binding protein in cardiovascular disease: A systemic review. *Clinica Chimica Acta, 474*, 44-53. https://doi.org/10.1016/j.cca.2017.09.007

Papadaki, A., Martínez-González, M., Alonso-Gómez, A., Rekondo, J., Salas-Salvadó, J., Corella, D., Ros, E., Fitó, M., Estruch, R., Lapetra, J., García-Rodriguez, A., Fiol, M., Serra-Majem, L., Pintó, X., Ruiz-Canela, M., Bulló, M., Serra-Mir, M., Sorlí, J. V., & Arós, F. (2017). Mediterranean diet and risk of heart failure: results from the PREDIMED randomized controlled trial. *European Journal of Heart Failure, 19*(9), 1179-1185. https://doi.org/10.1002/ejhf.750

Papandreou, C., Bulló, M., Hernández-Alonso, P., Ruiz-Canela, M., Li, J., Guasch-Ferré, M., Toledo, E., Clish, C., Corella, D., Estruch, R., Ros, E., Fitó, M., Alonso-Gómez, A., Fiol, M., Santos-Lozano, J. M., Serra-Majem, L., Liang, L., Martínez-González, M. A., Hu, F. B., & Salas-Salvadó, J. (2021). Choline Metabolism and Risk of Atrial Fibrillation and

Heart Failure in the PREDIMED Study. *Clinical Chemistry, 67*(1), 288-297. https://doi.org/10.1093/clinchem/hvaa224

Papandreou, C., Hernández-Alonso, P., Bulló, M., Ruiz-Canela, M., Li, J., Guasch-Ferré, M., Toledo, E., Clish, C., Corella, D., Estruch, R., Cofán, M., Fitó, M., Razquin, C., Arós, F., Fiol, M., Santos-Lozano, J. M., Serra-Majem, L., Liang, L., Martínez-González, M. A., Hu, F. B., & Salas-Salvadó, J. (2020). High Plasma Glutamate and a Low Glutamine-to-Glutamate Ratio Are Associated with Increased Risk of Heart Failure but Not Atrial Fibrillation in the Prevención con Dieta Mediterránea (PREDIMED) Study. *Journal of Nutrition, 150*(11), 2882-2889. https://doi.org/10.1093/jn/nxaa273

Petersen, S. E., Matthews, P. M., Bamberg, F., Bluemke, D. A., Francis, J. M., Friedrich, M. G., Leeson, P., Nagel, E., Plein, S., Rademakers, F. E., Young, A. A., Garratt, S., Peakman, T., Sellors, J., Collins, R., & Neubauer, S. (2013). Imaging in population science: cardiovascular magnetic resonance in 100,000 participants of UK Biobank - rationale, challenges and approaches. *Journal of Cardiovascular Magnetic Resonance, 15*(1), 46. https://doi.org/10.1186/1532-429x-15-46

Petersen, S. E., Matthews, P. M., Francis, J. M., Robson, M. D., Zemrak, F., Boubertakh, R., Young, A. A., Hudson, S., Weale, P., Garratt, S., Collins, R., Piechnik, S., & Neubauer, S. (2016). UK Biobank's cardiovascular magnetic resonance protocol. *Journal of Cardiovascular Magnetic Resonance, 18*, 8. https://doi.org/10.1186/s12968-016-0227-4

Petrone, A. B., Weir, N., Hanson, N. Q., Glynn, R., Tsai, M. Y., Gaziano, J. M., & Djoussé, L. (2013). Omega-6 fatty acids and risk of heart failure in the Physicians' Health Study. *American Journal of Clinical Nutrition, 97*(1), 66-71. https://doi.org/10.3945/ajcn.112.048991

Polidori, D., Iijima, H., Goda, M., Maruyama, N., Inagaki, N., & Crawford, P. A. (2018). Intra- and inter-subject variability for increases in serum ketone bodies in patients with type 2 diabetes treated with the sodium glucose co-transporter 2 inhibitor canagliflozin. *Diabetes, Obesity & Metabolism, 20*(5), 1321-1326. https://doi.org/10.1111/dom.13224

Ponikowski, P., Voors, A. A., Anker, S. D., Bueno, H., Cleland, J. G. F., Coats, A. J. S., Falk, V., González-Juanatey, J. R., Harjola, V.-P., Jankowska, E. A., Jessup, M., Linde, C., Nihoyannopoulos, P., Parissis, J. T., Pieske, B., Riley, J. P., Rosano, G. M. C., Ruilope, L. M., Ruschitzka, F., Rutten, F. H., van der Meer, P., & Group, E. S. D. (2016). 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC)Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *European Heart Journal, 37*(27), 2129-2200. https://doi.org/10.1093/eurheartj/ehw128

Rastogi, T., Ho, F. K., Rossignol, P., Merkling, T., Butler, J., Clark, A., Collier, T., Delles, C., Jukema, J. W., Heymans, S., Latini, R., Mebazaa, A., Pellicori, P., Sever, P., Staessen, J. A., Thijs, L., Cleland, J. G., Sattar, N., Zannad, F., & Girerd, N. (2022). Comparing and contrasting risk factors for heart failure in patients with and without history of myocardial infarction: data from HOMAGE and the UK Biobank. *European Journal of Heart Failure, 24*(6), 976-984. https://doi.org/10.1002/ejhf.2495

Razquin, C., Ruiz-Canela, M., Toledo, E., Hernández-Alonso, P., Clish, C. B., Guasch-Ferré, M., Li, J., Wittenbecher, C., Dennis, C., Alonso-Gómez, A., Fitó, M., Liang, L., Corella, D., Gómez-Gracia, E., Estruch, R., Fiol, M., Lapetra, J., Serra-Majem, L., Ros, E., Aros, F., Salas-Salvadó, J., Hu, F. B., & Martínez-González, M. A. (2021). Metabolomics of the tryptophan-kynurenine degradation pathway and risk of atrial fibrillation and heart failure: potential modification effect of Mediterranean diet. *American Journal of Clinical Nutrition, 114*(5), 1646-1654. https://doi.org/10.1093/ajcn/nqab238

Reina-Couto, M., Pereira-Terra, P., Quelhas-Santos, J., Silva-Pereira, C., Albino-Teixeira, A., & Sousa, T. (2021). Inflammation in Human Heart Failure: Major Mediators and Therapeutic Targets. *Frontiers in Physiology, 12*, 746494. https://doi.org/10.3389/fphys.2021.746494

Ritchie, S. C., Surendran, P., Karthikeyan, S., Lambert, S. A., Bolton, T., Pennells, L., Danesh, J., Di Angelantonio, E., Butterworth, A. S., & Inouye, M. (2023). Quality control and removal of technical variation of NMR metabolic biomarker data in ~120,000 UK Biobank participants. *Scientific Data, 10*(1), 64. https://doi.org/10.1038/s41597-023-01949-y

Ritterhoff, J., & Tian, R. (2023). Metabolic mechanisms in physiological and pathological cardiac hypertrophy: new paradigms and challenges. *Nature Reviews Cardiology*. https://doi.org/10.1038/s41569-023-00887-x

Roger, V. L. (2018). Heart Failure Epidemic. *Circulation, 138*(1), 25-28. https://doi.org/10.1161/CIRCULATIONAHA.118.028478

Ruiz, M., Labarthe, F., Fortier, A., Bouchard, B., Legault, J. T., Bolduc, V., Rigal, O., Chen, J., Ducharme, A., Crawford, P. A., Tardif, J. C., & Des Rosiers, C. (2017). Circulating acylcarnitine profile in human heart failure: A surrogate of fatty acid metabolic dysregulation in mitochondria and beyond. *American Journal of Physiology - Heart and Circulatory Physiology, 313*(4), 768-781. https://doi.org/10.1152/ajpheart.00820.2016

Ruiz-Canela, M., Guasch-Ferré, M., Razquin, C., Toledo, E., Hernández-Alonso, P., Clish, C. B., Li, J., Wittenbecher, C., Dennis, C., Alonso-Gómez, Á., Almanza-Aguilera, E., Liang, L., Corella, D., Gómez-Gracia, E., Estruch, R., Fiol, M., Lapetra, J., Serra-Majem, L., Ros, E., Arós, F., Salas-Salvadó, J., Hu, F. B., & Martínez-González, M. (2022). Plasma acylcarnitines and risk of incident heart failure and atrial fibrillation: the Prevención con dieta mediterránea study. *Revista Española de Cardiología, 75*(8), 649-658. https://doi.org/10.1016/j.rec.2021.10.005

Ruiz-Canela, M., Hruby, A., Clish, C. B., Liang, L., Martínez-González, M. A., & Hu, F. B. (2017). Comprehensive Metabolomic Profiling and Incident Cardiovascular Disease: A Systematic Review. *Journal of the American Heart Association, 6*(10). https://doi.org/10.1161/jaha.117.005705

Saleem, T. H., Algowhary, M., Kamel, F. E. M., & El-Mahdy, R. I. (2020). Plasma amino acid metabolomic pattern in heart failure patients with either preserved or reduced ejection fraction: The relation to established risk variables and prognosis. *Biomedical Chromatography*. https://doi.org/10.1002/bmc.5012

Sansbury, B. E., DeMartino, A. M., Xie, Z., Brooks, A. C., Brainard, R. E., Watson, L. J., DeFilippis, A. P., Cummins, T. D., Harbeson, M. A., Brittian, K. R., Prabhu, S. D., Bhatnagar, A., Jones, S. P., & Hill, B. G. (2014). Metabolomic analysis of pressure-overloaded and infarcted mouse hearts. *Circulation: Heart Failure, 7*(4), 634-642. https://doi.org/10.1161/circheartfailure.114.001151

Seferović, P. M., Polovina, M., Bauersachs, J., Arad, M., Ben Gal, T., Lund, L. H., Felix, S. B., Arbustini, E., Caforio, A. L. P., Farmakis, D., Filippatos, G. S., Gialafos, E., Kanjuh, V., Krljanac, G., Limongelli, G., Linhart, A., Lyon, A. R., Maksimović, R., Miličić, D., Milinković, I., Noutsias, M., Oto, A., Oto, Ö., Pavlović, S. U., Piepoli, M. F., Ristić, A. D., Rosano, G. M. C., Seggewiss, H., Ašanin, M., Seferović, J. P., Ruschitzka, F., Čelutkiene, J., Jaarsma, T., Mueller, C., Moura, B., Hill, L., Volterrani, M., Lopatin, Y., Metra, M., Backs, J., Mullens, W., Chioncel, O., de Boer, R. A., Anker, S., Rapezzi, C., Coats, A. J. S., & Tschöpe, C. (2019). Heart failure in cardiomyopathies: a position paper from the Heart Failure Association of the European Society of Cardiology. *European Journal of Heart Failure, 21*(5), 553-576. https://doi.org/10.1002/ejhf.1461

Seferović, P. M., Vardas, P., Jankowska, E. A., Maggioni, A. P., Timmis, A., Milinković, I., Polovina, M., Gale, C. P., Lund, L. H., Lopatin, Y., Lainscak, M., Savarese, G., Huculeci, R., Kazakiewicz, D., & Coats, A. J. S. (2021). The Heart Failure Association Atlas: Heart Failure Epidemiology and Management Statistics 2019. *European Journal of Heart Failure, 23*(6), 906-914. https://doi.org/10.1002/ejhf.2143

Sharma, A., Vaduganathan, M., Ferreira, J. P., Liu, Y., Bakris, G. L., Cannon, C. P., White, W. B., & Zannad, F. (2020). Clinical and Biomarker Predictors of Expanded Heart Failure Outcomes in Patients With Type 2 Diabetes Mellitus After a Recent Acute Coronary Syndrome: Insights From the EXAMINE Trial. *Journal of the American Heart Association, 9*(1), e012797. https://doi.org/10.1161/jaha.119.012797

Shemesh, E., Chevli, P. A., Islam, T., German, C. A., Otvos, J., Yeboah, J., Rodriguez, F., deFilippi, C., Lima, J. A. C., Blaha, M., Pandey, A., Vaduganathan, M., & Shapiro, M. D. (2023). Circulating ketone bodies and cardiovascular outcomes: the MESA study. *European Heart Journal, 44*(18), 1636-1646. https://doi.org/10.1093/eurheartj/ehad087

Shimokawa, H., Miura, M., Nochioka, K., & Sakata, Y. (2015). Heart failure as a general pandemic in Asia. *European Journal of Heart Failure, 17*(9), 884-892. https://doi.org/10.1002/ejhf.319

Siddesha, J. M., Valente, A. J., Yoshida, T., Sakamuri, S. S., Delafontaine, P., Iba, H., Noda, M., & Chandrasekar, B. (2014). Docosahexaenoic acid reverses angiotensin II-induced RECK suppression and cardiac fibroblast migration. *Cellular Signalling, 26*(5), 933-941. https://doi.org/10.1016/j.cellsig.2014.01.005

Soininen, P., Kangas, A. J., Würtz, P., Suna, T., & Ala-Korpela, M. (2015). Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. *Circulation: Cardiovascular Genetics, 8*(1), 192-206. https://doi.org/10.1161/circgenetics.114.000216

Soininen, P., Kangas, A. J., Würtz, P., Tukiainen, T., Tynkkynen, T., Laatikainen, R., Järvelin, M. R., Kähönen, M., Lehtimäki, T., Viikari, J., Raitakari, O. T., Savolainen, M. J., & Ala-Korpela, M. (2009). High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. *Analyst, 134*(9), 1781-1785. https://doi.org/10.1039/b910205a

Stefani, G. P., Capalonga, L., da Silva, L. R., & Dal Lago, P. (2020). β-Alanine and l-histidine supplementation associated with combined training increased functional capacity and maximum strength in heart failure rats. *Experimental Physiology, 105*(5), 831-841. https://doi.org/10.1113/ep088327

Steffen, B. T., Duprez, D., Szklo, M., Guan, W., & Tsai, M. Y. (2018). Circulating oleic acid levels are related to greater risks of cardiovascular events and all-cause mortality: The Multi-Ethnic Study of Atherosclerosis. *Journal of Clinical Lipidology, 12*(6), 1404-1412. https://doi.org/10.1016/j.jacl.2018.08.004

Stenemo, M., Ganna, A., Salihovic, S., Nowak, C., Sundström, J., Giedraitis, V., Broeckling, C. D., Prenni, J. E., Svensson, P., Magnusson, P. K. E., Lind, L., Ingelsson, E., Ärnlöv, J., & Fall, T. (2019). The metabolites urobilin and sphingomyelin (30:1) are associated with incident heart failure in the general population. *ESC Heart Faillure, 6*(4), 764-773. https://doi.org/10.1002/ehf2.12453

Stolfo, D., Uijl, A., Vedin, O., Strömberg, A., Faxén, U. L., Rosano, G. M. C., Sinagra, G., Dahlström, U., & Savarese, G. (2019). Sex-Based Differences in Heart Failure Across the Ejection Fraction Spectrum: Phenotyping, and Prognostic and Therapeutic Implications. *JACC: Heart Failure, 7*(6), 505-515. https://doi.org/10.1016/j.jchf.2019.03.011

Stryeck, S., Gastrager, M., Degoricija, V., Trbušić, M., Potočnjak, I., Radulović, B., Pregartner, G., Berghold, A., Madl, T., & Frank, S. (2019). Serum Concentrations of Citrate, Tyrosine, 2- and 3- Hydroxybutyrate are Associated with Increased 3-Month Mortality in Acute Heart Failure Patients. *Scientific Reports, 9*(1), 6743. https://doi.org/10.1038/s41598-019-42937-w

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., & Collins, R. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Medicine, 12*(3), e1001779. https://doi.org/10.1371/journal.pmed.1001779

Suinesiaputra, A., Sanghvi, M. M., Aung, N., Paiva, J. M., Zemrak, F., Fung, K., Lukaschuk, E., Lee, A. M., Carapella, V., Kim, Y. J., Francis, J., Piechnik, S. K., Neubauer, S., Greiser, A., Jolly, M. P., Hayes, C., Young, A. A., & Petersen, S. E. (2018). Fully-automated left ventricular mass and volume MRI analysis in the UK Biobank population cohort: evaluation of initial results. *International Journal of Cardiovascular Imaging, 34*(2), 281-291. https://doi.org/10.1007/s10554-017-1225-9

Sun, H., Olson, K. C., Gao, C., Prosdocimo, D. A., Zhou, M., Wang, Z., Jeyaraj, D., Youn, J.-Y., Ren, S., Liu, Y., Rau, C. D., Shah, S., Ilkayeva, O., Gui, W.-J., William, N. S., Wynn, R. M., Newgard, C. B., Cai, H., Xiao, X., Chuang, D. T., Schulze, P. C., Lynch, C., Jain, M. K., & Wang, Y. (2016). Catabolic Defect of Branched-Chain Amino Acids Promotes

Heart Failure. *Circulation, 133*(21), 2038-2049. https://doi.org/10.1161/CIRCULATIONAHA.115.020226

Tahir, U. A., Katz, D. H., Zhao, T., Ngo, D., Cruz, D. E., Robbins, J. M., Chen, Z.-Z., Peterson, B., Benson, M. D., Shi, X., Dailey, L., Andersson, C., Vasan, R. S., Gao, Y., Shen, C., Correa, A., Hall, M. E., Wang, T. J., Clish, C. B., Wilson, J. G., & Gerszten, R. E. (2021). Metabolomic Profiles and Heart Failure Risk in Black Adults: Insights From the Jackson Heart Study. *Circulation: Heart Failure, 14*(1), e007275. https://doi.org/10.1161/CIRCHEARTFAILURE.120.007275

Tanaka, M., Koyama, T., Sakurai, T., Kamiyoshi, A., Ichikawa-Shindo, Y., Kawate, H., Liu, T., Xian, X., Imai, A., Zhai, L., Hirabayashi, K., Owa, S., Yamauchi, A., Igarashi, K., Taniguchi, S., & Shindo, T. (2016). The endothelial adrenomedullin-RAMP2 system regulates vascular integrity and suppresses tumour metastasis. *Cardiovascular Research, 111*(4), 398-409. https://doi.org/10.1093/cvr/cvw166

Taylor, C. J., Ordóñez-Mena, J. M., Roalfe, A. K., Lay-Flurrie, S., Jones, N. R., Marshall, T., & Hobbs, F. D. R. (2019). Trends in survival after a diagnosis of heart failure in the United Kingdom 2000-2017: population based cohort study. *BMJ, 364*, l223. https://doi.org/10.1136/bmj.l223

Tenori, L., Hu, X., Pantaleo, P., Alterini, B., Castelli, G., Olivotto, I., Bertini, I., Luchinat, C., & Gensini, G. F. (2013). Metabolomic fingerprint of heart failure in humans: a nuclear magnetic resonance spectroscopy analysis. *International Journal of Cardiology, 168*(4), e113-115. https://doi.org/10.1016/j.ijcard.2013.08.042

Thomas, H. M. (2015). The Role of Echocardiography in Heart Failure. *Journal of Nuclear Medicine, 56*(Supplement 4), 31S. https://doi.org/10.2967/jnumed.114.150433

Tierney, A., Fry, D., Almond, R., Gordon, M., & Moffat, S. (2018). *Companion Document to Accompany HbA1c Biomarker Data*. https://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=2405

Toth, P. P., Philip, S., Hull, M., & Granowitz, C. (2019). Elevated Triglycerides (≥150 mg/dL) and High Triglycerides (200-499 mg/dL) Are Significant Predictors of New Heart Failure Diagnosis: A Real-World Analysis of High-Risk Statin-Treated Patients. *Vascular Health and Risk Management, 15*, 533-538. https://doi.org/10.2147/vhrm.S221289

Tromp, J., Beusekamp, J. C., Ouwerkerk, W., van der Meer, P., Cleland, J. G. F., Angermann, C. E., Dahlstrom, U., Ertl, G., Hassanein, M., Perrone, S. V., Ghadanfar, M., Schweizer, A., Obergfell, A., Filippatos, G., Dickstein, K., Collins, S. P., & Lam, C. S. P. (2022). Regional differences in precipitating factors of hospitalization for acute heart failure: insights from the REPORT-HF registry. *European Journal of Heart Failure, 24*(4), 645-652. https://doi.org/10.1002/ejhf.2431

Tsao, C. W., Aday, A. W., Almarzooq, Z. I., Anderson, C. A. M., Arora, P., Avery, C. L., Baker-Smith, C. M., Beaton, A. Z., Boehme, A. K., Buxton, A. E., Commodore-Mensah, Y., Elkind, M. S. V., Evenson, K. R., Eze-Nliam, C., Fugar, S., Generoso, G., Heard, D. G., Hiremath, S., Ho, J. E., Kalani, R., Kazi, D. S., Ko, D., Levine, D. A., Liu, J., Ma, J., Magnani, J. W., Michos, E. D., Mussolino, M. E., Navaneethan, S. D., Parikh, N. I., Poudel,

R., Rezk-Hanna, M., Roth, G. A., Shah, N. S., St-Onge, M.-P., Thacker, E. L., Virani, S. S., Voeks, J. H., Wang, N.-Y., Wong, N. D., Wong, S. S., Yaffe, K., & Martin, S. S. (2023). Heart Disease and Stroke Statistics—2023 Update: A Report From the American Heart Association. *Circulation, 147*(8), e93-e621. https://doi.org/10.1161/CIR.0000000000001123

UK Biobank. (2007). *UK Biobank: Protocol for a large-scale prospective epidemiological resource*. https://www.ukbiobank.ac.uk/media/gnkeyh2q/study-rationale.pdf

UK Biobank. (2015). *Imaging modality Cardiovascular Magnetic Resonance (CMR)*. https://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=349

UK Biobank. (2023a). *Hospital inpatient data*. https://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=138483

UK Biobank. (2023b). *Mortality data: linkage to death registries*. https://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=115559

Varbo, A., & Nordestgaard, B. G. (2018). Nonfasting Triglycerides, Low-Density Lipoprotein Cholesterol, and Heart Failure Risk. *Arteriosclerosis, Thrombosis, and Vascular Biology, 38*(2), 464-472. https://doi.org/10.1161/ATVBAHA.117.310269

Velavan, P., Huan Loh, P., Clark, A., & Cleland, J. G. F. (2007). The Cholesterol Paradox in Heart Failure. *Congestive Heart Failure, 13*(6), 336-341. https://doi.org/10.1111/j.1527-5299.2007.07211.x

Wang, C. H., Cheng, M. L., & Liu, M. H. (2018). Amino acid-based metabolic panel provides robust prognostic value additive to b-natriuretic peptide and traditional risk factors in heart failure. *Disease Markers, 2018*. https://doi.org/10.1155/2018/3784589

Wang, W., Zhang, F., Xia, Y., Zhao, S., Yan, W., Wang, H., Lee, Y., Li, C., Zhang, L., Lian, K., Gao, E., Cheng, H., & Tao, L. (2016). Defective branched chain amino acid catabolism contributes to cardiac dysfunction and remodeling following myocardial infarction. *American Journal of Physiology: Heart and Circulatory Physiology, 311*(5), H1160-h1169. https://doi.org/10.1152/ajpheart.00114.2016

Weis, E. M., Puchalska, P., Nelson, A. B., Taylor, J., Moll, I., Hasan, S. S., Dewenter, M., Hagenmüller, M., Fleming, T., Poschet, G., Hotz-Wagenblatt, A., Backs, J., Crawford, P. A., & Fischer, A. (2022). Ketone body oxidation increases cardiac endothelial cell proliferation. *EMBO Molecular Medicine, 14*(4), e14753. https://doi.org/10.15252/emmm.202114753

Wilk, J. B., Tsai, M. Y., Hanson, N. Q., Gaziano, J. M., & Djoussé, L. (2012). Plasma and dietary omega-3 fatty acids, fish intake, and heart failure risk in the Physicians' Health Study. *American Journal of Clinical Nutrition, 96*(4), 882-888. https://doi.org/10.3945/ajcn.112.042671

Willeit, P., Kaptoge, S., Welsh, P., Butterworth, A. S., Chowdhury, R., Spackman, S. A., Pennells, L., Gao, P., Burgess, S., Freitag, D. F., Sweeting, M., Wood, A. M., Cook, N. R., Judd, S., Trompet, S., Nambi, V., Olsen, M. H., Everett, B. M., Kee, F., Ärnlöv, J., Salomaa, V., Levy, D., Kauhanen, J., Laukkanen, J. A., Kavousi, M., Ninomiya, T., Casas, J.-

P., Daniels, L. B., Lind, L., Kistorp, C. N., Rosenberg, J., Mueller, T., Rubattu, S., Panagiotakos, D. B., Franco, O. H., de Lemos, J. A., Luchner, A., Kizer, J. R., Kiechl, S., Salonen, J. T., Wannamethee, S. G., de Boer, R. A., Nordestgaard, B. G., Andersson, J., Jørgensen, T., Melander, O., Ballantyne, C. M., DeFilippi, C., Ridker, P. M., Cushman, M., Rosamond, W. D., Thompson, S. G., Gudnason, V., Sattar, N., Danesh, J., & Di Angelantonio, E. (2016). Natriuretic peptides and integrated risk assessment for cardiovascular disease: an individual-participant-data meta-analysis. *The Lancet Diabetes & Endocrinology, 4*(10), 840-849. https://doi.org/10.1016/S2213-8587(16)30196-6

Wishart, D. S., Guo, A., Oler, E., Wang, F., Anjum, A., Peters, H., Dizon, R., Sayeeda, Z., Tian, S., Lee, B. L., Berjanskii, M., Mah, R., Yamamoto, M., Jovel, J., Torres-Calzada, C., Hiebert-Giesbrecht, M., Lui, V. W., Varshavi, D., Varshavi, D., Allen, D., Arndt, D., Khetarpal, N., Sivakumaran, A., Harford, K., Sanford, S., Yee, K., Cao, X., Budinski, Z., Liigand, J., Zhang, L., Zheng, J., Mandal, R., Karu, N., Dambrova, M., Schiöth, H. B., Greiner, R., & Gautam, V. (2022). HMDB 5.0: the Human Metabolome Database for 2022. *Nucleic Acids Research, 50*(D1), D622-d631. https://doi.org/10.1093/nar/gkab1062

Wittemans, L. B. L., Lotta, L. A., Oliver-Williams, C., Stewart, I. D., Surendran, P., Karthikeyan, S., Day, F. R., Koulman, A., Imamura, F., Zeng, L., Erdmann, J., Schunkert, H., Khaw, K. T., Griffin, J. L., Forouhi, N. G., Scott, R. A., Wood, A. M., Burgess, S., Howson, J. M. M., Danesh, J., Wareham, N. J., Butterworth, A. S., & Langenberg, C. (2019). Assessing the causal association of glycine with risk of cardio-metabolic diseases. *Nature Communications, 10*(1), 1060. https://doi.org/10.1038/s41467-019-08936-1

Wittenbecher, C., Eichelmann, F., Toledo, E., Guasch-Ferré, M., Ruiz-Canela, M., Li, J., Arós, F., Lee, C. H., Liang, L., Salas-Salvadó, J., Clish, C. B., Schulze, M. B., Martínez-González, M., & Hu, F. B. (2021). Lipid Profiles and Heart Failure Risk: Results From Two Prospective Studies. *Circulation Research, 128*(3), 309-320. https://doi.org/10.1161/circresaha.120.317883

Wollert, K. C., Kempf, T., & Wallentin, L. (2017). Growth Differentiation Factor 15 as a Biomarker in Cardiovascular Disease. *Clinical Chemistry, 63*(1), 140-151. https://doi.org/10.1373/clinchem.2016.255174

Yamagata, K., Suzuki, S., & Tagami, M. (2016). Docosahexaenoic acid prevented tumor necrosis factor alpha-induced endothelial dysfunction and senescence. *Prostaglandins, Leukotrienes, and Essential Fatty Acids, 104*, 11-18. https://doi.org/10.1016/j.plefa.2015.10.006

Yan, T., Zhu, S., Yin, X., Xie, C., Xue, J., Zhu, M., Weng, F., Zhu, S., Xiang, B., Zhou, X., Liu, G., Ming, Y., Zhu, K., Wang, C., & Guo, C. (2023). Burden, Trends, and Inequalities of Heart Failure Globally, 1990 to 2019: A Secondary Analysis Based on the Global Burden of Disease 2019 Study. *Journal of the American Heart Association, 12*(6), e027852. https://doi.org/10.1161/JAHA.122.027852

Yang, S., Hu, Y., Zhao, J., Jing, R., Wang, J., Gu, M., Niu, H., Chen, L., & Hua, W. (2021). Comprehensive plasma metabolites profiling reveals phosphatidylcholine species as potential predictors for cardiac resynchronization therapy response. *ESC Heart Failure, 8*(1), 280-290. https://doi.org/10.1002/ehf2.13037

York, M. K., Gupta, D. K., Reynolds, C. F., Farber-Eger, E., Wells, Q. S., Bachmann, K. N., Xu, M., Harrell, F. E., & Wang, T. J. (2018). B-Type Natriuretic Peptide Levels and Mortality in Patients With and Without Heart Failure. *Journal of the American College of Cardiology, 71*(19), 2079-2088. https://doi.org/10.1016/j.jacc.2018.02.071

Young, K. A., Scott, C. G., Rodeheffer, R. J., & Chen, H. H. (2021). Progression of Preclinical Heart Failure: A Description of Stage A and B Heart Failure in a Community Population. *Circulation: Cardiovascular Quality and Outcomes, 14*(5), e007216. https://doi.org/10.1161/CIRCOUTCOMES.120.007216

Yurista, S. R., Matsuura, T. R., Silljé, H. H. W., Nijholt, K. T., McDaid, K. S., Shewale, S. V., Leone, T. C., Newman, J. C., Verdin, E., van Veldhuisen, D. J., de Boer, R. A., Kelly, D. P., & Westenbrink, B. D. (2021). Ketone Ester Treatment Improves Cardiac Function and Reduces Pathologic Remodeling in Preclinical Models of Heart Failure. *Circulation: Heart Failure, 14*(1), e007684. https://doi.org/10.1161/circheartfailure.120.007684

Yurista, S. R., Silljé, H. H. W., Oberdorf-Maass, S. U., Schouten, E. M., Pavez Giani, M. G., Hillebrands, J. L., van Goor, H., van Veldhuisen, D. J., de Boer, R. A., & Westenbrink, B. D. (2019). Sodium-glucose co-transporter 2 inhibition with empagliflozin improves cardiac function in non-diabetic rats with left ventricular dysfunction after myocardial infarction. *European Journal of Heart Failure, 21*(7), 862-873. https://doi.org/10.1002/ejhf.1473

Zhang, Y., Pletcher, M. J., Vittinghoff, E., Clemons, A. M., Jacobs, D. R., Jr., Allen, N. B., Alonso, A., Bellows, B. K., Oelsner, E. C., Zeki Al Hazzouri, A., Kazi, D. S., de Ferranti, S. D., & Moran, A. E. (2021). Association Between Cumulative Low-Density Lipoprotein Cholesterol Exposure During Young Adulthood and Middle Age and Risk of Cardiovascular Events. *JAMA Cardiology, 6*(12), 1406-1413. https://doi.org/10.1001/jamacardio.2021.3508

Zhao, H., Shui, B., Zhao, Q., Hu, Z., Shu, Q., Su, M., Zhang, Y., & Ni, Y. (2021). Quantitative Metabolomics Reveals Heart Failure With Midrange Ejection Fraction as a Distinct Phenotype of Heart Failure. *Canadian Journal of Cardiology, 37*(2), 300-309. https://doi.org/10.1016/j.cjca.2020.03.024

Zhao, J., Yang, S., Jing, R., Jin, H., Hu, Y., Wang, J., Gu, M., Niu, H., Zhang, S., Chen, L., & Hua, W. (2020). Plasma Metabolomic Profiles Differentiate Patients With Dilated Cardiomyopathy and Ischemic Cardiomyopathy. *Frontiers in Cardiovascular Medicine, 7*, 597546. https://doi.org/10.3389/fcvm.2020.597546

Zhao, S., Kusminski, C. M., & Scherer, P. E. (2021). Adiponectin, Leptin and Cardiovascular Disorders. *Circulation Research, 128*(1), 136-149. https://doi.org/10.1161/CIRCRESAHA.120.314458

Zhong, Y., Wang, R., Yan, L., Lin, M., Liu, X., & You, T. (2017). Copeptin in heart failure: Review and meta-analysis. *Clinica Chimica Acta, 475*, 36-43. https://doi.org/10.1016/j.cca.2017.10.001

Zhuang, X.-d., Zhang, S.-z., Liao, L.-z., Lin, X., Zhou, H.-m., Zhong, X.-b., Sun, X.-t., Du, Z.-m., Zhang, M.-f., & Liao, X.-x. (2020). Serum Albumin and Incident Heart Failure. *Circulation: Genomic and Precision Medicine, 13*(6), e002989. https://doi.org/10.1161/CIRCGEN.120.002989

Ziaeian, B., Kominski, G. F., Ong, M. K., Mays, V. M., Brook, R. H., & Fonarow, G. C. (2017). National Differences in Trends for Heart Failure Hospitalizations by Sex and Race/Ethnicity. *Circulation: Cardiovascular Quality and Outcomes, 10*(7). https://doi.org/10.1161/circoutcomes.116.003552

Zile, M. R., Desai, A. S., Agarwal, R., Bharmi, R., Dalal, N., Adamson, P. B., & Maisel, A. S. (2020). Prognostic value of brain natriuretic peptide vs history of heart failure hospitalization in a large real-world population. *Clinical Cardiology, 43*(12), 1501-1510. https://doi.org/10.1002/clc.23468

Zoanni, B., Brioschi, M., Mallia, A., Gianazza, E., Eligini, S., Carini, M., Aldini, G., & Banfi, C. (2023). Novel insights about albumin in cardiovascular diseases: Focus on heart failure. *Mass Spectrometry Reviews, 42*(4), 1113-1128. https://doi.org/10.1002/mas.21743

Zordoky, B. N., Sung, M. M., Ezekowitz, J., al, R., Han, B., Bjorndahl, T. C., Bouatra, S., Anderson, T., Oudit, G. Y., Wishart, D. S., & Dyck, J. R. B. (2015). Metabolomic fingerprint of heart failure with preserved ejection fraction. *PloS One, 10*(5). https://doi.org/10.1371/journal.pone.0124844

Zymliński, R., Biegus, J., Sokolski, M., Siwołowski, P., Nawrocka-Millward, S., Todd, J., Jankowska, E. A., Banasiak, W., Cotter, G., Cleland, J. G., & Ponikowski, P. (2018). Increased blood lactate is prevalent and identifies poor prognosis in patients with acute heart failure without overt peripheral hypoperfusion. *European Journal of Heart Failure, 20*(6), 1011-1018. https://doi.org/10.1002/ejhf.1156