10-2020

Incremental prognostic value of a novel metabolite-based biomarker score in congestive heart failure patients

Peter McGranaghan
Baptist Health South Florida, PMcGranaghan@baptisthealth.net

Anshul Saxena
Baptist Health South Florida, anshuls@baptisthealth.net

Muni Rubens
Miami Cancer Institute, MuniR@baptisthealth.net

Joseph Salami
Baptist Health South Florida, josephsa@baptisthealth.net

Emir Veledar
Baptist Health South Florida, emirv@baptisthealth.net

Follow this and additional works at: https://scholarlycommons.baptisthealth.net/se-all-publications

ORIGINAL RESEARCH ARTICLE

Incremental prognostic value of a novel metabolite-based biomarker score in congestive heart failure patients

Peter McGranaghan[1,2] (iD), Hans-Dirk Düngen[1,3], Anshul Saxena[2], Muni Rubens[2], Joseph Salami[2], Jasmin Radenkovic[1,3], Doris Bach[1,3], Svetlana Apostolovic[4], Goran Loncar[5,6], Marija Zdravkovic[6,7], Elvis Tahirovic[1,3], Jovan Veskovic[1], Stefan Störk[8], Emir Veledar[2], Burkert Pieske[1,3,9], Frank Edelmann[1,3,9] and Tobias Daniel Trippel[1,3]* (iD)

[1]*Department of Internal Medicine and Cardiology, Charité Campus Virchow-Klinikum, Berlin, Germany;* [2]*Baptist Health South Florida, Coral Gables, FL, USA;* [3]*DZHK (German Centre for Cardiovascular Research), partner site Berlin, Augustenburger Platz 1, Berlin, 13353, Germany;* [4]*Clinic for Cardiovascular Diseases, Clinical Centre Niš, University of Niš, Niš, Serbia;* [5]*Department of Cardiology, Institute for Cardiovascular Diseases Dedinje, Belgrade, Serbia;* [6]*Faculty of Medicine, University of Belgrade, Belgrade, Serbia;* [7]*Department of Cardiology, University Hospital Medical Center Bezanijska Kosa, Belgrade, Serbia;* [8]*Comprehensive Heart Failure Center Würzburg, University and University Hospital Würzburg, Würzburg, Germany;* [9]*Berlin Institute of Health, Berlin, Germany*

Abstract

Aims The Cardiac Lipid Panel (CLP) is a newly discovered panel of metabolite-based biomarkers that has shown to improve the diagnostic value of N terminal pro B type natriuretic peptide (NT-proBNP). However, little is known about its usefulness in predicting outcomes. In this study, we developed a risk score for 4-year cardiovascular death in elderly chronic heart failure (CHF) patients using the CLP.

Methods and results From the Cardiac Insufficiency Bisoprolol Study in Elderly trial, we included 280 patients with CHF aged >65 years. A targeted metabolomic analysis of the CLP biomarkers was performed on baseline serum samples. Cox regression was used to determine the association of the biomarkers with the outcome after accounting for established risk factors. A risk score ranging from 0 to 4 was calculated by counting the number of biomarkers above the cut-offs, using Youden index. During the mean (standard deviation) follow-up period of 50 (8) months, 35 (18%) subjects met the primary endpoint of cardiovascular death. The area under the receiver operating curve for the model based on clinical variables was 0.84, the second model with NT-proBNP was 0.86, and the final model with the CLP was 0.90. The categorical net reclassification index was 0.25 using three risk categories: 0–60% (low), 60–85% (intermediate), and >85% (high). The continuous net reclassification index was 0.772, and the integrated discrimination index was 0.104.

Conclusions In patients with CHF, incorporating a panel of three metabolite-based biomarkers into a risk score improved the prognostic utility of NT-proBNP by predicting long-term cardiovascular death more precisely. This novel approach holds promise to improve clinical risk assessment in CHF patients.

Introduction

Chronic heart failure (CHF) is a leading cause of morbidity and mortality. Its prevalence continues to rise in developed countries, partly because of a shift in the age distribution of the population and improved treatment and care.[1,2] Clinicians should ensure that patients with CHF have the necessary knowledge and resources to make the best health decisions. Accurate and improved decision support methods, such as tools to predict the risk of mortality and prognosis of

patients, could help in making such shared decisions for treatment plans and risk management strategies. Recently, there has been an increase in the number of prognostic biomarkers being tested for CHF such as growth differentiation factor-15,[3,4] high-sensitivity C-reactive protein,[5,6] galectin-3,[7–9] and high-sensitivity troponin T.[10,11] However, the added value of these markers is still under debate, and long-term follow-up studies are lacking. Among biomarkers widely used in CHF, N terminal pro B type natriuretic peptide (NT-proBNP) is recognized as a standard reference for diagnosis and prognosis. Despite the clinical utility of NT-proBNP, some studies have reported a high intraindividual variance and high reference change values among patients with CHF.[12–14]

Metabolomic profiling, or metabolomics, can help meet the need for more robust prognostic biomarkers. This approach provides a holistic signature of biochemical activities in humans by detecting and quantifying low-weight molecules ($<$1500 Da) that could be associated with disease progression.[15–17] Studies of predictive metabolomic biomarkers in CHF have been published previously that support the overall hypothesis that circulating metabolites may be used for risk assessment of cardiovascular (CV) disease patients[18-20,21-29]. These studies appear promising, but validation and the additive value of these biomarkers are less established.

In a discovery phase untargeted metabolomic study by Mueller et al., comparing CHF patients to healthy controls, a novel panel of metabolites known as the Cardiac Lipid Panel (CLP) was found to improve the diagnostic performance over NT-proBNP alone.[30] Its prognostic performance, however, is unknown. Details of the CLP have been published previously.[30] In brief, the CLP is a biomarker panel consisting of three specific metabolomic features: triacylglycerol (TAG) 18:1/18:0/18:0, phosphatidylcholine (PC) 16:0/18:2, and the sum of the three isobaric sphingomyelin (SM) species SM d18:1/23:1, SM d18:2/23:0, and SM d17:1/24:1.

In this study, the prognostic value of the CLP was analysed in elderly patients with CHF. We developed a risk score for predicting 4-year CV death using cut-offs for the CLP, which improved predictive value of the standard reference biomarker, and traditional risk factors.

Materials and methods

Study population

The study sample was randomly selected from the Cardiac Insufficiency Bisoprolol Study in Elderly (CIBIS-ELD) trial, a multicentre, randomized, double-blind trial with ≥65-year-old patients with stable CHF. Details of the CIBIS-ELD trial have been published previously.[31,32] In brief, elderly patients with CHF were randomized in a 1:1 fashion to receive either bisoprolol or carvedilol, up titrated every 2 weeks for

12 weeks and then followed for 4 years. We only considered the baseline and 4-year follow-up time points for this study. From the 589 subjects with available blood samplesfrom CIBIS-ELD trial, patients were randomly selected and studied in a case cohort design. Following random selection, the cohort was filtered down based on the feasibility of performing the biomarker test, for instance if there was sufficient quantity of blood aliquot sample available for analysis, and whether blood samples passed quality assurance[33], resulting in a final set of 280 cases. The investigation conformed to the principles outlined in the Declaration of Helsinki.[34] The ethics committees of all participating centres approved the study protocol, and informed consent was signed by all participants prior to study participation.

Metabolite profiling

The serum samples were collected in 2006–2007 at the time of the CIBIS-ELD study initiation, stored at −80°C, and then shipped on dry ice in 2014 to the metabolomics lab for analysis. Metabolite profiling of the serum samples was performed using a kit developed for the routine measurement of the CLP. The dedicated protocol was designed for utilization in the clinical practice setting and based on a one-phase extraction of the samples using gas chromatography mass spectrometry (GC-MS), followed by liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis as previously described.[30] Sample and metabolite analysis quality assurance is part of the analytical protocol, so the metabolomic data that did not pass quality assurance were not included in this study[33]. The three CLP metabolomic features were generated at baseline, only for the previously mentioned samples.

Statistical analysis

Continuous variables were expressed as mean (standard deviation) and compared using *t*-test or Mann–Whitney *U* test, according to normal or non-normal distribution. Categorical variables were expressed as number (%). Comparisons among variables with more than two categories were performed using Wilcoxon rank sum test for continuous variables and Pearson's χ^2 test (or Fisher's exact test) or Mantel–Haenszel χ^2 test for categorical and ordinal data, respectively. All continuous predictor variables were log transformed to allow for direct comparison. Survival time was calculated from baseline until CV death or censoring at 4-year follow-up. Univariate Cox regression was performed on the CLP components and NT-proBNP, and multivariate Cox regression was also performed to adjust for clinical covariates. The considered clinical covariates were age, sex, body mass index, New York Heart Association class, creatinine, LDL cholesterol, triglycerides, left ventricular ejection fraction (LVEF), history of

diabetes, history of myocardial infarction, smoking history, hypertension, hyperlipidaemia, coronary artery disease, and medication including beta-blockers, aldosterone receptor blockers, angiotensin-converting enzyme inhibitors, anti-arrhythmic agents, aspirin, calcium channel blockers, diuretics, glycosides, nitrates, statins, sedative agents, and vitamin K antagonists. Hazard ratios and 95% confidence intervals were calculated for each univariate and multivariate model. The survival function for each model was generated using the predicted risk estimates following Cox regression.

To evaluate the predictive value of the CLP, three multivariable prediction models were built using Cox regression. The first model was built using the baseline clinical covariates only (Model A). Then, NT-proBNP was added to the first model (Model B). Finally, the CLP risk score was added to Model B (Model C). The CLP risk score was calculated as the sum of biomarkers above the Youden index cut-off.[35] There were four cut-off values, because four biomarkers are included in the score, three from the CLP and one from NT-proBNP. Each cut-off was calculated using Youden's index of the predicted probability from the Cox multivariate regression. Supplemental Data (*Data S1*) shows the equation for calculating the Youden cut-off. Based on the Youden cut-off, a value of 1 or 0 was assigned if the biomarker was above/below the cut-off value. A value of 1 was assigned in the direction of higher risk, that is, if a biomarker was protective (hazard ratio < 1), then a 1 value was assigned if the biomarker was below the Youden cut-off and vice versa. Then, all four values were summed to generate the final score for each subject. To measure the discrimination of each model, the area under the receiver operating curve (AUROC) and Harrell's concordance statistics were calculated for the 4-year survival of Models A, B, and C. Differences in Uno's concordance statistics were calculated for hypothesis testing of the change in AUROC of the three models.[36]

To measure risk reclassification, both continuous and categorical net reclassification indexes (NRIs) were calculated as well as integrated discrimination improvement (IDI).[37,38] The categorical NRI used three categories of <60%, 60–85%, and >85% corresponding to low, intermediate, and high risk, respectively. The continuous NRI does not depend on the choice of categories, but allocates any change in predicted risk in the correct direction.[39] IDI measures the ability of the new model to increase average sensitivity without reducing average specificity.

For sensitivity analysis, we performed logistic regression in addition to the Cox regression analysis, using the same independent and dependent variables in order to assess whether a different statistical model would yield similar results. We also tested two additional outcomes: the first was major adverse CV events defined as either myocardial infarction, transient ischaemic attack, stroke, or CV mortality, and the second outcome was all-cause mortality. Comparison of receiver operating curves following logistic regression was done

using the Mann–Whitney *U* test. To test the sensitivity of NRI variation in risk categories, we used the same number of risk categories (*n* = 3) but readjusted the cut-off values using two separate sets of cut-offs, which still corresponded to high, medium, and low. The first set was 70% and 90% followed by the second set of 80% and 95%. Statistical analysis was performed using SAS software version 9.4 and R software version 3.6.1.[40–42]

Results

Baseline characteristics

Figure 1 shows the study rationale and selection of subjects for this subcohort. The previously discovered CLP metabolites found to improve diagnosis of CHF were studied to assess their prognostic value. *Table 1* shows the baseline characteristics of the subsample population (*n* = 280) with a comparison to the source CIBIS-ELD cohort (*n* = 589). Mean patient age was 72.1 (4.9) years, 73.6% were men, 45% patients had heart failure with reduced ejection fraction (LVEF < 35%), 49% had heart failure with mid-range ejection fraction (LVEF 35–49%), 4% had heart failure with preserved ejection fraction (LVEF ≥ 50%), and the majority of patients were in New York Heart Association functional class II (67.5%). During the follow-up period (mean = 50 months, standard deviation = 8; median = 46 months), 35 (13%) died from CV causes.

Prognostic performance and risk reclassification

Table 2 shows the univariate (unadjusted) and multivariate (adjusted) models of the CLP risk score components. In the unadjusted model, two of the three CLP biomarkers (PC and SM) were significantly associated as well as NT-proBNP. In the adjusted model, the same two CLP biomarkers remained significant but NT-proBNP did not. Supporting Information, *Table S1* shows the the hazard ratios for the clinical variables included in the adjusted model. *Figure 2* shows the measures of discrimination (AUROC) for the three multivariable models with a comparison to its preceding model to test the level of significance after adding the respective covariate(s), and Supporting Information, *Table S2* shows Harrell's concordance statistics. The AUROC for Model A was 0.84, that of Model B was 0.86, and the final adjusted Model C was 0.90. The difference in AUROC after adding the CLP score (Model B vs. C) was significant (*P* = 0.02), whereas the difference after adding NT-proBNP to the clinical model (Model A vs. B) was insignificant (*P* = 0.47).

Figure 3 shows the number of subjects reclassified into each risk category for Model A to B to C. *Table 2* shows the risk reclassification of Models B and C as percentages of total

FIGURE 1 Study rationale for the prognostic biomarker study. AUC, area under the curve; CHF, chronic heart failure; CIBIS-ELD, Cardiac Insufficiency Bisoprolol Study in Elderly; CLP, Cardiac Lipid Panel; IDI, integrated discrimination improvement; NRI, net reclassification index; NT-proBNP, N terminal pro B type natriuretic peptide.
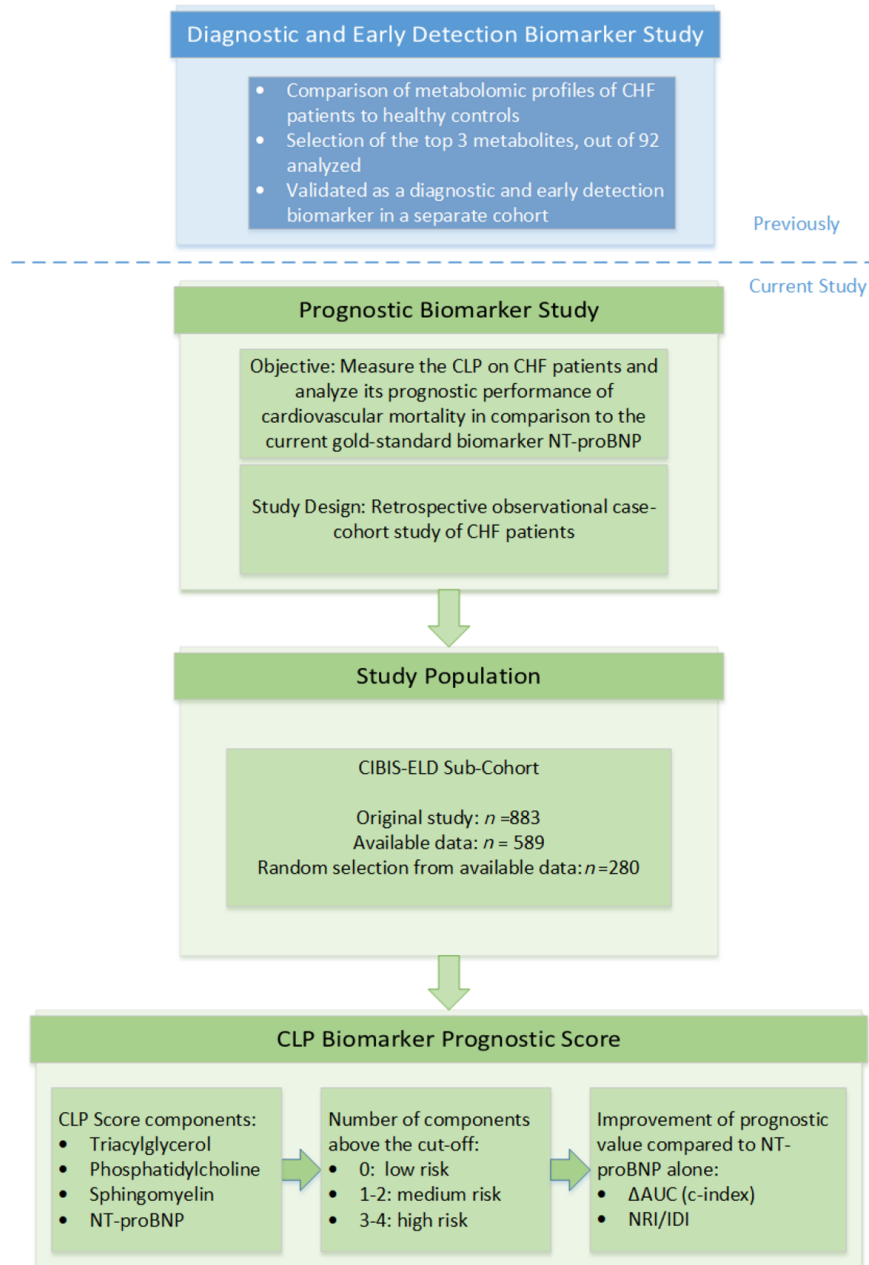
with events and non-events. Supporting Information, *Table S1a–b* shows the frequency of cases per risk category stratified by events and non-events from Models B and C. The overall categorical NRI was 0.25 using the three risk categories 0–60%, 60–85% and >85%, meaning 25% of the subjects were reclassified into the respective correct risk category after adding the CLP (*Table 3*). Accordingly, 59% of the reclassified cases were downgraded, and the other 41% were upgraded. Among patients experiencing events, the overall categorical NRI was 0.60, with 33% of those downgraded and 67% upgraded. For non-events, the categorical NRI was 0.19, with 70% of those downgraded and 30% upgraded. The overall continuous NRI was 0.472 and the IDI was 0.019. The CLP model (Model C) showed that its high-risk

Table 1 Baseline characteristics of the study participants compared with the source cohort

Characteristic	$n = 280$	$n = 589$	P value
Age (years), mean ± SD	72 ± 4.9	72 ± 4.9	0.4190[a]
NYHA (II/III), n	188/91	374/183	0.5424[c]
Male, n (%)	206 (74)	412 (71)	0.1389[b]
Body mass index (kg/m²), mean ± SD	26.8 ± 3.4	26.9 ± 3.9	0.4296[a]
Heart rate (bpm), mean ± SD	73 ± 13	74.7 ± 14	0.0031[a]
Systolic blood pressure (mm Hg), mean ± SD	134 ± 19	134 ± 19	0.2490[a]
Diastolic blood pressure (mm Hg), mean ± SD	81 ± 11	81 ± 11	0.3402[a]
Laboratory, mean ± SD			
Serum creatinine (μmol/l)	106 ± 29	107 ± 43	0.0096[a]
Haemoglobin (g/dL)	24.4 ± 34.8	14 ± 2	0.0325[a]
Sodium (mmol/L)	141.4 ± 3.3	141 ± 6.9	0.0765[a]
Uric acid (μmol/L)	273.2 ± 196.4	343 ± 121	0.0218[a]
Cholesterol (mmol/L)	5.5 ± 1.4	5.5 ± 1.4	0.2743[a]
HDL cholesterol (mmol/L)	1.2 ± 0.5	1.2 ± 0.5	0.4051[a]
LDL cholesterol (mmol/L)	3.4 ± 1.3	3.4 ± 1.2	0.348[a]
Triglycerides (mmol/L)	1.7 ± 1.0	1.8 ± 1.1	0.0283[a]
NT-proBNP (pg/mL)	793 (331–1765)[d]	873 (350–1931)[d]	0.0485[a]
Cardiac imaging, mean ± SD			
LVEF (%)	36 ± 9.5	37 ± 9.6	0.0899[a]
LVDed (mm)	58.8 ± 9.2	59.8 ± 9.3	0.0082[a]
LVDes (mm)	45.5 ± 9.7	46.5 ± 10.2	0.0089[a]
LVVed (mL)	152.7 ± 63.9	159 ± 67.7	0.0344[a]
LVVes (mL)	101.1 ± 51.6	105 ± 54.1	0.0705[a]
LAes (mm)	45.3 ± 7.2	45.2 ± 7.2	0.453[a]
E/e'	8 ± 4.3	11.1 ± 8.5	0.0025[a]
E/A	1 ± 0.8	1.1 ± 0.9	0.2928[a]
Deceleration time (ms)	226 ± 80	225 ± 79	0.7198[a]
Comorbidities, n (%)			
Diabetes	82 (29)	146 (25)	0.023[b]
Hypertension	224 (80)	469 (80)	0.7941[b]
Coronary artery disease	200 (71)	392 (67)	0.0382[b]
Smokers	125 (45)	257 (44)	0.7933[b]
Hyperlipidaemia	162 (58)	343 (59)	0.6822[b]
Medication, n (%)			
ACE inhibitor	247 (88)	509 (87)	0.527[b]
ARB	115 (41)	240 (41)	0.9643[b]
Glycoside	59 (21)	101 (17)	0.0216[b]
Aspirin	216 (77)	433 (74)	0.1273[b]
Nitrate	146 (52)	253 (43)	0.001[b]
Anti-arrhythmic agent	42 (15)	88 (15)	0.9512[b]
Statin	114 (41)	231 (40)	0.6044[b]

P values are compared with the available 589 subjects from the CIBIS-ELD cohort, which included this cohort of 280 subjects.
ACE, angiotensin-converting enzyme; ARB, angiotensin receptor blocker; E/A, ratio of the early (E) to late (A) ventricular filling velocities; E/e', ratio between early mitral inflow velocity and mitral annular early diastolic velocity; HDL, high-density lipoprotein; LDL, low-density lipoprotein; NT-proBNP, N-terminal pro–B-type natriuretic peptide; NYHA, New York Heart Association; LAes, left atrial end systole; LVEF, left ventricular ejection fraction; LVDed, left ventricular diameter end diastole; LVDes, left ventricular diameter end systole; LVVed, left ventricular volume end diastole; LVVes, left ventricular volume end systole.
[a] Wilcoxon rank sum test.
[b] Pearson's χ^2 test.
[c] Mantel–Haenszel χ^2.
[d] Median (interquartile range).

category contained predominantly subjects who experienced an event (77%), whereas the respective fraction in the NT-proBNP model (Model B) was only 42%.

Results were consistent in the sensitivity analysis using logistic regression. We found that the differences in AUROC values pointed in the same direction as the Cox regression models (Model C AUROC = 0.90 vs. Model B AUROC = 0.86, P = 0.02). The change in AUROC after adding NT-proBNP to the clinical model (Model A AUROC = 0.84) remained insignificant (P = 0.47; Supporting Information, *Figure S1*). We found similar results when testing the models using the two
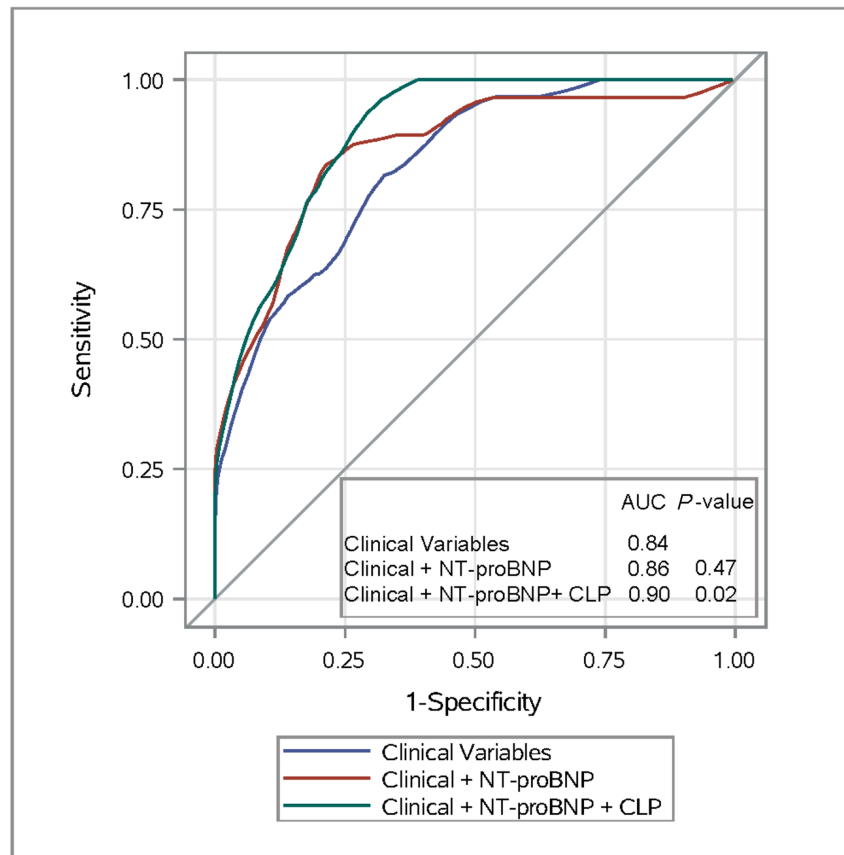
additional outcomes, major adverse CV event and all-cause mortality (Supporting Information, *Figures S2* and *S3*). Readjustment of risk categories using the two different sets of cut-offs also showed similar results with the original set of cut-off values. The overall NRI for the first set of 70% and 90% cut-offs was 0.28 (Supporting Information, *Table S6*), and the second set of 80% and 90% cut-offs was 0.36 (Supplemental Information, *Table S7*). In addition, Model C was still able to classify a higher proportion of cases with events in the high-risk group than Model B in each scenario (Supporting Information, *Tables S3–SS5*).

Table 2 CLP risk score components and HRs

CLP component	Unadjusted HR (95% CI)	P value	Adjusted HR (95% CI)	P value
SM	0.36 (0.16–0.82)	0.0143	0.18 (0.04–0.76)	0.0039
PC	0.76 (0.64–0.89)	0.0007	0.53 (0.38–0.75)	0.0003
TAG	0.69 (0.47–1.02)	0.0644	0.67 (0.35–1.25)	0.2069
NT-proBNP	1.49 (1.12–1.99)	0.007	1.60 (0.975–2.625)	0.0630

Adjusted Cox proportional hazard model considers the following clinical covariates: age, sex, body mass index, New York Heart Association class, creatinine, LDL cholesterol, triglycerides, left ventricular ejection fraction, history of diabetes, history of myocardial infarction, smoking history, hypertension, hyperlipidaemia, coronary artery disease, beta-blockers, aldosterone receptor blockers, angiotensin-converting enzyme inhibitors, anti-arrhythmic agents, aspirin, calcium channel blockers, diuretics, glycosides, nitrates, statins, sedative agents, and vitamin K antagonists.
CI, confidence interval; CLP, Cardiac Lipid Panel; HR, hazard ratio; NT-proBNP, N terminal pro B-type natriuretic peptide; SM, sum of the 3 isobaric sphingomyelin species: SM d18:1/23:1, SM d18:2/23:0, and SM d17:1/24:1; PC, phosphatidylcholine 16:0/18:2; TAG, triacylglycerol 18:1/18:0/18:0.

FIGURE 2 Discrimination analysis of the CLP biomarker risk score for 4-year cardiovascular mortality. AUC, area under the curve; CLP, Cardiac Lipid Panel; NT-proBNP, N terminal pro B type natriuretic peptide.

	AUC	P-value
Clinical Variables	0.84	
Clinical + NT-proBNP	0.86	0.47
Clinical + NT-proBNP+ CLP	0.90	0.02

Discussion

In this post hoc analysis of the CIBIS-ELD trial analysing 280 elderly patients with CHF, we showed that a risk score based on a novel panel of metabolites added prognostic value for the prediction of long-term CV mortality. A previous study already had reported that the CLP may improve the early detection and diagnosis of CHF.[30] However, to the best of our knowledge, the current study is first to estimate the prognostic performance of the CLP. Risk prediction models allow clinicians to accurately assess patient prognosis and facilitate more effective risk stratification and, ideally, a personalized treatment. Devising a more accurate biomarker panel for CHF risk prediction may aid clinicians with the difficult

FIGURE 3 Risk reclassification of subjects after adding N terminal pro B type natriuretic peptide to the clinical model followed by adding the Cardiac Lipid Panel biomarker score.
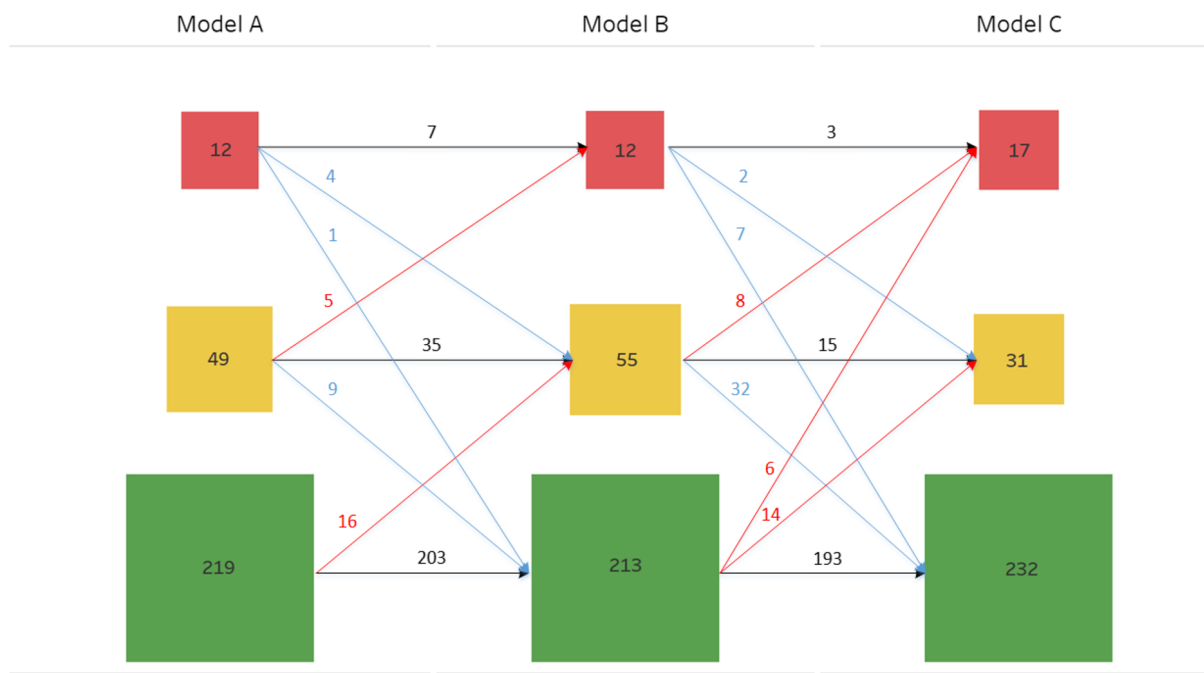

Table 3 Risk reclassification of total subjects, cases, and non-cases after adding the CLP risk score to the NT-proBNP based model

		Model B			
	Risk category	Low	Medium	High	Total
Model C	Low	69%	12%	2%	83%
		96%, 4%	84%, 16%	71%, 29%	
	Medium	5%	5%	1%	11%
		71%, 29%	80%, 20%	100%, 0%	
	High	2%	3%	1%	6%
		33%, 67%	25%, 75%	0%, 100%	
	Total	76%	20%	4%	100%

Percentage of subjects within each risk category of each Model A and B only. Events and non-events are proportions of the group total and are comma separated with red denoting events and blue denoting non-events.
Model B is the clinical covariates + NT-proBNP, Model C is clinical covariates + NT-proBNP + CLP score. Total subjects, $n = 280$; total events, $n = 35$.
The considered clinical covariates were age, sex, body mass index, New York Heart Association class, creatinine, LDL cholesterol, triglycerides, left ventricular ejection fraction, history of diabetes, history of myocardial infarction, smoking history, hypertension, hyperlipidaemia, coronary artery disease, and medication including beta-blockers, aldosterone receptor blockers, angiotensin-converting enzyme inhibitors, anti-arrhythmic agents, aspirin, calcium channel blockers, diuretics, glycosides, nitrates, statins, sedative agents, and vitamin K antagonists.
Categorical net reclassification index was calculated according to risk cut-offs of <60%, 60–85%, and >85% corresponding to risk categories low, medium, and high, respectively.

decisions surrounding the management of such high-risk patients. Conversely, identifying patients at lower risk may help reassure both clinicians and patients.

In the current study, two out of the three CLP components as well as NT-proBNP were independently associated with the outcome, so our next step was to build a risk score using these four components. The CLP risk score showed improved discrimination and risk reclassification in comparison with NT-proBNP alone, which is the current reference standard. Adding NT-proBNP to the clinical model slightly but insignificantly improved discrimination, while adding the CLP yielded a significant change in AUROC. Risk reclassification was improved by adding the CLP as it correctly identified a higher proportion of high-risk patients experiencing an event. For non-events, the majority of reclassified cases (70%) were downgraded. This indicates that added information of CLP also aided the proper classification of low-risk patients. Although both Models B and C misclassified some patients who did not experience any event in the high-risk group, the model with CLP had higher specificity as it classified

35% more patients with an event into the high-risk group. The continuous NRI also showed that Model C, compared with Model B, produced higher (i.e. more accurate) risk estimates for patients experiencing an event and lower risk estimates for those who are not.

Application of a single biomarker such as NT-proBNP for outcome prediction is primarily limited by insufficient specificity, resulting in a high false positive rate or low positive predictive value.[43,44] Because NT-proBNP is really a marker of elevated atrial pressures and volume overload, it may be of limited use in well-compensated, clinically stable heart failure patients. Hence, supporting this marker at the metabolic level may provide additional prognostic value and potentially phenotypic information.A combination of several metabolomic features into a biomarker panel or a risk score may provide a better prognosis utility over single biomarkers. A systematic review[21] reported that 6 out of 12 articles[22–27] developed a score by combining between 4 and up to 16 metabolites to predict CV risk. Recently, Lanfear et al. identified and then validated a panel of 13 circulating metabolites as a predictor of mortality risk in HF patients after accounting for conventional clinical risk factors and NT-proBNP levels.[28] Another prospective population-based study deriving a risk score from four metabolites and validating this score in two cohorts found improved risk reclassification of CHF patients using the biomarker score, although discrimination was not significantly enhanced.[29] A meta-analysis of 18 metabolomic prediction studies of CV disease outcomes reported an average change in c-statistic of 0.0417 (standard error 0.008) after adding metabolite-based information, which is consistent with our results. Of note, the metabolite score subgroup performed best (n = 5 studies),[45] although publication bias and heterogeneity were reported regarding variations in cohorts, study design, and metabolite profiling approaches.

In addition to investigating the improvement of the prognostic performance of CV outcomes, it is conceivable that metabolomic findings may also foster a better understanding of the pathophysiology and biological mechanisms involved in the development of CHF events. Altered lipid metabolism and dyslipidaemia are known to be associated with inflammation and oxidative stress, which are primary drivers of the pathological changes in CHF. The CLP metabolites belong to three different lipid classes, sphingomyelin (SM) phosphatidylcholine (PC), and triglycerides (TAG), and may be involved in different dysregulated metabolic pathways in CHF such as cell stress, inflammation, and atherosclerosis, although future studies are needed to assess whether the CLP biomarkers are representative of altered biological pathways. It has been previously shown that pathway-specific biomarkers/scores consisting of high-sensitivity C-reactive protein (inflammation), soluble urokinase plasminogen activator receptor (inflammation), fibrin degradation products (thrombosis), and heat shock protein 70 (cell stress) significantly improved the prediction of adverse cardiac events in high-risk populations.

These studies also reported similar increases in c-statistics as this study after adding the pathway-specific biomarkers to predictive models.[46–48]

The combination of the CLP's metabolomic features with NT-proBNP may help overcome well-known limitations of NT-proBNP regarding clinical risk factors like age, gender, body mass index, and LVEF. A strength of this study is the high mean age, because elderly patients are underrepresented in CHF trials although CHF is responsible for a great deal of morbidity and mortality in the aging population.[49] Moreover, study samples were derived from a well-characterized cohort including high-quality assessment of outcome events. In future studies, we would like to further elucidate the prognostic utility of the CLP and externally validate its clinical effectiveness by including a larger cohort with more women and patients with early stage CHF and testing different biological matrices (e.g. plasma). Following these studies and regulatory approval, it is conceivable that this biomarker panel can be tested alongside the standard NTpro-BNP test in the clinical setting for a more precise risk assessment of CHF patients (Supporting Information, *Figure S4*).

Study limitations

Our findings can only be interpreted in the context of this specific subcohort and the CLP metabolites, which limits the generalizability of our findings. We were limited by the ability to perform the CLP analysis on separate cohorts, but these proof of concept data can be used as a reference point for additional and larger validation cohorts in the future. The CLP was originally discovered and intended as a diagnostic biomarker, and we cannot assume that it is also a powerful prognostic algorithm as these are still preliminary findings needing validation. Ideally, a prospective derivation validation design using an untargeted metabolite profiling approach should be used to discover a novel prognostic biomarker; however, we were limited on available data and resources. Our findings can only be interpreted as exploratory.

The sample selection criteria, based on the availability and quality of blood samples, may have introduced selection bias for subjects who were more willing or prone to have blood withdrawn and may have excluded patients who were not able to provide sufficient blood possibly due to other CV risk factors, socio-economic status, or comorbidities. The serum samples used in this study may have been affected by the long-term storage prior to the CLP assay, as lipid parameters are known to be subject to in vitro degradation. The quality assurance methods used in this CLP protocol only apply to the sample preparation and measurement requirements for the identification and analysis of the CLP features and do not adjust for any potential effects of prolonged storage. Although NT-proBNP is the gold standard biomarker for CHF

patients, we did not find a significant increase in prognostic power after adding this biomarker to the clinical model, possibly due to the homogeneity of our population of elderly, stable CHF patients, in which it may be of limited use given the fact that NT-proBNP is a marker of volume overload and elevated atrial pressures. Other common cardiac biomarkers, such as troponins or C-reactive protein, should also be evaluated for their incremental prognostic power because a more comprehensive biomarker profile for prognosis may be a better solution than including only CLP plus NT-proBNP.

The samples from the population in this study may have been affected by other medications or a combination of co-morbidities that can affect the lipid metabolites in the CLP. The cut-off values used to generate the CLP risk score using Youden's index are specific to this cohort and not universally applicable, as a large validation cohort(s) would be required to create a generalized equation that could be used in the daily routine management of CHF. The NRI as well as the IDI can be affected by the event rate, which is low in our study.. Although all biomarkers were log transformed, they were not normally distributed, which could affect the concordance of the NRI and IDI. The choice of cut-offs for categorical NRI to determine incremental predictive performance was challenging, as there seems to be no standardized guideline for choosing NRI cut-offs. We found that the NRI was sensitive to changes in the definition of risk categories; however, results did not differ in the sensitivity analysis.

Although this biomarker was developed for routine clinical use, it is currently only available in specialized labs equipped with mass spectrometry equipment.While the CLP is still a research tool awaiting further translation to the routine lab, as an ELISA test for example, this study represents the first step towards that direction.

Conclusions

Our findings demonstrate that the CLP risk score comprising a panel of three lipid-based metabolomic features meaningfully improved the prediction of CV mortality and reclassified patients to their proper risk categories. This new panel of lipid metabolites may complement currently used biomarkers such as NT-proBNP. Thus, the metabolomics approach may potentially translate into clinical applications such as routinely applied risk stratification and targeted treatments for CHF patients.

Acknowledgements

Conflict of interest

Funding

Supporting information

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

Data S1. Supplemental equations.
Figure S1. Sensitivity analysis of the discrimination performance of the CLP biomarker risk score for 4-year cardiovascular mortality using logistic regression.
Figure S2. Sensitivity analysis of the discrimination performance of the CLP biomarker risk score for Major Adverse Cardiovascular Events (MACE) using logistic regression.
Figure S3. Sensitivity analysis of the discrimination performance of the CLP biomarker risk score for all-cause mortality using logistic regression.
Figure S4. Proposed patient flow for testing the CLP in addition to NT-pro-BNP.
Table S1. Clinical variables and multivariable HRs.
Table S2. Harrell's Concordance Statistics for the discrimination of Models A, B, and C.
Table S3. (A) Frequency of events and non-events per risk category of Model B by cardiovascular mortality status using risk categories of 60%–85%. (B) Frequency of events and non-events per risk category

of Model C by cardiovascular mortality status using risk categories of 60%–85%.

Table S4. *(A) Sensitivity Analysis of the Frequency of events and non-events per risk category of Model B by cardiovascular mortality status of Readjusted Risk Categories 70–90%. (B) Sensitivity Analysis of the frequency of events and non-events per risk category of Model C by cardiovascular mortality status using the readjusted risk categories 70–90%.*

Table S5. *(A) Sensitivity analysis of the frequency of events and non-events per risk category of Model B by cardiovascular mortality status using the readjusted risk categories 80–95%. (B) Sensitivity analysis*

of the frequency of events and non-events per risk category of Model B by cardiovascular mortality status using the readjusted risk categories 80–95%.

Table S6. *Sensitivity analysis for risk re-classification after adding the CLP risk score to the NT-proBNP based model using the readjusted risk categories 70–90%.*

Table S7. *Sensitivity Analysis of Adjusted Risk Categories for Risk re-classification after adding the CLP risk score to the NT-proBNP based model (n = 280) 80%–95%.*

References

1. Bui AL, Horwich TB, Fonarow GC. Epidemiology and risk profile of heart failure. *Nat Rev Cardiol* 2011; **8**: 30–41.

2. Ho KK, Pinsky JL, Kannel WB, Levy D. The epidemiology of heart failure: the Framingham study. *J Am Coll Cardiol* 1993; **22**: A6–A13.

3. Kempf T, von Haehling S, Peter T, Allhoff T, Cicoira M, Doehner W, Ponikowski P, Filippatos GS, Rozentryt P, Drexler H, Anker SD. Prognostic utility of growth differentiation factor-15 in patients with chronic heart failure. *J Am Coll Cardiol* 2007; **50**: 1054–1060.

4. Anand IS, Kempf T, Rector TS, Tapken H, Allhoff T, Jantzen F, Kuskowski M, Cohn JN, Drexler H, Wollert KC. Serial measurement of growth-differentiation factor-15 in heart failure: relation to disease severity and prognosis in the Valsartan Heart Failure Trial. *Circulation* 2010; **122**: 1387–1395.

5. Niizeki T, Takeishi Y, Kitahara T, Suzuki S, Sasaki T, Ishino M, Kubota I. Combination of conventional biomarkers for risk stratification in chronic heart failure. *J Cardiol* 2009; **53**: 179–187.

6. Ky B, French B, Levy WC, Sweitzer NK, Fang JC, Wu AH, Goldberg LR, Jessup M, Cappola TP. Multiple biomarkers for risk prediction in chronic heart failure. *Circ Heart Fail* 2012; **5**: 183–190.

7. Lok DJ, Van Der Meer P, Lipsic E, Van Wijngaarden J, Hillege HL, van Veldhuisen DJ. Prognostic value of galectin-3, a novel marker of fibrosis, in patients with chronic heart failure: data from the DEAL-HF study. *Clin Res Cardiol* 2010; **99**: 323–328.

8. de Boer RA, Lok DJ, Jaarsma T, van der Meer P, Voors AA, Hillege HL, van Veldhuisen DJ. Predictive value of plasma galectin-3 levels in heart failure with reduced and preserved ejection fraction. *Ann Med* 2011; **43**: 60–68.

9. Wang CH, Cheng ML, Liu MH, Kuo LT, Shiao MS. Metabolic profile provides prognostic value better than galectin-3 in patients with heart failure. *J Cardiol* 2017; **70**: 92–98.

10. Masson S, Anand I, Favero C, Barlera S, Vago T, Bertocchi F, Maggioni AP, Tavazzi L, Tognoni G, Cohn JN, Latini R. Serial measurement of cardiac troponin T using a highly sensitive assay in patients with chronic heart failure: data from 2 large randomized clinical trials. *Circulation* 2012; **125**: 280–288.

11. Egstrup M, Schou M, Tuxen CD, Kistorp CN, Hildebrandt PR, Gustafsson F, Faber J, Goetze JP, Gustafsson I. Prediction of outcome by highly sensitive troponin T in outpatients with chronic systolic left ventricular heart failure. *Am J Cardiol* 2012; **110**: 552–557.

12. Bruins S, Fokkema MR, Römer JW, DeJongste MJ, van der Dijs FP, van den Ouweland JM, Muskiet FA. High intraindividual variation of B-type natriuretic peptide (BNP) and amino-terminal proBNP in patients with stable chronic heart failure. *Clin Chem* 2004; **50**: 2052–2058.

13. Wu AH, Smith A, Wieczorek S, Mather JF, Duncan B, White CM, McGill C, Katten D, Heller G. Biological variation for N-terminal pro-and B-type natriuretic peptides and implications for therapeutic monitoring of patients with congestive heart failure. *Am J Cardiol* 2003; **92**: 628–631.

14. Schou M, Gustafsson F, Nielsen PH, Madsen LH, Kjaer A, Hildebrandt PR. Unexplained week-to-week variation in BNP and NT-proBNP is low in chronic heart failure patients during steady state. *Eur J Heart Fail* 2007; **9**: 68–74.

15. Dodd MS, Atherton HJ, Carr CA, Stuckey DJ, West JA, Griffin JL, Radda GK, Clarke K, Heather LC, Tyler DJ. Impaired in vivo mitochondrial Krebs cycle activity after myocardial infarction assessed using hyperpolarized magnetic resonance spectroscopy. *Circ Cardiovasc Imaging* 2014; **7**: 895–904.

16. Gupte AA, Hamilton DJ, Cordero-Reyes AM, Youker KA, Yin Z, Estep JD, Stevens RD, Wenner B, Ilkayeva O, Loebe M, Peterson LE. Mechanical unloading promotes myocardial energy recovery in human heart failure. *Circ Cardiovasc Genet* 2014; **7**: 266–276.

17. Bedi KC Jr, Snyder NW, Brandimarto J, Aziz M, Mesaros C, Worth AJ, Wang LL, Javaheri A, Blair IA, Margulies KB, Rame JE. Evidence for intramyocardial disruption of lipid metabolism and increased myocardial ketone utilization in advanced human heart failure. *Circulation* 2016; **133**: 706–716.

18. Tang WW, Wang Z, Fan Y, Levison B, Hazen JE, Donahue LM, Wu Y, Hazen SL. Prognostic value of elevated levels of intestinal microbe-generated metabolite trimethylamine-N-oxide in patients with heart failure: refining the gut hypothesis. *J Am Coll Cardiol* 2014; **64**: 1908–1914.

19. Ahmad T, Kelly JP, McGarrah RW, Hellkamp AS, Fiuzat M, Testani JM, Wang TS, Verma A, Samsky MD, Donahue MP, Ilkayeva OR. Prognostic implications of long-chain acylcarnitines in heart failure and reversibility with mechanical circulatory support. *J Am Coll Cardiol* 2016; **67**: 291–299.

20. Cheng ML, Wang CH, Shiao MS, Liu MH, Huang YY, Huang CY, Mao CT, Lin JF, Ho HY, Yang NI. Metabolic disturbances identified in plasma are associated with outcomes in patients with heart failure: diagnostic and prognostic value of metabolomics. *J Am Coll Cardiol* 2015; **65**: 1509–1520.

21. Ruiz-Canela M, Hruby A, Clish CB, Liang L, Martínez-González MA, Hu FB. Comprehensive metabolomic profiling and incident cardiovascular disease: a systematic review. *J Am Heart Assoc* 2017; **6**: e005705.

22. Shah SH, Bain JR, Muehlbauer MJ, Stevens RD, Crosslin DR, Haynes C, Dungan J, Newby LK, Hauser ER, Ginsburg GS, Newgard CB. Association of a peripheral blood metabolic profile with coronary artery disease and risk of subsequent cardiovascular events. *Circ Cardiovasc Genet* 2010; **3**: 207–214.

23. Shah SH, Sun JL, Stevens RD, Bain JR, Muehlbauer MJ, Pieper KS, Haynes C, Hauser ER, Kraus WE, Granger CB,

Newgard CB. Baseline metabolomic profiles predict cardiovascular events in patients at risk for coronary artery disease. *Am Heart J* 2012; **163**: 844–850.

24. Rizza S, Copetti M, Rossi C, Cianfarani MA, Zucchelli A, Luzi A, Pecchioli C, Porzio O, Di Cola G, Urbani A, Pellegrini F. Metabolomics signature improves the prediction of cardiovascular events in elderly subjects. *Atherosclerosis* 2014; **232**: 260–264.

25. Vaarhorst AA, Verhoeven A, Weller CM, Böhringer S, Göraler S, Meissner A, Deelder AM, Henneman P, Gorgels AP, van den Brandt PA, Schouten LJ. A metabolomic profile is associated with the risk of incident coronary heart disease. *Am Heart J* 2014; **168**: 45–52.

26. Kume S, Araki SI, Ono N, Shinhara A, Muramatsu T, Araki H, Isshiki K, Nakamura K, Miyano H, Koya D, Haneda M. Predictive properties of plasma amino acid profile for cardiovascular disease in patients with type 2 diabetes. *PLoS ONE* 2014; **9**: e101219.

27. Zheng Y, Yu B, Alexander D, Manolio TA, Aguilar D, Coresh J, Heiss G, Boerwinkle E, Nettleton JA. Associations between metabolomic compounds and incident heart failure among African Americans: the ARIC study. *Am J Epidemiol* 2013; **178**: 534–542.

28. Lanfear DE, Gibbs JJ, Li J, She R, Petucci C, Culver JA, Tang WW, Pinto YM, Williams LK, Sabbah HN, Gardell SJ. Targeted metabolomic profiling of plasma and survival in heart failure patients. *JACC: Heart Failure* 2017; **5**: 823–832.

29. Würtz P, Havulinna AS, Soininen P, Tynkkynen T, Prieto-Merino D, Tillin T, Ghorbani A, Artati A, Wang Q, Tiainen M, Kangas AJ. Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. *Circulation* 2015; **131**: 774–785.

30. Mueller-Hennessen M, Düngen HD, Lutz M, Trippel TD, Kreuter M, Sigl J, Müller OJ, Tahirovic E, Witt H, Ternes P, Carvalho S. A novel lipid biomarker panel for the detection of heart failure with reduced ejection fraction. *Clin Chem* 2017; **63**: 267–277.

31. Düngen HD, Apostolović S, Inkrot S, Tahirović E, Krackhardt F, Pavlović M, Putniković B, Lainščak M, Gelbrich G, Edelmann F, Wachter R. Bisoprolol vs. carvedilol in elderly patients with heart failure: rationale and design of the CIBIS-ELD trial. *Clin Res Cardiol* 2008; **97**: 578–586.

32. Düngen HD, Apostolović S, Inkrot S, Tahirović E, Töpper A, Mehrhof F, Prettin C, Putniković B, Nešković AN, Krotin M, Sakač D. Titration to target dose of bisoprolol vs. carvedilol in elderly patients with heart failure: the CIBIS-ELD trial. *Eur J Heart Fail* 2011; **13**: 670–680.

33. KamlageB, Maldonado SG, Bethan B, Peter E, Schmitz O, Liebenberg V, Schatz P. Quality Markers Addressing Preanalytical Variations of Blood and Plasma Processing Identified by Broad and Targeted Metabolite Profiling. *Clin Chem* 2014; **60**: 399–412.

34. Rickham PP. Human experimentation. Code of ethics of the world medical association. Declaration of Helsinki. *Br Med J* 1964; **2**: 177.

35. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950; **3**: 32–35.

36. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med* 2011; **30**: 1105–1117.

37. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008; **27**: 157–172.

38. Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* 2011; **30**: 11–21.

39. Pencina MJ, D'agostino RB, Pencina KM, Janssens AC, Greenland P. Interpreting incremental value of markers added to risk prediction models. *Am J Epidemiol* 2012; **176**: 473–481.

40. SAS Institute Inc. *Base SAS 9.4 Procedures Guide*, 5th ed. Cary, NC: SAS Institute Inc.; 2015 http://support.sas.com (13 February 2020).

41. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2019. https://www.R-project.org/ (13 February 2020).

42. Kassambara A, Kosinski M, Biecek P. survminer: drawing survival curves using 'ggplot2'. R package version 0.4.6. 2019. https://CRAN.R-project.org/package=survminer

43. Maisel A, Mueller C, Adams K Jr, Anker SD, Aspromonte N, Cleland JG, Cohen-Solal A, Dahlstrom U, DeMaria A, Di Somma S, Filippatos GS. State of the art: using natriuretic peptide levels in clinical practice. *Eur J Heart Fail* 2008; **10**: 824–839.

44. Jensen J, Ma LP, Bjurman C, Hammarsten O, Fu ML. Prognostic values of NTpro BNP/BNP ratio in comparison with NTpro BNP or BNP alone in elderly patients with chronic heart failure in a 2-year follow up. *Int J Cardiol* 2012; **155**: 1–5.

45. McGranaghan P, Saxena A, Rubens M, Radenkovic J, Bach D, Schleußner L, Pieske B, Edelmann F, Trippel TD. Predictive value of metabolomic biomarkers for cardiovascular disease risk: a systematic review and meta-analysis. *Biomarkers* 2020; **25**: 101–111.

46. Eapen DJ, Manocha P, Patel RS, Hammadah M, Veledar E, Wassel C, Nanjundappa RA, Sikora S, Malayter D, Wilson PW, Sperling L. Aggregate risk score based on markers of inflammation, cell stress, and coagulation is an independent predictor of adverse cardiovascular outcomes. *J Am Coll Cardiol* 2013; **62**: 329–337.

47. Eapen DJ, Manocha P, Ghasemzadeh N, Patel RS, Al Kassem H, Hammadah M, Veledar E, Le NA, Pielak T, Thorball CW, Velegraki A. Soluble urokinase plasminogen activator receptor level is an independent predictor of the presence and severity of coronary artery disease and of future adverse events. *J Am Heart Assoc* 2014; **3**: e001118.

48. Ghasemzadeh N, Hayek SS, Ko YA, Eapen DJ, Patel RS, Manocha P, Al Kassem H, Khayata M, Veledar E, Kremastinos D, Thorball CW. Pathway-specific aggregate biomarker risk score is associated with burden of coronary artery disease and predicts near-term risk of myocardial infarction and death. *Circ Cardiovasc Qual Outcomes* 2017; **10**: e001493.

49. Shibata MC, Soneff CM, Tsuyuki RT. Utilization of evidence-based therapies for heart failure in the institutionalized elderly. *Eur J Heart Fail* 2005; **7**: 1122–1125.