

# **SUPPLEMENTAL MATERIAL**

## **Supplemental Methods**

### **Data S1. List of Enrolling Centers**

Barnes-Jewish Hospital, St. Louis, MO

Cleveland Clinic Foundation, Cleveland, OH

Dartmouth-Hitchcock Medical Center, Lebanon, NH

Intermountain Heart Institute, Murray, UT

Massachusetts General Hospital, Boston, MA

Morristown Medical Center, Morristown, NJ

Stanford Medical Center, Palo Alto, CA

University of Texas Southwestern Medical Center, Dallas, TX

University of Utah Hospital, Salt Lake City, UT

Vanderbilt University Medical Center, Nashville, TN

## Data S2.

### Metabolite profiling using liquid chromatography-mass spectrometry.

Methods for metabolite quantification were reproduced directly from other work to enhance rigor and reproducibility

(<https://www.patentsencyclopedia.com/app/20080261317>; Date Accessed 1 May 2023)<sup>50</sup>. Venous blood was processed within 30 minutes of collection and stored at -80°C.

Amino acids and amines, sugars and ribonucleotides, and organic acids were separated by liquid chromatography. Columns were connected in parallel with an automated switching valve on a robotic sample loader (Leap Technologies). A triple-quadrupole mass spectrometer (API4000, Applied Biosystem/ Sciex) operated in automated switching polarity mode with a turbo ion spray LC-MS interface under selected reaction monitoring conditions. Either positive or negative ions were selected for targeted tandem mass spectrometry (MS/MS) analysis using selective reaction conditions. Quantification was performed by integrating peak areas for parent/daughter ion pairs. Metabolite concentrations  $\leq 0$  were treated as missing (0.001%). We excluded metabolites with  $\geq 10\%$  missingness (4/225), and imputed missing values for any metabolite with  $< 10\%$  missingness by half of the minimum value detected for that metabolite (0.1% of metabolite measurements). Two subjects had metabolite quantification duplicated, we used the mean concentration for these subjects in analysis.

### **Principal components analysis to summarize echocardiographic data into 3 axes**

**of cardiac structure/function.** Given the relatedness among the different echocardiographic measures (**Figure S2**), we elected to summarize echocardiographic measures into different, related “axes” of remodeling/function using principal components analysis (PCA; with echocardiographic measures log-transformed to reduce skewness and standardized to unit variance). In this unsupervised PCA approach, the relations among phenotypes were used to collapse the 12 echocardiographic measures into 3 composite axes in an unbiased manner (selected based on examination of variance explained on a scree plot). Broadly, the contribution of each of the 12 echocardiographic measures to each of 3 PCs was quantitated as the loading for the PC (**Figure 2A**). Study participants had a score for each of the 3 PCs that summarized the aspect of their cardiac structure/function captured by that PC. Varimax post-rotation for 3 PCs was used to improve interpretability of loadings across PCs. Scores for each participant represented composite axes of cardiac structure/function and used as the dependent variable in subsequent LASSO models, thereby capturing a broad array of related echocardiographic measures to supervise selection of metabolites most closely related to those axes. While we acknowledge there are many different additional approaches that could be used here (e.g., pre-selecting phenotypes and using those for penalized regression, clustering methods, etc.), we felt that this approach would preserve power to develop optimal models for discovery while limiting type 1 error (due to multiple biomarker testing) that may reduce reproducibility.

**Recalibration of LASSO models for use in the single center validation cohort.** Given that the metabolites quantified in the single-center validation cohort did not completely match the metabolites measured in our multi-center cohort, we used the metabolites that overlapped between the single-center validation and the multi-center cohort (78 metabolites) to refit models in the multi-center derivation cohort. In these models, the original metabolite score (based on

the “full” LASSO regressions across all metabolites in derivation) was the dependent variable, and the overlapping metabolites were the independent variables. We used a LASSO model for this recalibration effort to mitigate overfitting, generating coefficients that could be applied to overlapping metabolites in the single-center validation cohort to generate the 3 metabolite scores.

**Table S1: Metabolite Correlation with Echocardiographic Parameters from LASSO**

Metabolite	HMDB	LASSO Coefficient			Biological Significance
		PC1	PC2	PC3	
Phosphocreatine	HMDB0001511	-0.084	-	-	Derivative of amino acid creatine. Plays important role in ATP generation for cardiac contraction <sup>51,52</sup> .
Nicotinamide N-Oxide	HMDB0002730	-0.087	-	-0.014	Antagonizes CXCR2 which recruits granulocytes <sup>53</sup> .
N-Docosanoyl Taurine	NA	0.072	-	-	Taurine-conjugated fatty acid.
Adenine	HMDB0000034	0.068	-	-	Nucleic acid.
Serine	HMDB0000187	0.053	0.019	-	Amino acid.
N-Palmitoyltaurine	HMDB0240594	0.059	-	-	Derivative of hexadecenoic acid (fatty acid).
Tryptophan	HMDB0000929	0.045	-	-	Amino acid.
N-Acetyl-L-Phenylalanine	HMDB0000512	0.052	-	-	Derivative of amino acid phenylalanine.
Aconitic acid	HMDB0000072	-0.044	-	-	TCA cycle intermediate. Associated with aortic stenosis <sup>54</sup> .
Bilirubin	HMDB0000054	0.043	-	-	Derivative of heme. Associated with mortality among subjects undergoing TAVI <sup>55</sup> .
C9 acylcarnitine	HMDB0013288	0.045	-	-	Medium chain acylcarnitine.
Oleoyl glycine	HMDB0013631	-0.038	-	-	Long chain fatty acyl glycine.
Creatine	HMDB0000064	-0.039	0.051	0.036	Amino acid with role in myocardial contraction <sup>52</sup> .
Phosphocholine	HMDB0001565	-0.038	-	-	Product of choline kinase to dephosphorylate ATP.
2'-deoxycytidine	HMDB0000014	0.045	-	-	Deoxyribonucleoside.
N-Acetyl-L-Glutamic acid	HMDB0001138	-0.032	-	-	Derivative of L-glutamic acid.
Aspartic acid	HMDB0000191	0.031	0.074	-	Amino acid.
N-Acetyl-L-Ornithine	HMDB0003357	-0.020	-	-	Derivative of amino acids arginine and proline.
2-Hydroxybutyric acid	HMDB0000008	-0.024	-	-	Organic acid from amino acid metabolism (threonine and methionine). A marker of insulin resistance <sup>56</sup> .
Sarcosine	HMDB0000271	-0.020	-	-	Derivative of amino acid glycine.
C26 acylcarnitine	HMDB0002356	0.020	-	-	Long chain acylcarnitine. An intermediate of fatty acid metabolism, associated with LV remodeling in AS <sup>12</sup> .
Saccharopine	HMDB0000279	-0.025	-	-	Derivative of amino acid lysine.
Glycocholic acid	HMDB0000138	-0.023	-	-	Bile acid conjugate of glycine and choline.
Kynurenine	HMDB0000684	0.033	-	-	Metabolite of tryptophan (amino acid) that has been

					shown to suppress immune response <sup>57</sup> . Associated with survival among patients with heart failure <sup>31</sup> .
Lactic acid	HMDB0000190	-0.030	-	-	Organic acid derivative of glucose metabolism. May play a in modulating hypertrophy and heart failure <sup>58</sup> .
Choline	HMDB0000097	0.013	-	-	Associated with cardiovascular disease <sup>36,59</sup> . Associated with LV remodeling in AS <sup>12</sup> .
DDHAP / Glyceraldehyde 3-phosphate	HMDB0001112	-0.019	-0.037	-	Derivative of glucose metabolism.
DiHOME	HMDB0004705	-0.018	-	-	Linoleic acid metabolite. Increases fatty acid uptake in skeletal muscle <sup>60</sup> .
Dihomo-γ-Linolenoyl Ethanolamide	HMDB0013625	0.020	0.008	-	Endocannabinoid.
Uracil	HMDB0000300	-0.008	-	-	Nucleobase of RNA. May play a role in cardiac hypertrophy <sup>30</sup> .
Glucose/Fructose/Galactose	HMDB0000122 (glucose)	-0.022	-	-	Monosaccharides.
Cystine	HMDB0000192	-0.010	-	-	Amino acid. Associated with diastolic dysfunction <sup>61</sup> .
Glycochenodeoxycholic acid	HMDB0304944	-0.003	-	-	Bile salt.
N-carbomoyl-beta-alanine	HMDB0000026	0.009	-0.028	-	Derivative of amino acid uracil.
C18.2 carnitine	HMDB0006469	0.017	-	-	An intermediate of fatty acid metabolism, associated with LV remodeling in AS <sup>12</sup> . Related to coronary artery disease <sup>33</sup> .
Histamine	HMDB0000870	-0.001	-	-	Imidazole. May play a role in the development of heart failure and cardiac fibrosis after MI <sup>62,63</sup> .
Cyclic AMP	HMDB0000058	-0.007	-	-	Derivative of adenosine triphosphate used in intracellular signaling for a variety of pathways.
Anserine	HMDB0000194	-0.004	-	-	A dipeptide derivative of carnosine.
Cytidine	HMDB0000089	0.003	-	-	Nucleoside component of RNA.
α-Ketoglutaric acid	HMDB0000208	-	-0.100	-	TCA cycle intermediate. In a murine model, supplementation lessened pressure-overload cardiac hypertrophy and preserved systolic function <sup>35</sup> .
Beta-alanine	HMDB0000056	-	-0.086	-	Amino acid. Supplementation shown to improve response to strength training in humans <sup>64</sup> and functional capacity

					in a murine model of heart failure <sup>65</sup> .
Glutamate	HMDB0000148	-	-0.059	-	Amino acid and a common neurotransmitter. Used in some formulations of cardioplegia due to possible improvements in outcomes <sup>66</sup> . Shown to reduce oxidative injury in a murine model of myocardial infarction <sup>67</sup> . Associated with increased risk of stroke <sup>68</sup> .
Uridine	HMDB0000296	-	0.058	-	Nucleic acid. Supplementation reduced ischemic reperfusion injury in a murine model <sup>69</sup> . Associated with incident heart failure in the Jackson Heart Study <sup>36</sup> .
Uridine (anode)	HMDB0000296	-	0.018	-	
Arginine	HMDB0000517	-	0.052	-	Amino acid and precursor of nitric oxide in the vascular endothelium. Supplementation in subjects with ischemic heart failure had improved function and dimensions <sup>70</sup> .
Glycine	HMDB0000123	-	0.057	-	Amino acid. Elevated in subjects with severe heart failure <sup>71</sup> . Glycine metabolism is different after exercise in subjects with hypertrophic cardiomyopathy <sup>72</sup> .
UDP-glucose / UDP-galactose	HMDB0000286 (glucose)	-	-0.050	-	Nucleotide sugar. Involved in carbohydrate metabolism. Implicated as a coronary vasoconstrictor <sup>73</sup> .
Spermidine	HMDB0001257	-	-0.046	-	Polyamine. Elevated in murine models of pressure overload and infarct hearts <sup>74</sup> . May be helpful in prognosticating outcomes in heart failure <sup>28</sup> .
2-aminoadipic acid	HMDB0302754	-	-0.031	-	Product of lysine metabolism. Associated with incident heart failure in the Jackson Heart Study <sup>36</sup> and development of T2DM <sup>75</sup> and atherosclerosis <sup>76</sup> .
N-oleoyl dopamine	HMDB0255218	-	0.020	-	Fatty amide. Protective against ischemia-reperfusion injury <sup>34</sup> .
Tyrosine	HMDB0000158	-	-0.023	-	Amino acid. A disorder of tyrosine catabolism, alkaptonuria, is associated with aortic stenosis <sup>77</sup> .
N-Carbamoyl-BAIBA	NA	-	-	0.102	Product of thymine catabolism
S-adenosyl-L-homocysteine	HMDB0000939	-	-	0.076	Sole metabolic precursor to homocysteine and a derivative of amino acids methionine and



					cysteine. Associated with coronary heart disease <sup>32</sup> .
20-Hydroxy N-Arachidonoyl-Taurine	NA	-	-	0.012	Fatty acid amide of amino acid taurine.
Histidine	HMDB0000177	-	-	-0.069	Amino acid. Improved functional capacity in a murine model of heart failure <sup>65</sup> .
N-acetyl-L-methionine	HMDB0011745	-	-	0.055	Derivative of amino acid methionine. May be helpful in reducing ischemia-reperfusion injury <sup>78</sup> .
Valine	HMDB0000883	-	-	-0.044	Branched chain amino acid. Defective metabolism may be implicated in heart failure pathogenesis <sup>21</sup> and cardiac insulin resistance <sup>79</sup> .
Oxaloacetic acid	HMDB0000223	-	-	0.028	Organic acid involved in numerous metabolic pathways.
Leucine	HMDB0000687	-	-	-0.006	Branched chain amino acid. Defective metabolism may be implicated in heart failure pathogenesis <sup>21</sup> and cardiac insulin resistance <sup>79</sup> .
Malic acid	HMDB0000156	-	-	0.004	TCA cycle intermediate. Associated with atrial fibrillation and heart failure <sup>80</sup> . Reduced ischemia reperfusion injury in mice <sup>81</sup> .

ATP = adenosine triphosphate; CXCR2 = chemokine (CXC motif) receptor 2; TCA = tricarboxylic acid; TAVI = transcatheter aortic valve implantation; LV = left ventricular; AS = aortic stenosis; RNA = Ribonucleic acid; MI = myocardial infarction; T2DM = type 2 diabetes mellitus.

**Table S2:** Spearman correlation between phenotype scores, metabolite scores, select echocardiographic variables (exemplary of each PC-based phenotype category), and multi-morbidity measures from the multi-center derivation cohort. N represents the number of observations on which the correlation is based (not all individuals had every measure). The P-value reported is the nominal P-value for the correlation, **Figure 4** in the manuscript uses an asterisk to indicate which P-values passed a false discovery rate of 5% (Benjamini-Hochberg). Of note, phenotype scores were not calculated in the multi-center validation cohort due to incomplete echocardiographic data (see **Table 1**). eGFR = estimate glomerular filtration rate, FEV1 = Forced expiratory volume in the first second; PC = principal component; LVEDVI = left ventricular end-diastolic volume index; LVEF = left ventricular ejection fraction.

<b>Correlation coefficient</b>									
	PC1	Metabolite -PC1	LVEDVI	PC2	Metabolite -PC2	LVEF	PC3	Metabolite -PC3	Mean E/e'
Age (N=519)	-0.002	0.022	-0.107	0.045	0.108	0.009	0.141	0.067	0.036
Kansas City Cardiomyopathy Questionnaire summary score (N=487)	-0.033	-0.044	-0.038	0.167	0.236	0.130	-0.222	-0.286	-0.169
Average grip strength (N=475)	0.141	0.116	0.207	-0.118	-0.253	-0.190	-0.164	-0.138	-0.195
Average gait speed (N=458)	0.116	-0.036	0.127	0.005	0.141	-0.031	-0.182	-0.281	-0.173
Mini-Cog total score (N=484)	-0.024	0.013	0.023	-0.037	-0.006	-0.031	-0.102	-0.051	-0.030
FEV1, % Predicted (N=410)	-0.103	-0.166	-0.121	0.204	0.256	0.218	-0.085	-0.157	-0.082
Hemoglobin (N=516)	-0.059	-0.147	-0.083	-0.142	-0.088	-0.033	-0.216	-0.359	-0.242
Platelets (N=515)	-0.020	-0.097	-0.043	-0.035	0.030	-0.014	-0.020	0.016	0.064
Uric Acid (N=504)	0.058	0.170	0.096	-0.054	-0.199	-0.096	0.129	0.299	0.045
eGFR (N=517)	-0.019	-0.158	-0.033	-0.041	0.042	0.038	-0.211	-0.509	-0.203

<b>P-value of correlation</b>									
Age	9.69E-01	6.10E-01	1.49E-02	3.06E-01	1.37E-02	8.39E-01	1.26E-03	1.25E-01	4.15E-01
Kansas City Cardiomyopathy Questionnaire summary score	4.63E-01	3.36E-01	4.01E-01	2.16E-04	1.45E-07	3.94E-03	7.58E-07	1.22E-10	1.80E-04
Average grip strength	2.04E-03	1.11E-02	5.49E-06	1.04E-02	2.35E-08	3.02E-05	3.18E-04	2.67E-03	1.83E-05
Average gait speed	1.30E-02	4.37E-01	6.37E-03	9.15E-01	2.48E-03	5.04E-01	8.95E-05	8.67E-10	1.95E-04
Mini-Cog total score	5.92E-01	7.72E-01	6.09E-01	4.18E-01	8.94E-01	5.02E-01	2.44E-02	2.62E-01	5.05E-01
FEV1, % Predicted	3.66E-02	7.29E-04	1.39E-02	3.29E-05	1.52E-07	8.81E-06	8.67E-02	1.45E-03	9.75E-02
Hemoglobin	1.83E-01	7.95E-04	5.84E-02	1.20E-03	4.66E-02	4.53E-01	7.60E-07	<2.20E-16	2.56E-08
Platelets	6.43E-01	2.76E-02	3.31E-01	4.24E-01	4.90E-01	7.53E-01	6.45E-01	7.10E-01	1.45E-01
Uric Acid	1.97E-01	1.26E-04	3.07E-02	2.25E-01	6.94E-06	3.18E-02	3.82E-03	7.08E-12	3.17E-01
eGFR	6.61E-01	3.11E-04	4.52E-01	3.51E-01	3.37E-01	3.90E-01	1.31E-06	<2.20E-16	3.13E-06

**Table S3:** Spearman correlation between phenotype scores, metabolite scores, select echocardiographic variables (exemplary of each PC-based phenotype category), and multi-morbidity measures in the multi-center validation cohort. N represents the number of observations on which the correlation is based (not all individuals had every measure). The P-value reported is the nominal P-value for the correlation, **Figure 4** in the manuscript uses an asterisk to indicate which P-values passed a false discovery rate of 5% (Benjamini-Hochberg). Of note, phenotype scores were not calculated in the multi-center validation cohort due to incomplete echocardiographic data (see **Table 1**). eGFR = estimate glomerular filtration rate, FEV1 = Forced expiratory volume in the first second.

<b>Correlation coefficient</b>			
	Metabolite-PC1	Metabolite-PC2	Metabolite-PC3
Age (N=286)	0.150	0.100	0.093
Left ventricular ejection fraction (N=260)	-0.177	0.158	-0.273
Left ventricular end diastolic volume index (N=120)	0.139	-0.173	0.237
Mean E/e' (N=79)	0.170	0.005	0.359
Average gait speed (N=226)	-0.065	0.151	-0.316
Average grip strength (N=251)	0.033	-0.142	-0.114
Kansas City Cardiomyopathy Questionnaire summary score (N=267)	-0.017	0.170	-0.200
FEV1, % predicted (N=192)	0.023	0.189	-0.225
Mini-Cog total score (N=265)	0.066	0.062	0.010

Hemoglobin (N=284)	-0.168	-0.114	-0.298
Platelets (N=284)	-0.080	0.018	0.033
Uric acid (N=276)	0.192	-0.202	0.300
eGFR (N=283)	-0.202	0.146	-0.523
<b>P-value of correlation</b>			
Age	1.11E-02	9.18E-02	1.18E-01
Left ventricular ejection fraction	4.17E-03	1.08E-02	8.13E-06
Left ventricular end diastolic volume index	1.31E-01	5.94E-02	9.11E-03
Mean E/e'	1.34E-01	9.66E-01	1.17E-03
Average gait speed	3.31E-01	2.30E-02	1.28E-06
Average grip strength	6.03E-01	2.48E-02	7.15E-02
Kansas City Cardiomyopathy Questionnaire summary score	7.88E-01	5.45E-03	1.05E-03
FEV1, % predicted	7.56E-01	8.69E-03	1.74E-03
Mini-Cog total score	2.81E-01	3.15E-01	8.75E-01
Hemoglobin	4.46E-03	5.56E-02	3.04E-07
Platelets	1.78E-01	7.58E-01	5.81E-01
Uric acid	1.34E-03	7.29E-04	3.87E-07
eGFR	6.42E-04	1.43E-02	<2.20E-16

**Table S4: Cox models for all-cause mortality.** Hazard ratios represent risk for a 1 standard deviation increase. “Adjusted” models in the derivation and multi-center validation cohorts were adjusted for age, sex, body mass index, history of diabetes mellitus, history of coronary artery disease, history of atrial fibrillation or flutter, estimated glomerular filtration rate, high sensitivity troponin, and N-terminal pro hormone B-type natriuretic peptide. “Adjusted” models in the single-center validation cohort were adjusted for age, sex, body mass index, history of diabetes mellitus, history of coronary artery disease, history of atrial fibrillation or flutter, estimated glomerular filtration rate, and B-type natriuretic peptide. “Sensitivity” models include additional adjustments for mean aortic valve gradient and left ventricular ejection fraction. PC = principal component.

		Metabolite-based				Phenotype-based			
		N	Deaths	Hazard Ratio (95% CI)	P-value	N	Deaths	Hazard Ratio (95% CI)	P-value
Derivation	<b>Unadjusted</b>								
	PC1	516	205	1.27 (1.11-1.45)	0.001	516	205	1.01 (0.88 – 1.16)	0.88
	PC2	516	205	0.84 (0.73-0.97)	0.02	516	205	0.84 (0.73 – 0.96)	0.01
	PC3	516	205	1.75 (1.55-1.99)	<0.001	516	205	1.40 (1.22 – 1.62)	<0.001
	<b>Adjusted</b>								
	PC1	494	198	0.98 (0.84-1.15)	0.81	494	198	0.83 (0.70 – 0.98)	0.02
	PC2	494	198	0.94 (0.80-1.10)	0.42	494	198	0.90 (0.78 – 1.05)	0.18
	PC3	494	198	1.54 (1.25-1.90)	<0.001	494	198	1.09 (0.91 – 1.29)	0.34
	<b>Sensitivity</b>								
	PC1	494	198	1.01 (0.86-1.19)	0.87			-	
PC2	494	198	0.95 (0.81-1.12)	0.53			-		
PC3	494	198	1.55 (1.26-1.92)	<0.001			-		
Multi-center Validation	<b>Unadjusted</b>								
	PC1	282	121	1.10 (0.92-1.33)	0.28			-	
	PC2	282	121	0.77 (0.64-0.93)	0.008			-	
	PC3	282	121	1.58 (1.34-1.87)	<0.001			-	
	<b>Adjusted</b>								
	PC1	274	116	0.94 (0.76-1.16)	0.58			-	
	PC2	274	116	0.90 (0.73-1.11)	0.32			-	
PC3	274	116	1.37 (1.05-1.78)	0.02			-		
Single-center Validation	<b>Unadjusted</b>								
	PC1	257	93	1.17 (0.94-1.44)	0.16			-	
	PC2	257	93	0.93 (0.76-1.13)	0.44			-	
	PC3	257	93	1.37 (1.13-1.67)	0.001			-	
	<b>Adjusted</b>								
	PC1	251	91	1.11 (0.86-1.43)	0.43			-	
	PC2	251	91	0.88 (0.72-1.09)	0.25			-	
PC3	251	91	1.43 (1.06-1.92)	0.02			-		

**Table S5:** Results of the "recalibrated" LASSO model using metabolites common to both internal derivation and external validation cohorts. LASSO = least absolute shrinkage and selection operator; PC = principal component.

metabolite	LASSO Coefficient		
	PC1	PC2	PC3
glycine	0.01494987	0.139102307	0.030239001
alanine	-0.251512522	-0.113791394	0.024166223
serine	0.143051865	0.106803934	-0.041300673
threonine	0.050589769	0.033504175	0.040275094
methionine	-	-0.05081661	-
aspartate	0.015995975	-	0.080618582
glutamate	-	-0.16867989	-0.033268046
asparagine	-	0.042200172	-0.086920091
glutamine	-0.005139735	0.022095756	-
histidine	-0.069711788	0.074003134	-0.332893973
arginine	-0.030562809	0.212603096	-0.016608614
lysine	-	0.041159659	-0.001146626
valine	-0.101041903	0.000266757	-0.167696977
leucine	-	0.045187493	-0.110701583
phenylalanine	0.084663329	-	0.048414897
tyrosine	-	-0.114599243	0.002947603
tryptophan	0.101032064	-0.013813357	-0.029229131
proline	-0.053258912	-0.012065449	-0.01105175
cis.trans.hydroxyproline	-	-	-
ornithine	-	-0.068995475	-
citrulline	0.016689991	0.029839466	0.012033922
taurine	0.025119423	-	-
5.hydroxytryptophan	-	-	-0.017393582
5.HIAA	-	-0.006672996	0.042201013
cystamine	-0.062069023	0.013839851	-0.001509095
cysteamine	0.030232505	-0.000299186	-0.007223417
GABA	-	-0.01236756	-0.013089138
dimethylglycine	0.017769162	-	-
homocysteine	-	-0.008340423	-0.001346062
ADMA.SDMA	0.010550956	0.052037309	0.166046764
NMMA	-	-	-
allantoin	-	0.005085548	0.115265873

aminoisobutyric.acid	0.008684823	0.020307994	-
carnitine	-	-0.015765839	-0.056224754
1.methylhistamine	-	-0.027932122	0.023886073
5..adenosylhomocysteine	-	-	0.01108117
3.hydroxyanthranilic.acid	-	-	-0.05508892
N.carbomoyl.beta.alanine	0.113717016	-0.18009016	0.281661982
niacinamide	-0.157019631	0.006085806	-
betaine	0.016686918	0.011834496	-0.011957004
choline	0.009601636	-	0.006595445
phosphocholine	-0.1747666	-	-0.021421928
alpha.glycerophosphocholine	-	0.037644223	0.005717554
spermidine	-0.04104223	-0.191238112	-0.095427795
creatine	-0.158601676	0.132805505	0.188000128
creatinine	0.086683549	-	0.114546392
adenosine	0.037085836	0.003767399	-0.022240968
cytosine	-	0.010832363	-
xanthosine	0.00130912	-0.063711329	0.024627807
cAMP	-0.060711268	-0.034277629	-0.001141732
isoleucine	-	-0.021463649	0.017886777
xanthine	-0.048163564	0.032269234	-0.038850884
xanthurate	0.11585772	-0.095660451	-
kynurenine	0.172680721	0.033820158	0.086083182
uridine	-0.012756245	0.24380548	-0.079669817
citicholine	-0.00029907	-0.134291446	-
beta.alanine	-0.089808244	-0.305192791	-
C2.carnitine	-0.073445173	-0.008632177	-
C3.carnitine	-	-	-0.048265682
C3.malonyl.carnitine	-	-0.067866411	-
C4.butyl.carnitines	0.053656022	0.000160457	0.056067457
C4.methylmalonyl.carnitine	-0.063571497	-0.016661022	0.017603944
C5.valeryl.carnitines	-	-	-
C5.glutaryl.carnitine	0.013413055	-	0.073555365
C6.carnitine	-	-	0.014374416
C7.carnitine	-	-0.038240411	0.005419971
C8.carnitine	-	0.035161499	-0.016166012
C9.carnitine	0.17198752	-	0.02357368
C10.carnitine	-0.03710882	-	-
C12.carnitine	-	-	0.04619551
C14.carnitine	0.075879056	0.038998555	-0.038372988

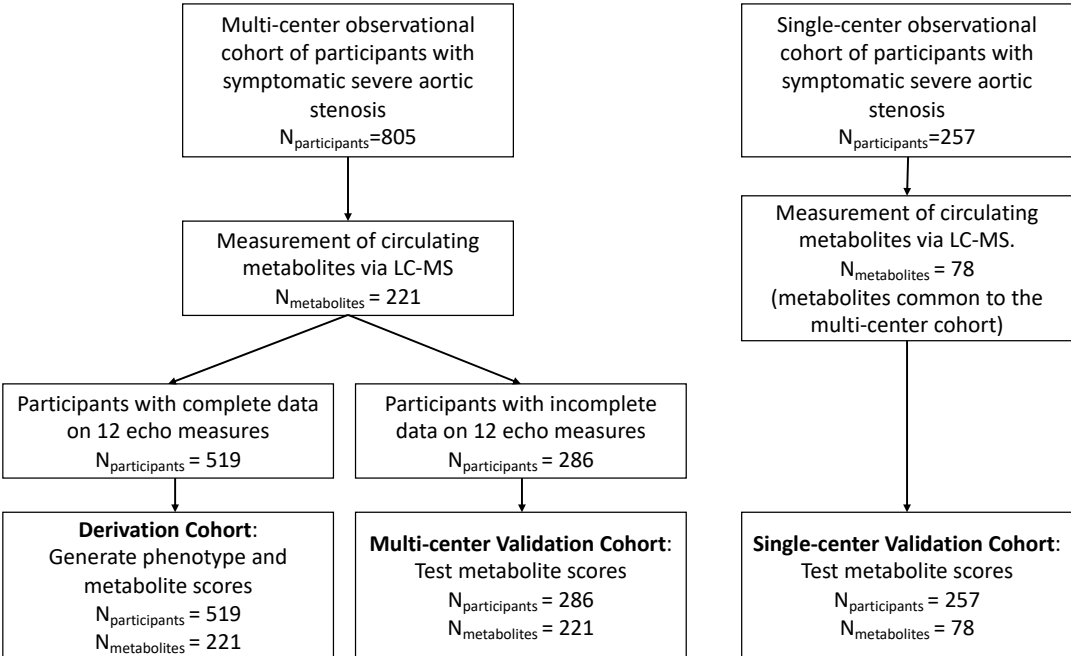


C16.carnitine	-	-	0.005230519
C18.carnitine	-0.034616529	0.024214834	-0.081105255
C18.1.carnitine	-	-0.090844535	0.06323825
C18.2.carnitine	0.14104691	-	0.043582665
C26.carnitine	0.051564122	0.044507161	-0.024568913
anandamide	-0.046417867	-	-
Cystine	-0.059824958	-0.034687137	-0.025504188

**Table S6:** Measures of multimorbidity and frailty in the derivation and multi-center validation cohort. Continuous variables are reported as median (25<sup>th</sup>, 75<sup>th</sup> percentile); % missing. Categorical variables are reported as n (%); % missing. FEV<sub>1</sub> = Forced expiratory volume in the first second.

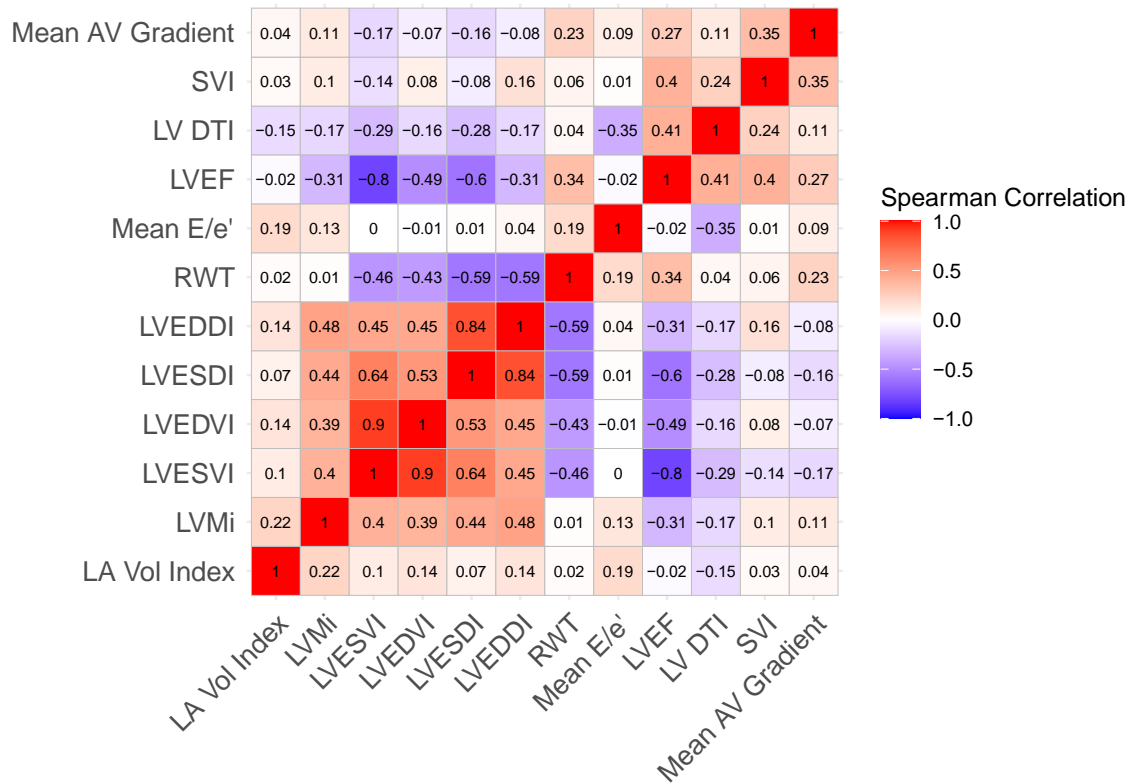
<b>Characteristic</b>	<b>Derivation N = 519</b>	<b>Multi-center Validation N = 286</b>
Kansas City Cardiomyopathy Questionnaire Score	47 (29, 66); 6.2%	46 (30, 64); 6.6%
Gait speed average (m/s)	0.72 (0.55, 0.88); 12%	0.67 (0.51, 0.84); 21%
Grip strength average (kg)	19 (14, 26); 8.5%	19 (13, 27); 12%
Mini-Cog total score	3.00 (2.00, 5.00); 6.7%	3.00 (2.00, 4.00); 7.3%
Percent predicted FEV <sub>1</sub>	82 (67, 102); 21%	75 (60, 90); 33%
Hemoglobin (mg/dL)	12.20 (10.90, 13.30); 0.6%	12.50 (11.20, 13.72); 0.7%
Platelets (per liter)	199 (158, 244); 0.8%	190 (156, 239); 0.7%
Uric acid (mg/dL)	6.30 (5.00, 7.70); 2.9%	6.30 (5.10, 8.00); 3.5%

**Figure S1: CONSORT diagram**



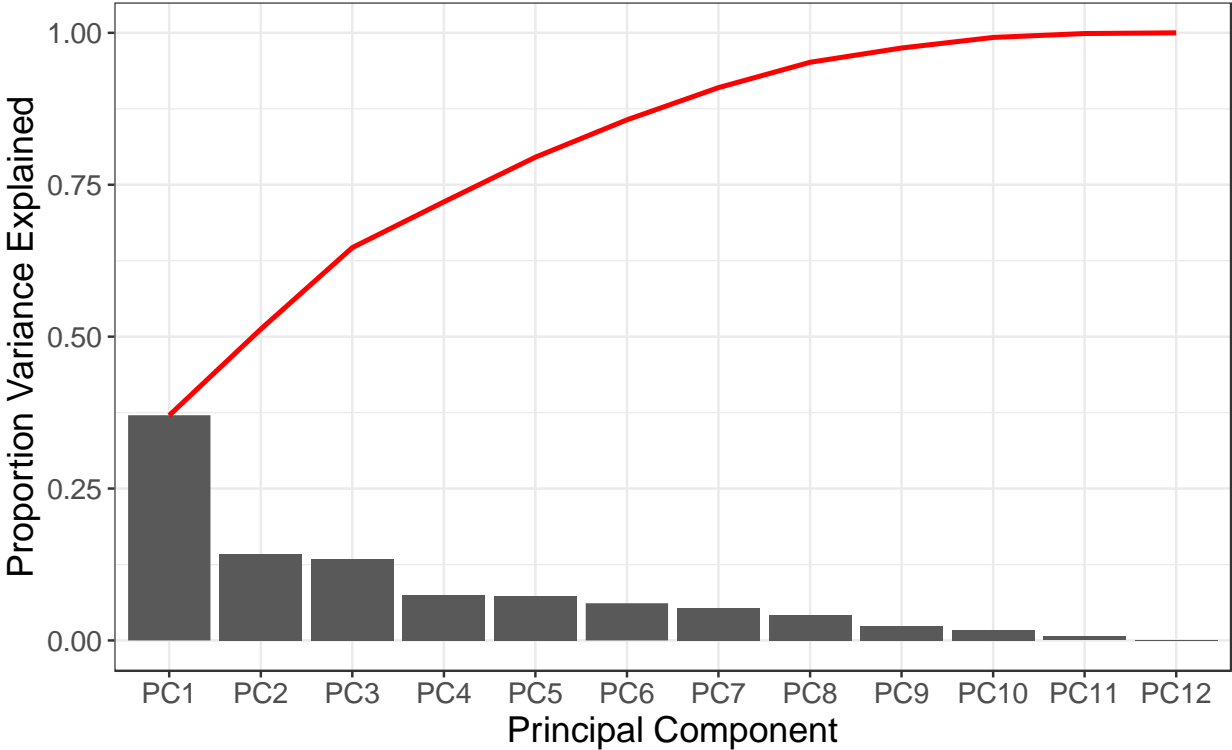
LC-MS = liquid chromatography – mass spectrometry

**Figure S2: Correlation of echocardiographic parameters**



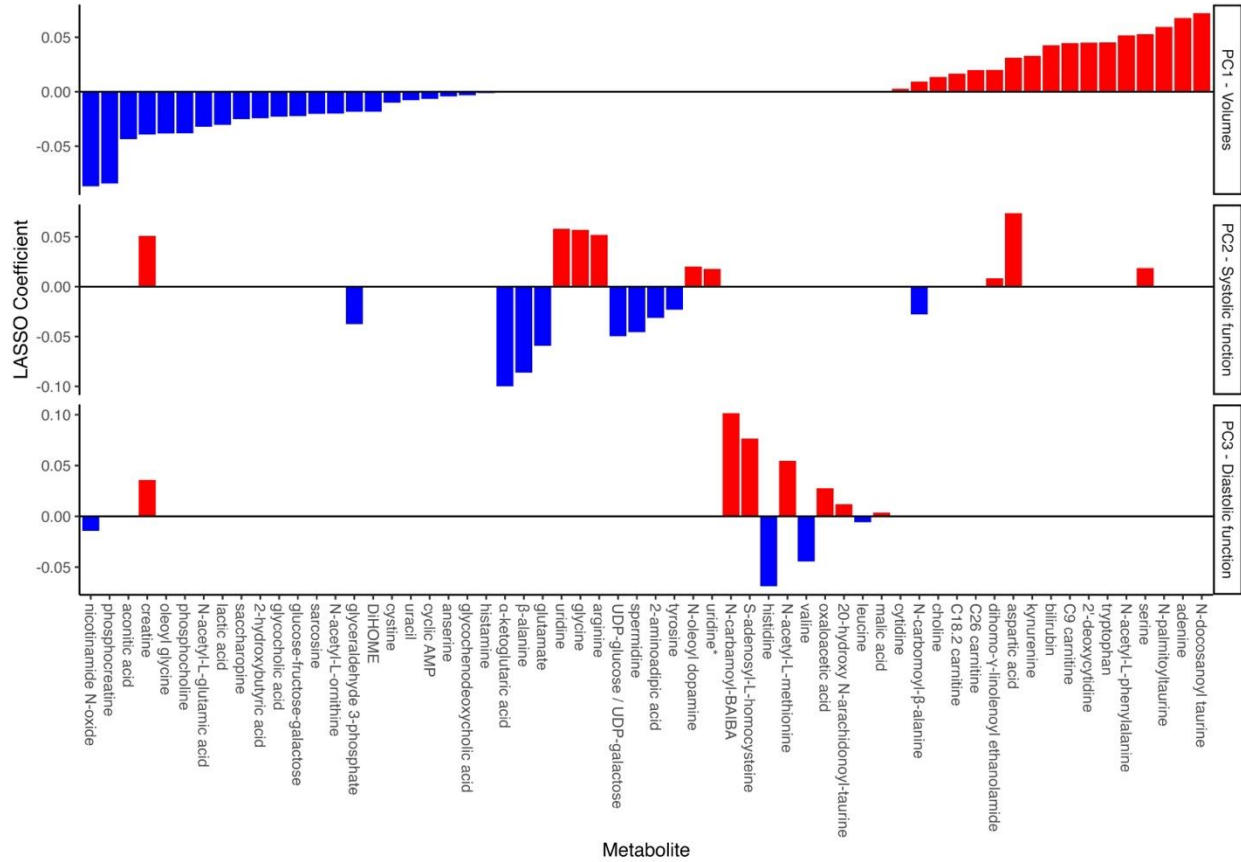
Spearman correlation across 12 echocardiographic parameters used in principal components analysis. LA Vol = left atrium volume index; LVMi = left ventricular mass index; LVESVI = left ventricular end-systolic volume index; LVEDVI = left ventricular end-diastolic volume index; LVESDI = left ventricular end-systolic diameter index; LVEDDI = left ventricular end-diastolic diameter index; RWT = relative wall thickness; LVEF = left ventricular ejection fraction; LV DTI = tissue Doppler S velocity of lateral mitral annulus; SVI = stroke volume index; AV = aortic valve.

**Figure S3: Scree plot from principal components analysis**



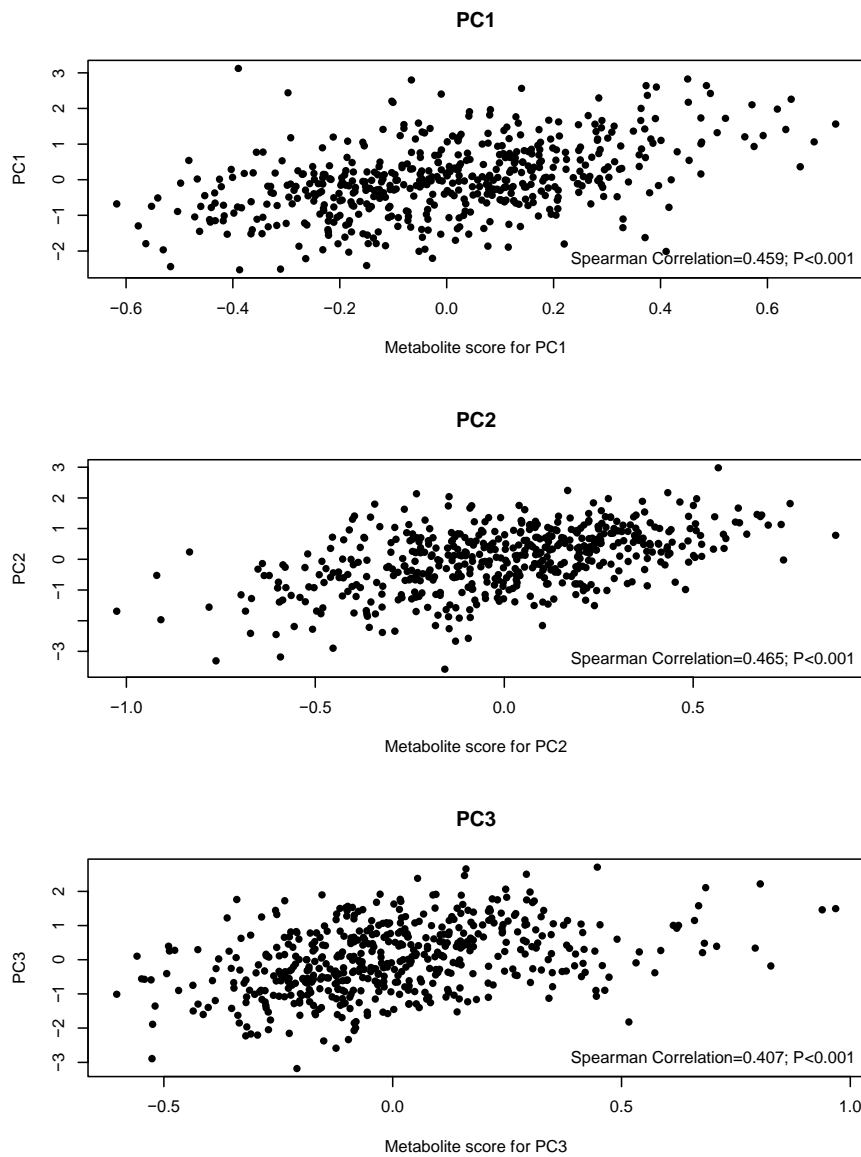
Scree plot demonstrating proportion of variance explained (bars) and cumulative variance explained (red line), suggesting 3 principal components. PC = principal component.

**Figure S4: Relationship between circulating metabolites and echocardiographic PCs.**



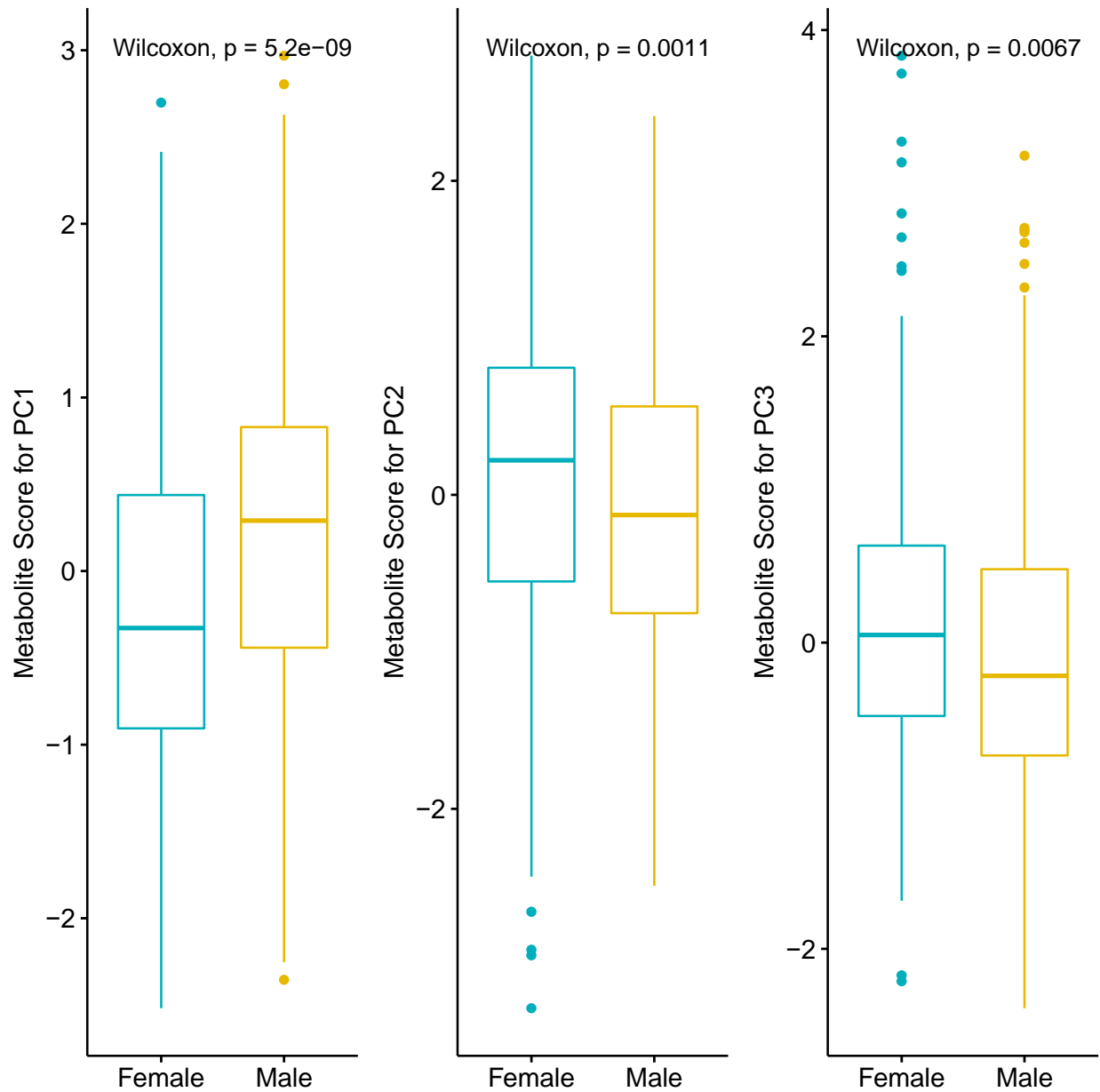
All 60 metabolites selected by LASSO are shown with their corresponding loading with each echocardiographic PC (see **Table S1** for coefficients). LASSO = least absolute shrinkage and selection operator; PC = principal component. \*HILIC negative ion mode.

**Figure S5: Relation of metabolite scores to parent phenotypes**



Relation of parent phenotype PC scores with metabolite scores. The associations were moderate in magnitude. Reported  $R^2$  is from model optimization at the optimal lambda. PC = principal component.

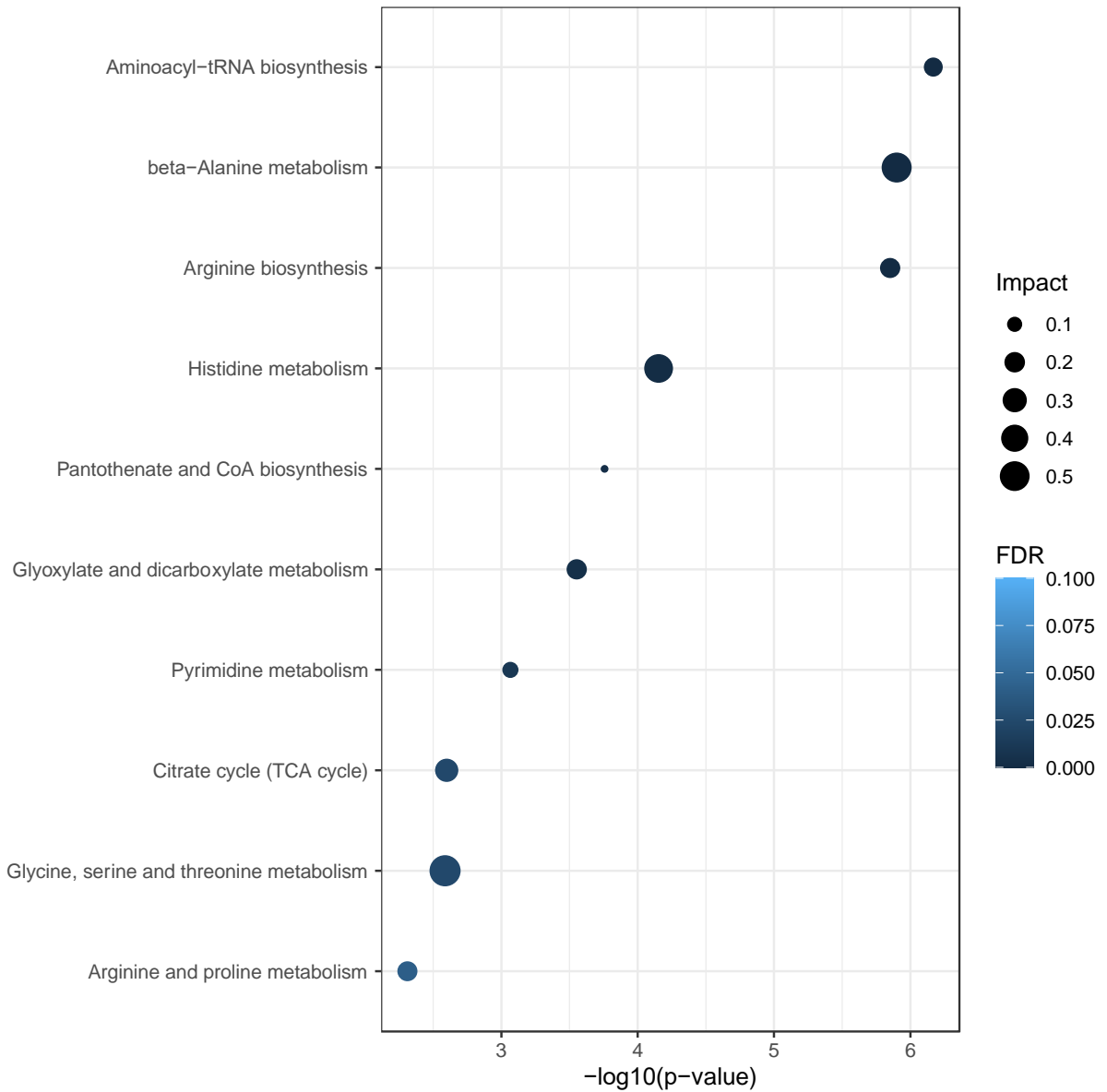
**Figure S6: Comparison of metabolite scores across sex.**



Box-plots for metabolite scores across sex (comparison by Wilcoxon test). While we observed statistically significant differences by sex, there was a broad degree of overlap suggesting no clinically meaningful differences in metabolite score by sex. PC = principal component.



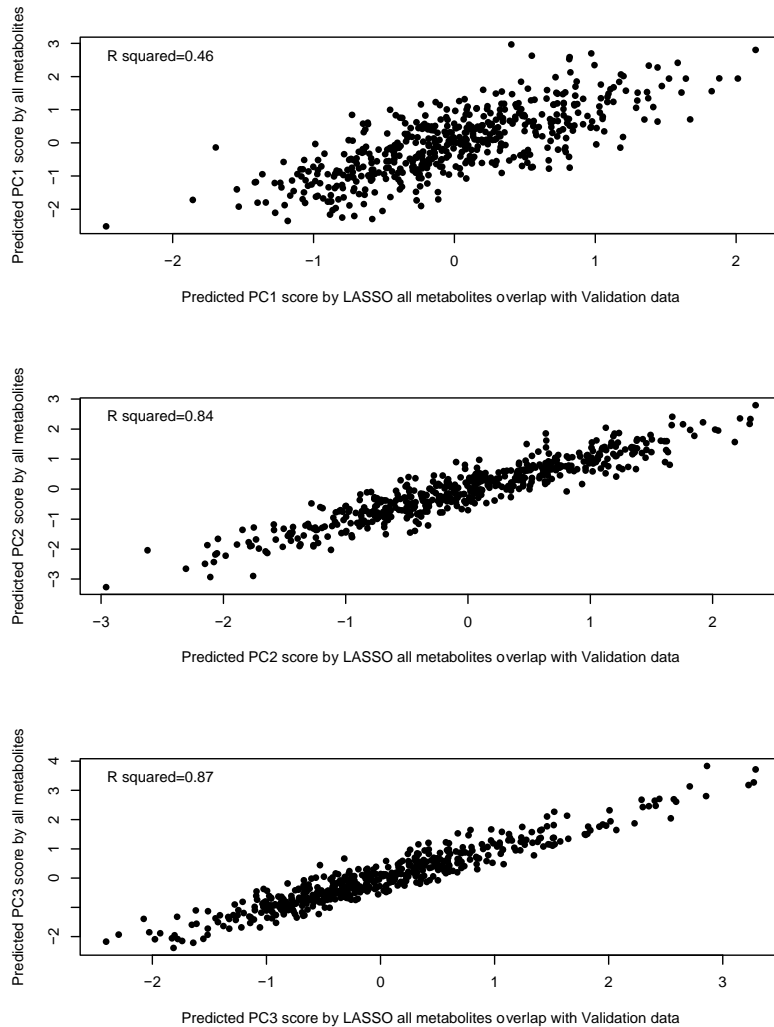
**Figure S7: Pathway analysis of metabolite scores**



Metabolic pathway analysis of 52 metabolites with Human Metabolome Database (HMDB) identifiers (out of 60 metabolites across all identified metabolite scores).

Pathways with an FDR < 0.05 are shown. FDR = false discovery rate.

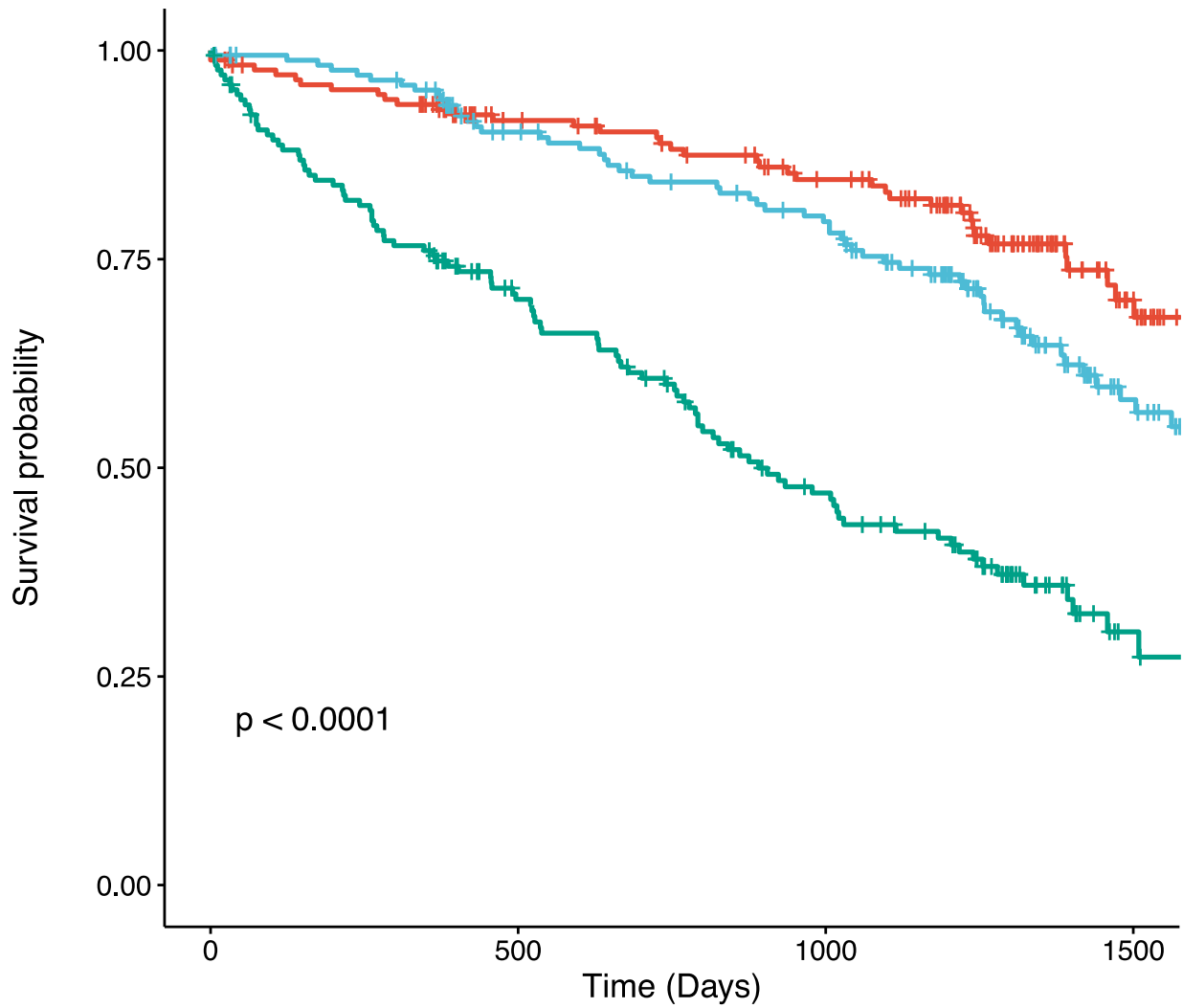
**Figure S8: Correlation of recalibrated metabolite scores with original metabolite scores**



Recalibration of scores for the single-center validation cohort (as described in **Methods** and **Results**). Axes correspond to the full score (Y axis; based on all metabolites in our derivation sample) versus the recalibrated “reduced” score (X-axis; based on metabolites that overlap between the multi-center derivation cohort and the single-center validation cohort).  $R^2$  is moderate to excellent (up to 87% for PC3). PC = principal component.

**Figure S9: Validation of the prognostic utility of the metabolite score for PC3 in separate cohorts.**

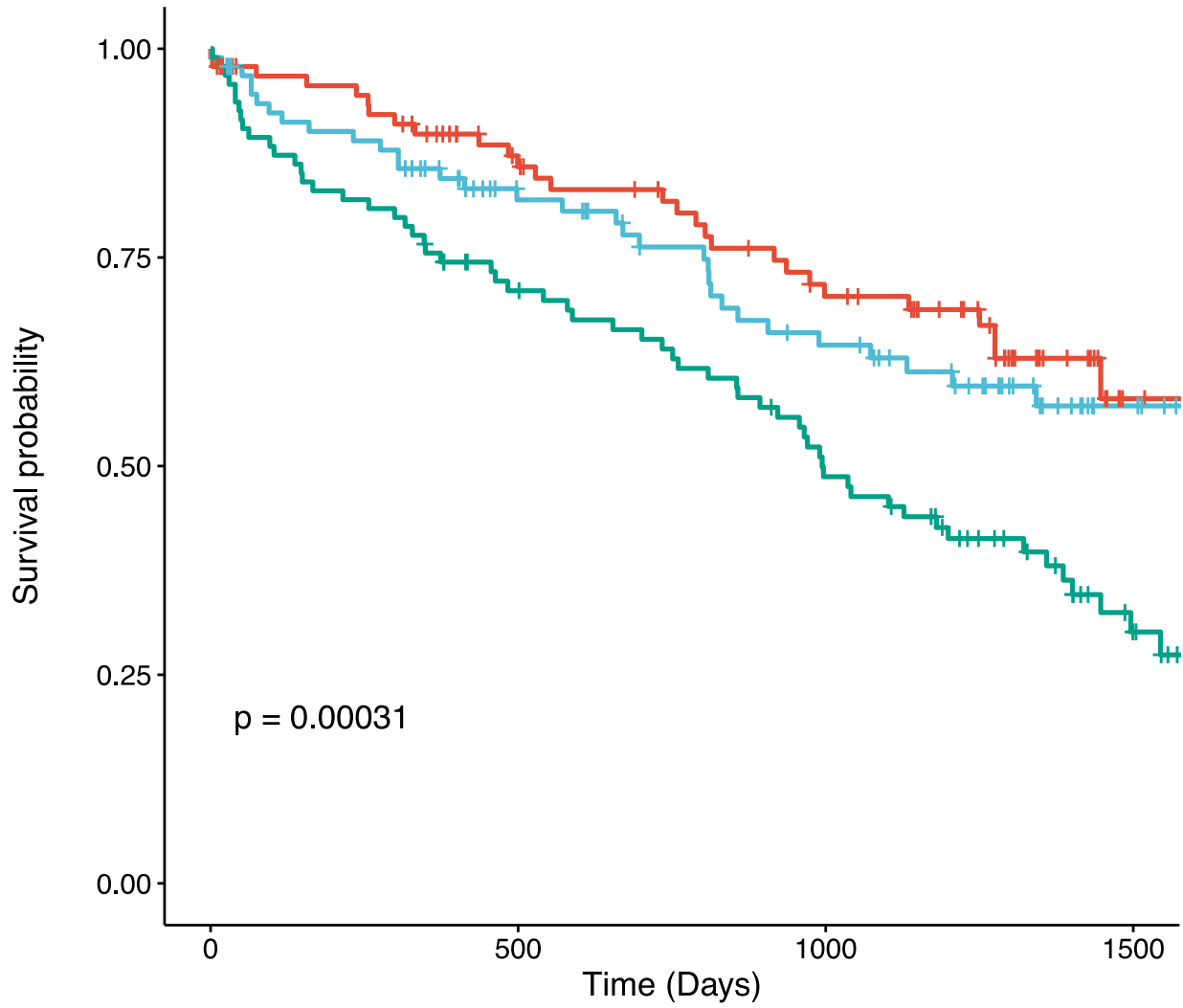
**Figure S9A**



Metabolite-PC3 Score	Number at risk			
	0	500	1000	1500
Tertile 1	173	136	113	35
Tertile 2	172	138	116	38
Tertile 3	171	104	62	10

Time (Days)

Figure S9B

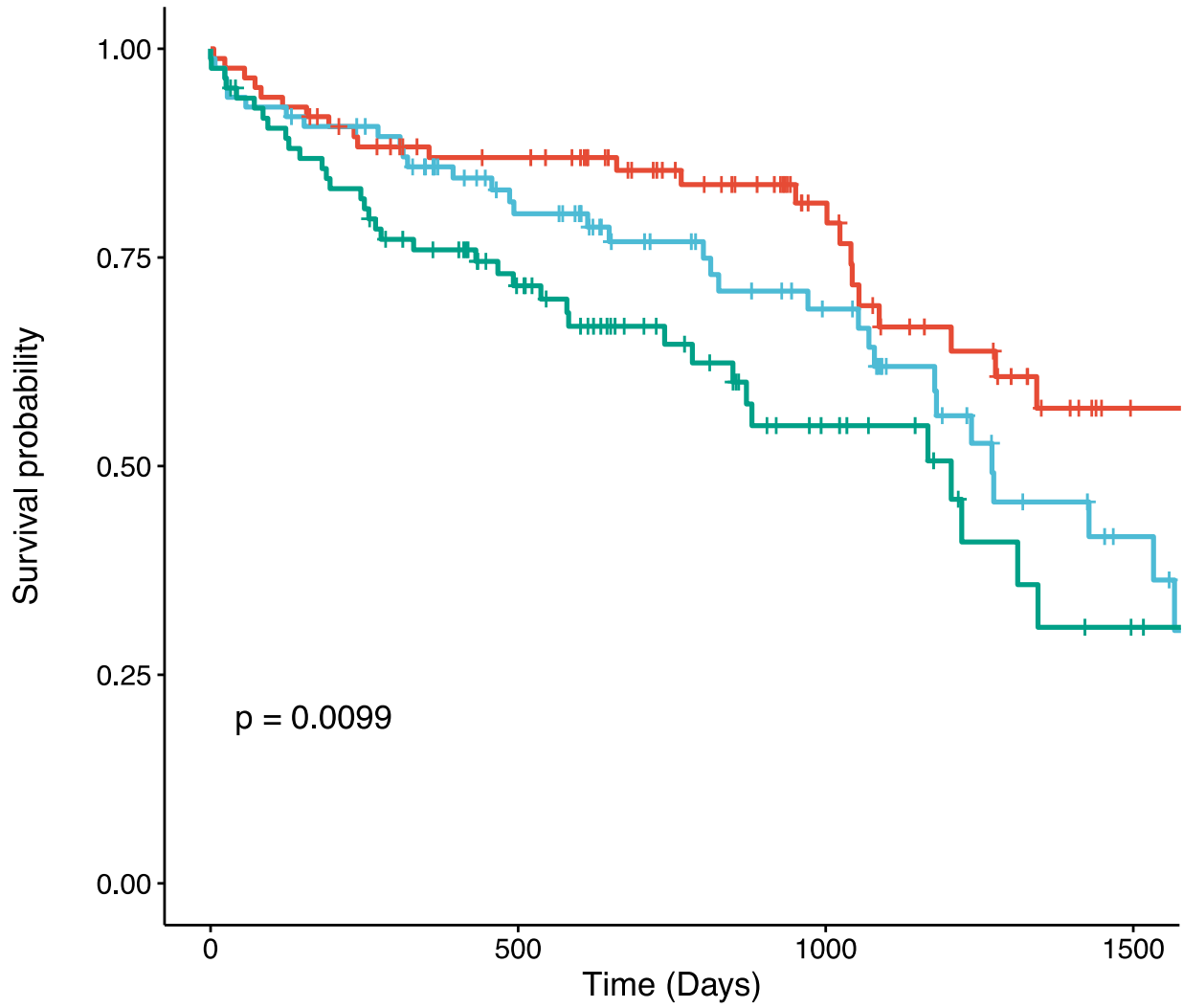


Metabolite-PC3 Score

	Number at risk			
	0	500	1000	1500
Tertile 1	94	65	48	7
Tertile 2	94	61	43	15
Tertile 3	94	62	41	12

Time (Days)

Figure S9C



Metabolite-PC3 Score

	Number at risk			
	0	500	1000	1500
Tertile 1	86	66	34	7
Tertile 2	86	56	31	8
Tertile 3	85	48	17	4

Time (Days)

Kaplan-Meier estimates of survival among the multi-center derivation cohort (A), multi-center validation cohort (B), and single-center validation cohort (C) stratified by tertiles of PC3 metabolite score. For visualization plots are truncated at 1500 days. P-values from logrank tests are reported. PC = principal component.