Blood-based cardiometabolic phenotypes in atrial fibrillation and their associated risk: EAST-AFNET 4 biomolecule study.

Larissa Fabritz[1,2,3,4, 5], Winnie Chua[5], Victor R. Cardoso[5], Christoph Al-Taie[1,2,3], Katrin Borof[1,2], Anna Suling[6], Linda Krause[6], Shinwan Kany[1,3,7,8], Christina Magnussen[1,3,9], Karl Wegscheider[6], Guenter Breithardt[10], Harry Crijns[11], A. John Camm[12], George Gkoutos[5], Patrick T. Ellinor[7,8], Andreas Goette[13], Ulrich Schotten[4,14], Ursula-Henrike Wienhues-Thelen[15], Tanja Zeller[1,2,3], Renate B. Schnabel[1,2,3]Antonia Zapf[6], Paulus Kirchhof[1,2,4,5]

[1]Department of Cardiology, University Heart and Vascular Center Hamburg, University Medical Center Hamburg Eppendorf, Hamburg, Germany

[2]University Center of Cardiovascular Science, University Heart and Vascular Center Hamburg, University Medical Center Hamburg Eppendorf, Hamburg, Germany

[3]German Center for Cardiovascular Research (DZHK), partner site Hamburg/Kiel/Lübeck, Germany

[4]AFNET, Münster, Germany

[5]Institute of Cardiovascular Sciences, University of Birmingham, United Kingdom

[6]Institute of Medical Biometry and Epidemiology, University Medical Center Hamburg Eppendorf, Hamburg, Germany

[7]Cardiovascular Disease Initiative, The Broad Institute of MIT and Harvard, Cambridge, MA, USA

[8]Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA

[9]Center for Population Health Innovation (POINT), University Heart and Vascular Center Hamburg, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[10]University Hospital Münster, Münster, Germany

[11]Department of Cardiology, University Hospital Maastricht, Maastricht, The Netherlands

[12]Clinical Sciences, St George´s University, London, UK

[13]Vincenz-Krankenhaus Paderborn, Germany

[14]Department of Physiology, Maastricht University, Maastricht, The Netherlands

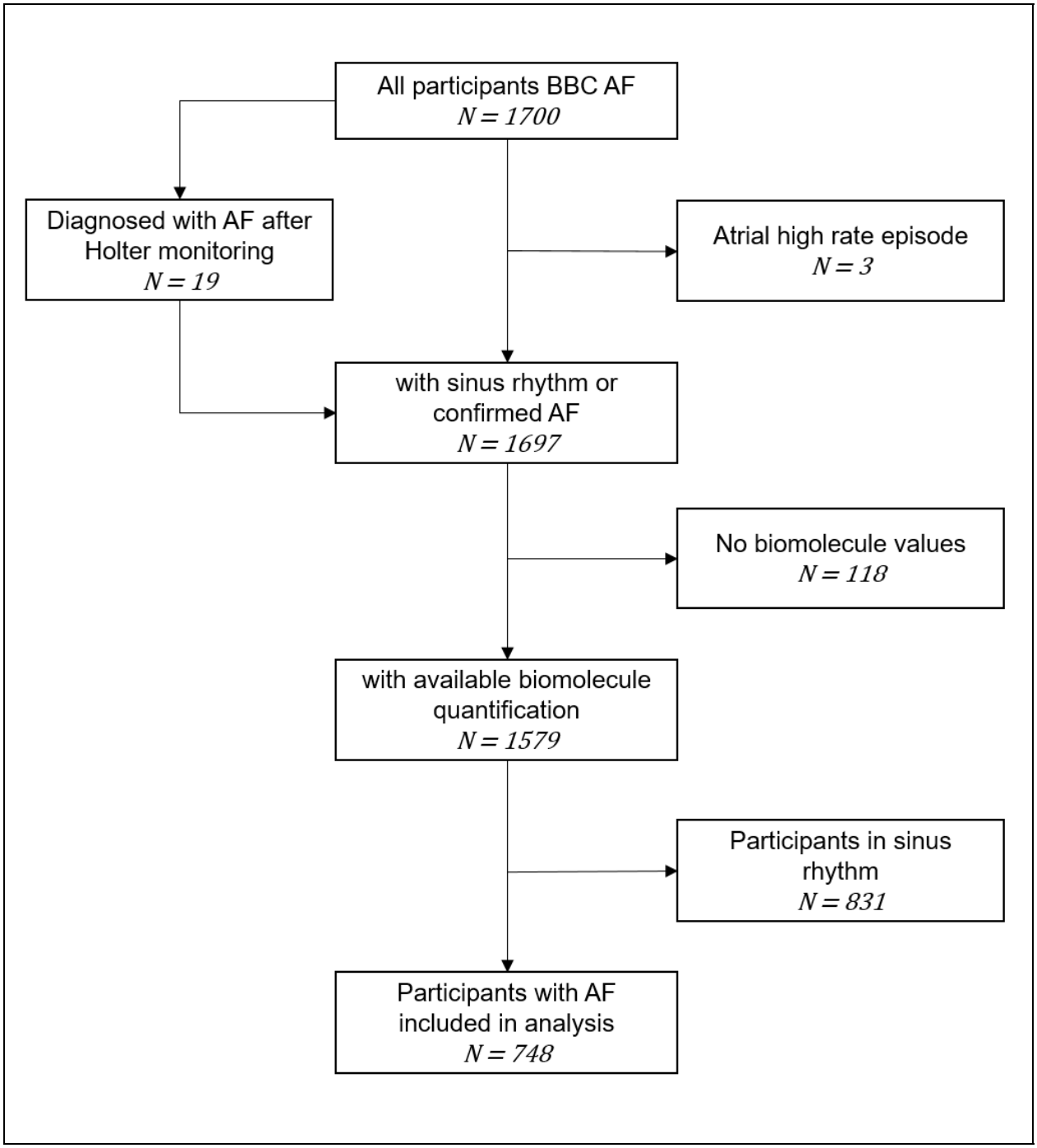
[15]Roche Diagnostics, Penzberg, Germany

Short title: Biomolecule-based patient clusters in atrial fibrillation

Supplementary Table 1: comparison of clinical features in the derivation cohort (EAST-AFNET4) and validation cohort (BBC-AF atrial fibrillation sub-cohort)

	Cohort		**p-value**
	EAST AFNET 4 n = 1,586	**BBC AF** n = 748	
Median age (years (IQR))	71 (66, 76)	73 (64, 80)	<0.001
Female sex (n (%))	713 (45%)	288 (39%)	0.003
Mean Body Mass Index (SD)	29.4 (5.3)	29.9 (6.7)	0.5
Hypertension (n (%))	1,400 (88%)	399 (53%)	<0.001
Diabetes mellitus (n (%))	396 (25%)	181 (24%)	0.700
Heart failure (n (%))	475 (30%)	200 (27%)	0.110
AF pattern			
No history	0 (0%)	47 (6.3%)	
First episode	560 (35%)	46 (6.1%)	
Paroxysmal	590 (37%)	337 (45%)	
Paroxysmal atrial flutter	0 (0%)	7 (0.9%)	
Long-standing persistent	0 (0%)	83 (11%)	
Persistent	436 (27%)	140 (19%)	
Permanent	0 (0%)	88 (12%)	

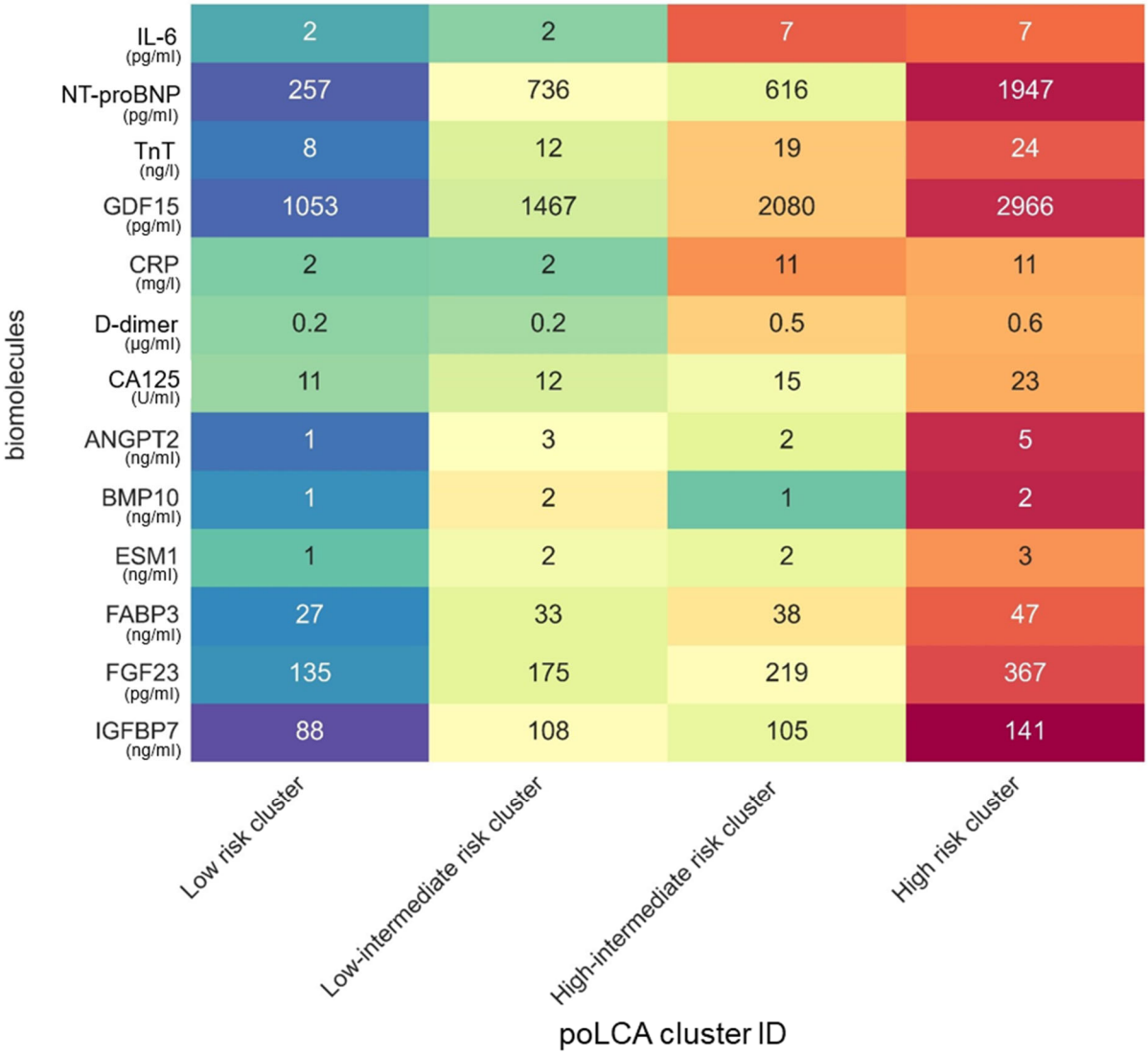
Supplementary Figure 1: PRISMA flow chart of the validation cohort, patients with atrial fibrillation enrolled into BBC-AF



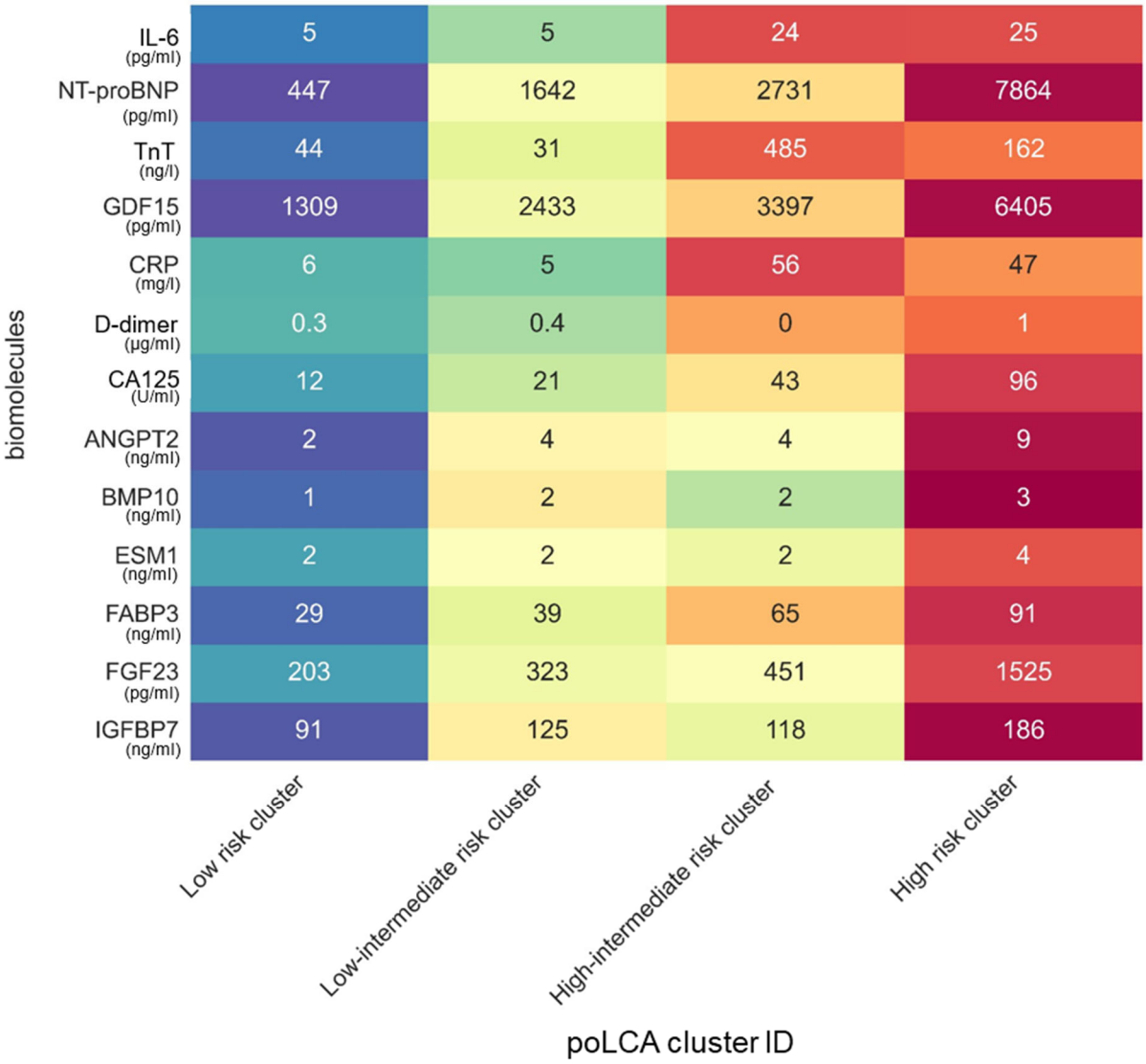
Supplementary Table 2: Distribution of patients and clinical characteristics into the four biomolecule-based clusters in the validation cohort, BBC-AF.

Characteristic	**Cluster in BBC AF**				**p-value**
	Low risk cluster N = 268 (36%)	**Low-intermediate risk cluster** N = 185 (25%)	**High-intermediate risk cluster** N = 123 (16%)	**High risk cluster** N = 172 (23%)	
Age (years)	67 (58, 74)	76 (68, 81)	74 (68, 80)	78 (70, 84)	<0.001
Age ≥ 65 (years)	157 (59%)	153 (83%)	99 (80%)	150 (87%)	<0.001
CHA_2DS_2VASc score	3 (2, 4)	4 (3, 5)	4 (3, 5)	4 (3, 5)	<0.001
Gender					0.070
Female	90 (34%)	81 (44%)	43 (35%)	74 (43%)	
BMI	30 (26, 33)	28 (25, 33)	28 (24, 33)	29 (25, 33)	0.034
Missing	9	10	8	9	
BMI ≥ 30	124 (48%)	70 (40%)	45 (39%)	70 (43%)	0.3
Missing	9	10	8	9	
Arterial hypertension	142 (53%)	96 (52%)	70 (57%)	91 (53%)	0.8
Diabetes mellitus	46 (17%)	37 (20%)	36 (29%)	62 (36%)	<0.001
Severe coronary artery disease	43 (16%)	37 (20%)	44 (36%)	50 (29%)	<0.001
Stable heart failure	33 (12%)	45 (24%)	36 (29%)	86 (50%)	<0.001
Prior stroke or TIA	10 (3.8%)	11 (6.0%)	5 (4.1%)	13 (7.6%)	0.3
Missing	2	1	0	1	
LVEF < 50%	37 (15%)	55 (33%)	51 (44%)	85 (51%)	<0.001
Missing	21	16	7	5	
CAD	51 (19%)	46 (25%)	51 (41%)	51 (30%)	<0.001
Missing	2	3	0	2	
COPD	12 (4.5%)	26 (14%)	20 (16%)	20 (12%)	<0.001
Missing	2	1	1	1	

Supplementary Figure 2A: biomarker signature by poLCA cluster in EAST-AFNET 4. The heatmap illustrates distinct biomolecule patterns, e.g. intermediate concentrations for NT-proBNP and IGFBP7, but high concentrations of BMP10, in the intermediate-low risk cluster, or higher concentrations of IL-6 and CRP in the high-intermediate risk cluster. Low concentrations are coded in blue, high concentrations in red.



Supplementary Figure 2B: mean biomarker signature by poLCA cluster in BBC AF. Distinct patterns that are comparable to the derivation cohort emerge, e.g. intermediate concentrations for NT-proBNP and IGFBP7, but high concentrations of BMP10, in the intermediate-low risk cluster, or higher concentrations of IL-6 and CRP in the high-intermediate risk cluster. Low concentrations are coded in blue, high concentrations in red.

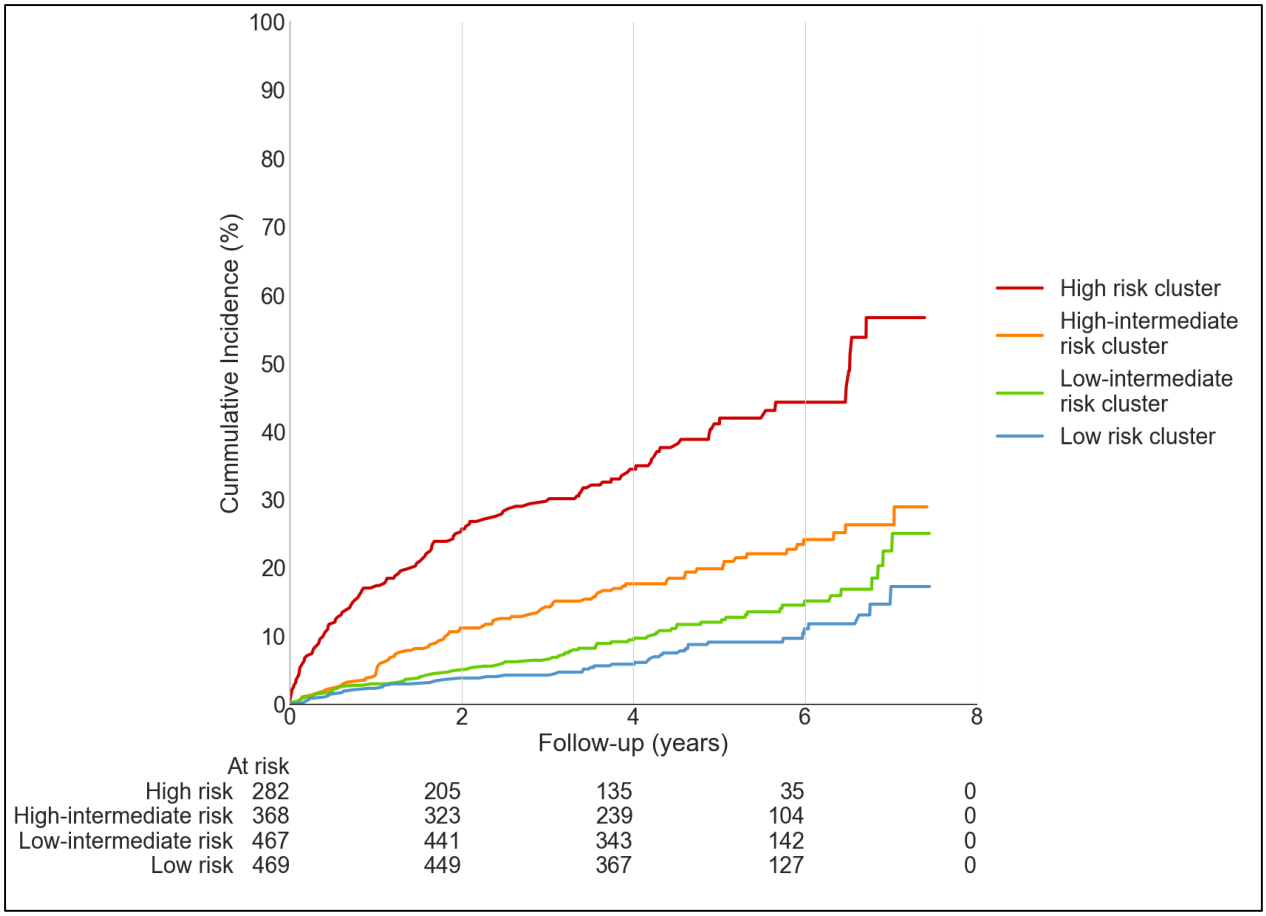


Sensitivity Analysis I: Clustering using *k*-means

Supplementary Table 3: crosstab comparing the partitioning of the EAST-AFNET 4 participants to cluster groups resulting from k-means and poLCA clustering. The adjusted rand index is 0.66133.

	poLCA cluster				
	Low risk cluster	**Low-intermediate risk cluster**	**High-intermediate risk cluster**	**High risk cluster**	**All**
***k*-means cluster**					
Low risk cluster	444	21	4	0	**469**
Low-intermediate risk cluster	34	406	14	13	**467**
High-intermediate risk cluster	24	75	263	6	**368**
High risk cluster	0	10	21	251	**282**
All	**502**	**512**	**302**	**270**	**1586**

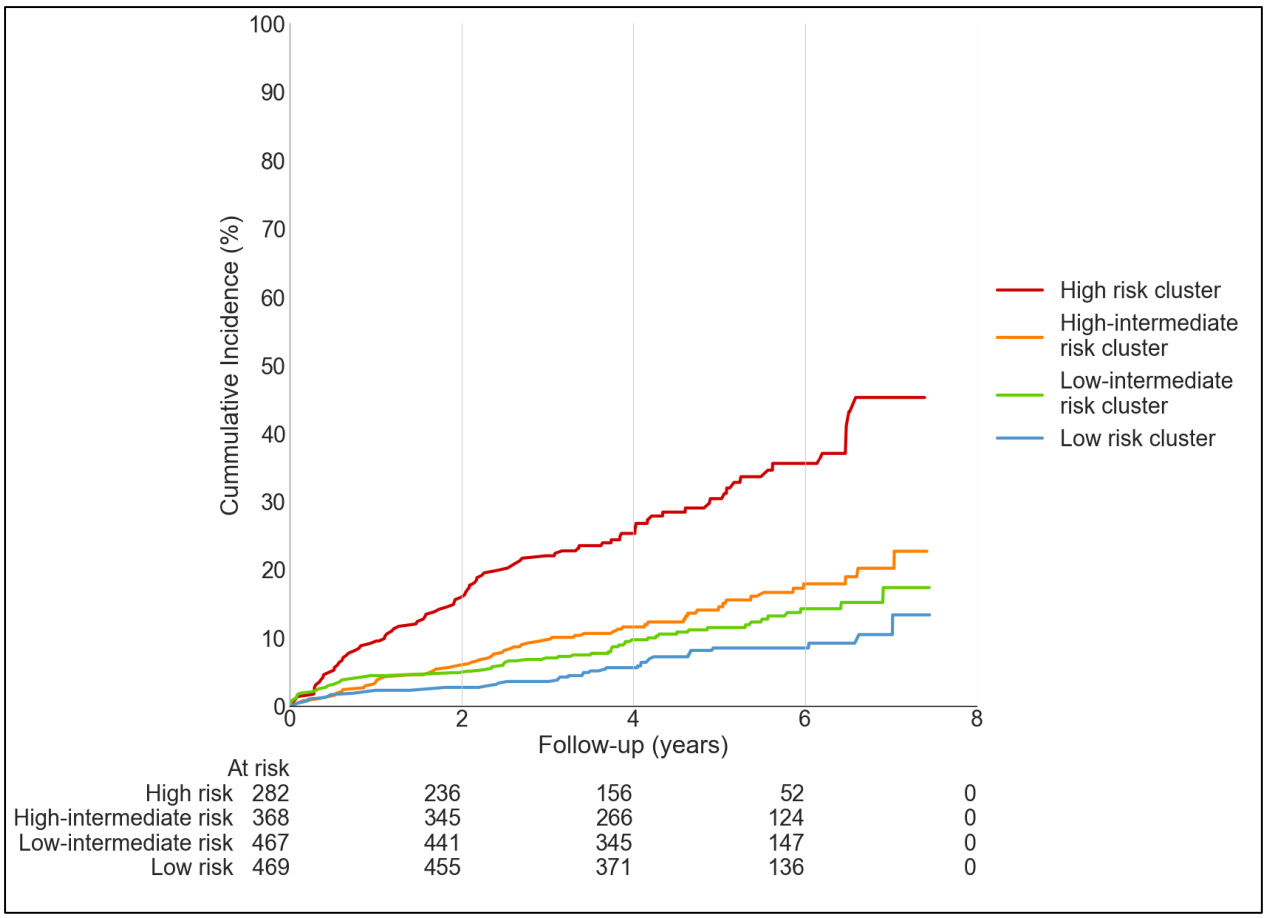
Supplementary Figure 3A: Aalen-Johansen curves for cluster groups from k-means clustering model for the first primary composite outcome in EAST-AFNET 4.



Supplementary Table 4a: efficacy HRs inclusive 95% CI for the first primary outcome in EASTAFNET 4 for cluster group by k-means; on the left from unadjusted Cox PH model, on the right additionally adjusted for age and sex.

Predictors	**HR (95% CI)**	**p**	**HR (95% CI)**	**p**
Low-intermediate risk cluster	1.47 [1.11 – 1.94]	**0.007**	1.26 [0.95 – 1.68]	0.109
High-intermediate risk cluster	2.39 [1.83 – 3.14]	**< 0.001**	1.98 [1.49 – 2.64]	**< 0.001**
High risk cluster	5.78 [4.43 – 7.55]	**< 0.001**	4.64 [3.39 – 6.36]	**< 0.001**
age			1.03 [1.02 – 1.05]	**< 0.001**
male sex			1.33 [1.08 – 1.63]	**0.006**

Supplementary Figure 3B: Aalen-Johansen curves for cluster groups from k-means clustering model for the first primary safety outcome in EAST-AFNET 4.



Supplementary Table 4b: safety HRs inclusive 95% CI for the first primary outcome in EAST-AFNET 4 for cluster group by k-means; on the left from unadjusted Cox PH model, on the right additionally adjusted for age and sex.

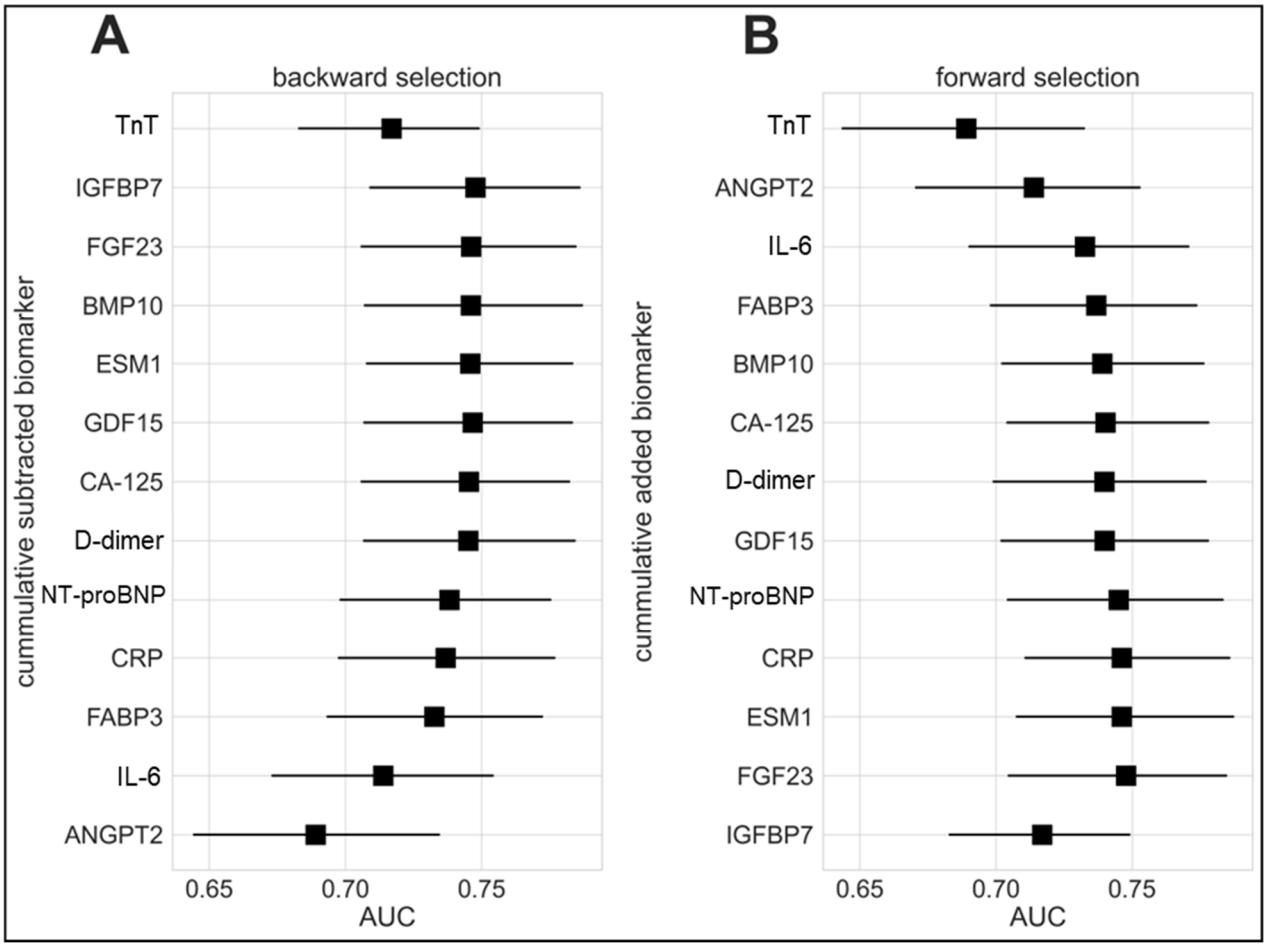
Predictors	**HR (95% CI)**	**p**	**HR (95% CI)**	**p**
Low-intermediate risk cluster	1.53 [1.04 – 2.25]	**0.033**	1.24 [0.85 – 1.83]	0.269
High-intermediate risk cluster	1.96 [1.29 – 2.97]	**0.002**	1.49 [1.02 – 2.19]	**0.041**
High risk cluster	4.75 [3.46 – 6.53]	**< 0.001**	3.38 [2.43 – 4.70]	**< 0.001**
age			1.05 [1.03 – 1.07]	**< 0.001**
male sex			1.19 [0.93 – 1.52]	0.166

Supplementary Table 5: Early rhythm control, the randomized intervention, is effective in all four patient clusters in the EAST-AFNET 4 biomolecule study. Shown are event rates, Hazard ratios (HRs) and their 95% confidence interval (CI) for early rhythm control (ERC) and HRs for the cluster group for each cluster for the first primary outcome and its components stemming from unadjusted Cox models. Unsupervised poLCA model assigns patients to arbitrary cluster group IDs. Hazard ratios give the hazard reduction in patients randomized to early rhythm control with usual care in the low risk cluster set as the reference. All numbers give hazard ratios and their 95% confidence intervals in brackets.

	EAST - Cluster by poLCA			
	Low risk (blue) cluster	**Low-intermediate risk (green) cluster**	**High-intermediate risk (orange) cluster**	**High risk (red) cluster**
First primary outcome	52 (10%)	69 (13%)	76 (25%)	104 (39%)
Early Rhythm Control	0.91 (0.53, 1.57)	0.62 (0.38, 1.00)	0.59 (0.37, 0.95)	0.62 (0.42, 0.92)
Usual care	reference	1.33 (0.93, 1.92)	2.70 (1.89, 3.85)	5.16 (3.68, 7.23)
Death from cv causes	10 (2.0%)	17 (3.3%)	21 (7.0%)	42 (16%)
Early Rhythm Control	0.94 (0.27, 3.27)	0.64 (0.24, 1.69)	0.85 (0.36, 2.03)	0.45 (0.23, 0.86)
Usual care	reference	1.69 (0.77, 3.69)	3.78 (1.78, 8.02)	9.60 (4.81, 19.16)
Stroke	10 (2.0%)	16 (3.1%)	11 (3.6%)	14 (5.2%)
Early Rhythm Control	2.19 (0.57, 8.48)	0.13 (0.03, 0.58)	0.94 (0.29, 3.08)	0.36 (0.11, 1.14)
Usual care	reference	1.58 (0.72, 3.47)	1.96 (0.83, 4.62)	3.12 (1.38, 7.03)
hospitalization due to worsening of heart failure	20 (4.0%)	37 (7.2%)	43 (14%)	65 (24%)
Early Rhythm Control	0.75 (0.31, 1.81)	1.12 (0.58, 2.13)	0.51 (0.27, 0.97)	0.67 (0.41, 1.10)
Usual care	reference	1.84 (1.07, 3.17)	4.01 (2.36, 6.82)	8.07 (4.89, 13.34)
hospitalization due to acute coronary syndrome	17 (3.4%)	14 (2.7%)	18 (6.0%)	14 (5.2%)
Early Rhythm Control	0.82 (0.32, 2.13)	0.52 (0.17, 1.55)	0.41 (0.15, 1.16)	0.69 (0.24, 1.99)
Usual care	reference	0.81 (0.40, 1.64)	1.89 (0.97, 3.66)	1.84 (0.90, 3.73)

Sensitivity Analysis I: Cox proportional hazard models using backward and forward selection of biomolecules.

*Supplementary Figure 4: **A**: starting with all biomarkers from above to below one biomarker is subsequently subtracted from the Cox PH model leading to a change in AUC. **B**: starting with the null-model biomarkers are subsequently added to the model. Forward selection starts with TnT and backward selection never drops TnT. Also for ANGPT2 and IL-6 both model selection methods agree upon that those are important markers for discriminatory power of the models for the first primary composite outcome. Also both agree that IGFBP7 is least important in the linear model.*



Sensitivity Analysis II: Effect of removing/adding biomarkers directly in supervised Cox proportional hazard model (no prior clustering on biomarkers)

*Supplementary Table 6a: different Cox proportional hazard model instances with different sets of biomarkers from **backward selection** for the first primary composite outcome in EAST-AFNET 4. AUC Area under the curve*

predictors	**AUC**	**AUC lower**	**AUC upper**
IL-6, NT-proBNP, TnT, GDF15, CRP, D-dimer, CA125, ANGPT2, BMP10, ESM1, FABP3, FGF23, IGFBP7	0.717	0.683	0.749
IL-6, NT-proBNP, TnT, GDF15, CRP, D-dimer, CA125, ANGPT2, BMP10, ESM1, FABP3, FGF23	0.748	0.709	0.786
IL-6, NT-proBNP, TnT, GDF15, CRP, D-dimer, CA125, ANGPT2, BMP10, ESM1, FABP3	0.746	0.706	0.785
IL-6, NT-proBNP, TnT, GDF15, CRP, D-dimer, CA125, ANGPT2, ESM1, FABP3	0.746	0.707	0.787
IL-6, NT-proBNP, TnT, GDF15, CRP, D-dimer, CA125, ANGPT2, FABP3	0.746	0.708	0.783
IL-6, NT-proBNP, TnT, CRP, D-dimer, CA125, ANGPT2, FABP3	0.747	0.707	0.783
IL-6, NT-proBNP, TnT, CRP, D-dimer, ANGPT2, FABP3	0.745	0.706	0.782
IL-6, NT-proBNP, TnT, CRP, ANGPT2, FABP3	0.745	0.707	0.784
IL-6, TnT, CRP, ANGPT2, FABP3	0.738	0.698	0.775
IL-6, TnT, ANGPT2, FABP3	0.737	0.698	0.777
IL-6, TnT, ANGPT2	0.733	0.693	0.772
TnT, ANGPT2	0.714	0.673	0.754
TnT	0.689	0.644	0.734

*Supplementary Table 6b: different Cox proportional model instances with different sets of biomarkers from **forward selection** for the first primary composite outcome in EAST-AFNET 4. AUC Area under the curve*

predictors	**AUC**	**AUC lower**	**AUC upper**
TnT	0.689	0.644	0.732
TnT, ANGPT2	0.714	0.671	0.753
IL-6, TnT, ANGPT2	0.733	0.690	0.771
IL-6, TnT, ANGPT2, FABP3	0.737	0.698	0.774
IL-6, TnT, ANGPT2, BMP10, FABP3	0.739	0.702	0.776
IL-6, TnT, CA125, ANGPT2, BMP10, FABP3	0.740	0.704	0.778
IL-6, TnT, D-dimer, CA125, ANGPT2, BMP10, FABP3	0.740	0.699	0.777
IL-6, TnT, GDF15, D-dimer, CA125, ANGPT2, BMP10, FABP3	0.740	0.702	0.778
IL-6, NT-proBNP, TnT, GDF15, D-dimer, CA125, ANGPT2, BMP10, FABP3	0.745	0.704	0.783
IL-6, NT-proBNP, TnT, GDF15, CRP, D-dimer, CA125, ANGPT2, BMP10, FABP3	0.746	0.711	0.786
IL-6, NT-proBNP, TnT, GDF15, CRP, D-dimer, CA125, ANGPT2, BMP10, ESM1, FABP3	0.746	0.708	0.787
IL-6, NT-proBNP, TnT, GDF15, CRP, D-dimer, CA125, ANGPT2, BMP10, ESM1, FABP3, FGF23	0.748	0.705	0.785
IL-6, NT-proBNP, TnT, GDF15, CRP, D-dimer, CA125, ANGPT2, BMP10, ESM1, FABP3, FGF23, IGFBP7	0.717	0.683	0.749

Sensitivity Analysis III: Hazard ratios for patients groups defined by quartiles of individual biomolecules.

Supplementary Table 7: Comparison of hazard ratios for the four biomolecule cluster groups and for quartiles of patients defined by concentrations of individual biomolecules. For this analysis, patients were grouped into four equal groups by biomolecule quartiles. Shown are hazard ratios and the 95% confidence intervals for the primary outcome. Hazard ratios for the biomolecules TnT and NT-proBNP are shown in Table 4 alongside other established risk estimators.

	Low risk (blue) cluster / lowest quartile	**Low-intermediate risk (green) cluster / second-lowest quartile**	**High-intermediate risk (orange) cluster / second-highest quartile**	**High risk (red) cluster/ highest quartile**
EAST-AFNET 4 (derivation data set)				
Biomolecule clusters	1 (reference)	1.3 [0.9, 1.9]	2.7 [1.9, 3.6]	5.2 [3.7, 7.2]
IL-6	1 (reference)	1.5 [0.9 – 2.5]	1.8 [1.2 – 2.7]	3.7 [2.5 – 5.6]
IGFBP7	1 (reference)	1.0 [0.7 – 1.4]	1.3 [0.9 – 1.8]	2.9 [2.2 – 3.8]
CRP	1 (reference)	1.4 [0.9 – 2.1]	1.4 [0.9 – 1.9]	2.4 [1.7 – 3.3]
GDF15	1 (reference)	1.8 [1.4 – 2.4]	2.6 [1.9 – 3.6]	4.2 [3.2 – 5.6]
FGF23	1 (reference)	1.1 [0.8 – 1.5]	1.7 [1.3 – 2.3]	2.3 [1.7 – 3.2]
FABP3	1 (reference)	1.1 [0.8 – 1.5]	1.5 [1.0 – 2.2]	2.8 [1.9 – 4.1]
ESM1	1 (reference)	1.2 [0.9 – 1.6]	1.1 [0.8 – 1.6]	1.9 [1.5 – 2.7]
D-dimer	1 (reference)	1.2 [0.8 – 1.9]	2.3 [1.5 – 3.5]	2.6 [1.7 – 3.9]
CA125 (quartiles)	1 (reference)	0.8 [0.5 – 1.0]	0.8 [0.6 – 1.0]	1.4 [1.0 – 1.9]
BMP10	1 (reference)	1.3 [0.9 – 1.9]	1.8 [1.3 – 2.5]	1.9 [1.3 – 2.9]
ANGPT2	1 (reference)	1.4 [0.9 – 2.2]	2.1 [1.5 – 2.9]	2.5 [1.9 – 3.3]
Serum creatinine	1 (reference)	1.20 [0.85 -1.69]	1.21 [0.85 – 1.73]	1.85 [1.35 –2.55]
BBC-AF (validation data set)				
Biomolecule clusters BBC-AF (validation)	1 (reference)	4.0 [2.3 – 7.0]	8.3 [4.80 – 14.4]	14.1 [8.4 – 23.7]
IL-6	1 (reference)	2.4 [1.3 – 4.5]	5.1 [2.8 – 9.1]	8.4 [4.8 – 14.8]
IGFBP7	1 (reference)	1.7 [1.0 – 3.0]	2.7 [1.6 – 4.5]	7.1 [4.3 – 11.4]
CRP	1 (reference)	1.5 [0.9 – 2.5]	2.4 [1.5 – 3.8]	3.6 [2.3 – 5.6]
GDF15	1 (reference)	1.8 [1.0 – 3.4]	4.2 [2.4 – 7.3]	10.6 [6.2 – 18.1]
FGF23	1 (reference)	1.3 [0.8 – 2.2]	2.2 [1.4 – 3.7]	5.3 [3.4 – 8.4]
FABP3	1 (reference)	1.8 [1.0 – 3.2]	4.2 [2.5 – 7.0]	7.3 [4.4 – 12.0]
ESM1	1 (reference)	1.5 [1.0 – 2.4]	1.8 [1.1 – 2.8]	3.2 [2.1 – 5.0]
D-dimer	1 (reference)	2.0 [1.2 – 3.4]	3.1 [1.9 – 5.0]	4.8 [3.0 – 7.7]
CA125	1 (reference)	1.6 [1.0 – 2.8]	2.1 [1.3 – 3.5]	6.0 [3.8 – 9.4]
BMP10	1 (reference)	2.5 [1.4 – 4.5]	4.3 [2.5 – 7.4]	8.0 [4.8 – 13.5]
ANGPT2	1 (reference)	2.3 [1.3 – 4.2]	5.1 [3.0 – 8.9]	8.8 [5.1 – 15.0]
Serum creatinine	1 (reference)	0.98 [0.61 -1.57]	1.55 [1.00 – 2.40]	2.76 [1.84 –4.13]

Supplementary Table 8a: Shown is the unique, common and total explained variance for each variable for assignment of patients into the high risk cluster, measured by pseudo R^2 in descending order of percentage variance explained for High risk cluster by poLCA in EAST-AFNET 4

High risk cluster	**Unique**	**Common**	**Total**
IGFBP7	0.014	0.284	0.298
NT-proBNP	0.009	0.249	0.258
BMP10	0.023	0.218	0.241
GDF15	0.007	0.221	0.228
ANGPT2	0.007	0.220	0.227
FGF23	0.006	0.178	0.184
TnT	0.002	0.156	0.157
FABP3	0.003	0.148	0.151
IL-6	0.001	0.138	0.138
ESM1	0.001	0.091	0.092
CRP	0.003	0.071	0.073
CA125	0.004	0.068	0.071
D-dimer	0.002	0.065	0.068

Supplementary Table 8b: Shown is the unique, common and total explained variance for each variable for assignment of patients into the high-intermediate risk cluster, measured by pseudo R^2 in descending order of percentage variance explained for high risk cluster by poLCA in EAST-AFNET 4

High-intermediate risk cluster	**Unique**	**Common**	**Total**
IL-6	0.059	0.124	0.183
CRP	0.010	0.098	0.108
TnT	0.010	0.046	0.056
GDF15	0.010	0.047	0.056
D-dimer	0.006	0.041	0.047
BMP10	0.027	0.014	0.041
FABP3	0.000	0.026	0.027
FGF23	0.000	0.009	0.009
CA125	0.000	0.001	0.002
ANGPT2	0.003	-0.003	0.000
NT-proBNP	0.002	-0.002	0.000
ESM1	0.000	0.000	0.000
IGFBP7	0.007	-0.007	0.000

Supplementary Table 8c: Shown is the unique, common and total explained variance for each variable for assignment of patients into the Low-intermediate high-risk cluster, measured by pseudo R^2 in descending order of percentage variance explained for high risk cluster by poLCA in EAST-AFNET 4

Low-intermediate risk cluster	Unique	Common	Total
IL-6	0.033	0.035	0.068
CRP	0.011	0.048	0.059
BMP10	0.005	0.033	0.038
D-dimer	0.013	0.016	0.029
IGFBP7	0.015	0.005	0.019
NT-proBNP	0.008	0.008	0.016
ANGPT2	0.006	0.007	0.013
GDF15	0.002	0.002	0.004
ESM1	0.001	0.002	0.004
CA125	0.002	0.001	0.003
TnT	0.000	0.002	0.002
FABP3	0.000	0.000	0.000
FGF23	0.000	0.000	0.000

Supplementary Table 8d: Shown is the unique, common and total explained variance for each variable for assignment of patients into the low risk cluster, measured by pseudo R^2 in descending order of percentage variance explained for High risk cluster by poLCA in EAST-AFNET 4

Low risk cluster	Unique	Common	Total
IGFBP7	0.022	0.316	0.338
NT-proBNP	0.017	0.274	0.291
GDF15	0.011	0.260	0.271
ANGPT2	0.009	0.230	0.239
TnT	0.010	0.217	0.227
FABP3	0.003	0.187	0.190
BMP10	0.003	0.175	0.178
FGF23	0.004	0.169	0.172
IL-6	0.002	0.157	0.159
ESM1	0.003	0.091	0.094
CRP	0.000	0.063	0.063
D-dimer	0.000	0.049	0.049
CA125	0.000	0.039	0.039

Supplementary Table 9a: four levels of dominance and resulting percentage relative importance for each biomolecule and poLCA high risk cluster.

High risk cluster					
	Interactional Dominance	**Individual Dominance**	**Average Partial Dominance**	**Total Dominance**	**Percentage Relative Importance**
NT-proBNP	0.048	0.322	0.112	0.123	17.004
IGFBP7	0.045	0.328	0.110	0.121	16.799
BMP10	0.038	0.295	0.108	0.117	16.166
ANGPT2	0.034	0.254	0.084	0.093	12.874
GDF15	0.019	0.194	0.052	0.061	8.384
FGF23	0.012	0.170	0.040	0.048	6.641
FABP3	0.015	0.144	0.041	0.047	6.441
TnT	0.005	0.121	0.025	0.031	4.295
ESM1	0.006	0.067	0.019	0.022	2.999
CA 125	0.001	0.079	0.013	0.018	2.436
CRP	0.003	0.061	0.012	0.015	2.070
IL-6	0.000	0.084	0.010	0.015	2.043
D-dimer	0.007	0.047	0.011	0.013	1.847

Supplementary Table 9b: four levels of dominance and resulting percentage relative importance for each biomolecule and poLCA high-intermediate risk cluster.

High-intermediate risk cluster					
	Interactional Dominance	**Individual Dominance**	**Average Partial Dominance**	**Total Dominance**	**Percentage Relative Importance**
BMP10	0.066	0.073	0.086	0.083	28.190
IL-6	0.051	0.100	0.079	0.078	26.356
CRP	0.004	0.062	0.031	0.032	10.709
TnT	0.016	0.033	0.024	0.024	8.265
NT-proBNP	0.016	0.008	0.021	0.020	6.598
D-dimer	0.005	0.027	0.013	0.014	4.569
GDF15	0.010	0.017	0.013	0.013	4.282
IGFBP7	0.011	0.004	0.012	0.011	3.707
ANGPT2	0.001	0.007	0.009	0.008	2.815
FGF23	0.006	0.003	0.005	0.005	1.785
FABP3	0.000	0.010	0.004	0.004	1.507
ESM1	0.001	0.001	0.003	0.003	0.901
CA 125	0.000	0.002	0.001	0.001	0.317

Supplementary Table 9c: four levels of dominance and resulting percentage relative importance for each biomolecule and poLCA low-intermediate risk cluster.

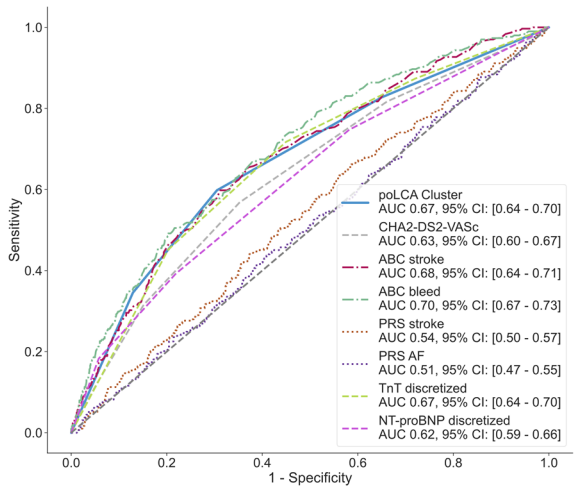
Low-intermediate risk cluster					
	Interactional Dominance	**Individual Dominance**	**Average Partial Dominance**	**Total Dominance**	**Percentage Relative Importance**
IL-6	0.021	0.082	0.045	0.046	34.305
CRP	0.020	0.074	0.041	0.042	31.019
D-dimer	0.006	0.027	0.012	0.012	9.333
TnT	0.003	0.020	0.006	0.007	5.368
GDF15	0.002	0.021	0.006	0.007	5.285
BMP10	0.002	0.001	0.005	0.005	3.428
FGF23	0.002	0.011	0.003	0.004	2.914
ANGPT2	0.005	0.000	0.004	0.004	2.846
FABP3	0.000	0.011	0.002	0.003	1.964
NT-proBNP	0.001	0.001	0.002	0.002	1.287
CA 125	0.000	0.006	0.001	0.001	1.083
IGFBP7	0.000	0.002	0.001	0.001	0.784
ESM1	0.000	0.001	0.000	0.001	0.382

Supplementary Table 9d: four levels of dominance and resulting percentage relative importance for each biomolecule and poLCA low risk cluster.

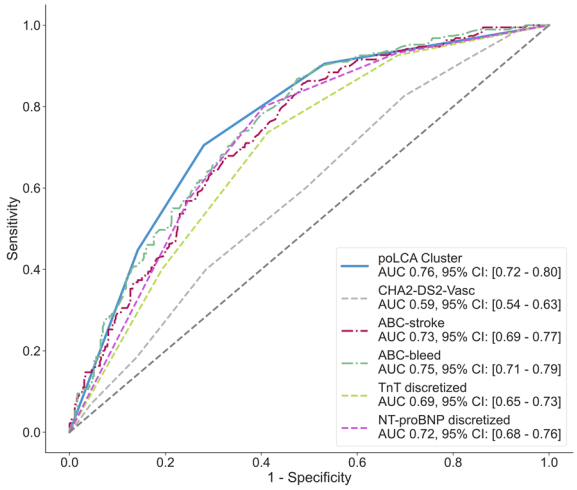
Low risk cluster					
	Interactional Dominance	**Individual Dominance**	**Average Partial Dominance**	**Total Dominance**	**Percentage Relative Importance**
IGFBP7	0.058	0.291	0.112	0.122	17.591
NT-proBNP	0.039	0.256	0.094	0.102	14.759
ANGPT2	0.043	0.223	0.086	0.093	13.486
GDF15	0.028	0.216	0.066	0.074	10.774
TnT	0.028	0.180	0.059	0.066	9.527
BMP10	0.034	0.158	0.058	0.064	9.292
FABP3	0.017	0.144	0.042	0.048	6.906
FGF23	0.018	0.123	0.033	0.039	5.639
IL-6	0.012	0.121	0.031	0.037	5.326
ESM1	0.009	0.057	0.016	0.018	2.673
CA 125	0.004	0.043	0.009	0.011	1.633
CRP	0.002	0.044	0.009	0.011	1.558
D-dimer	0.000	0.031	0.004	0.006	0.837

*Supplementary Figure 5: Receiver Operation Characteristics Curves for different Cox PH models for the **first primary composite outcome**. All models were fitted against the whole EAST-AFNET 4 (train-) dataset and used to make predictions against the whole EAST-AFNET 4 (train-) dataset. **Plot A** shows model performances / discriminatory powers on EAST-AFNET 4 (train-) dataset as Area under the ROC curve (AUC) at two years. **Plot B** shows model performances / discriminatory powers on BBC-AF (test-) dataset as Area under the ROC curve (AUC) at two years.*

A

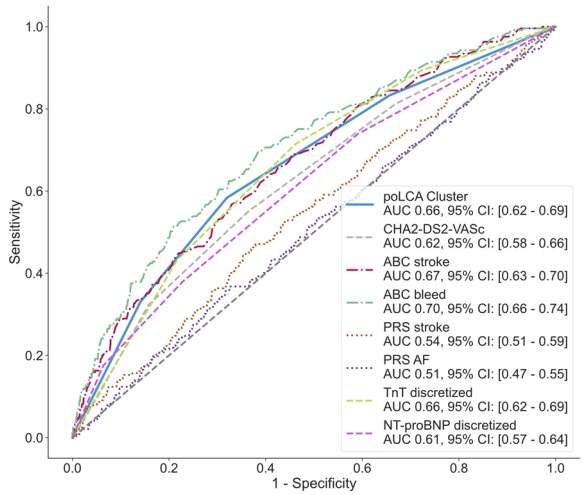


B

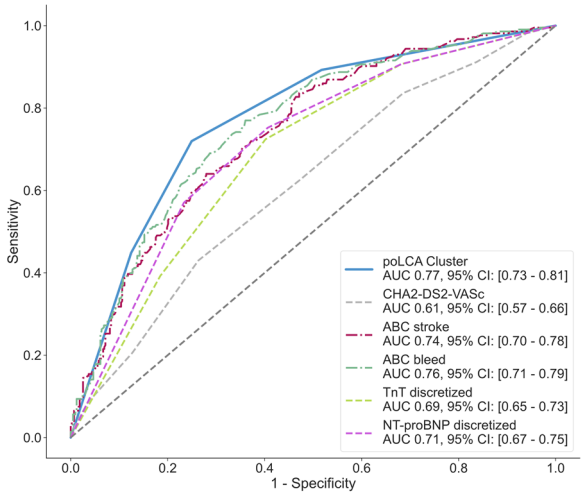


*Supplementary Figure 6: Receiver Operation Characteristics Curves for different Cox PH models for the **safety outcome**. All models were fitted against the whole EAST-AFNET 4 (train-) dataset and used to make predictions against the whole EAST-AFNET 4 (train-) dataset. **Plot A** shows model performances / discriminatory powers on EAST-AFNET 4 (train-) dataset as Area under the ROC curve (AUC) at two years. **Plot B** shows model performances / discriminatory powers on BBC-AF (test-) dataset as Area under the ROC curve (AUC) at two years.*

A

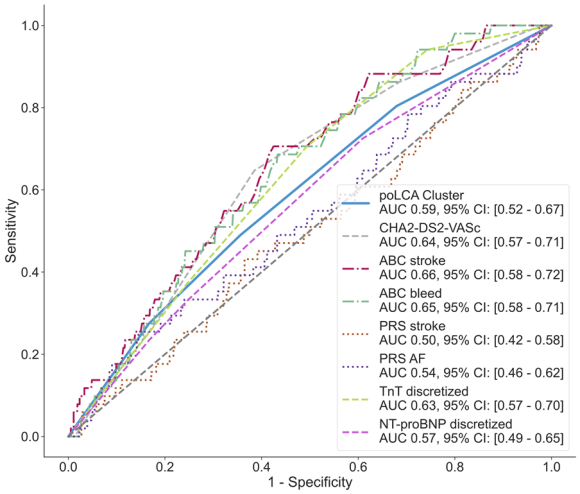


B

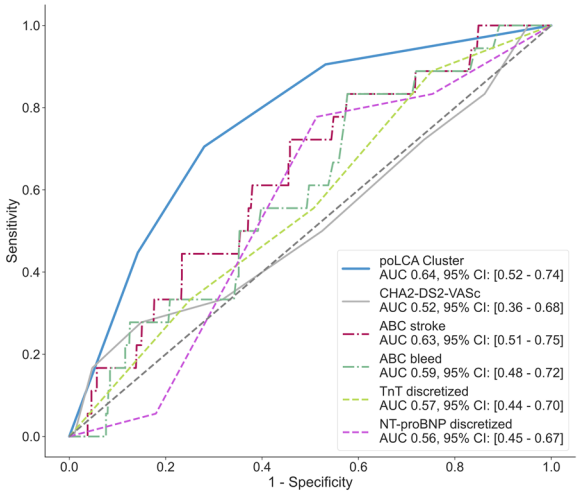


*Supplementary Figure 7: Receiver Operation Characteristics Curves for different Cox PH models **stroke**. All models were fitted against the whole EAST-AFNET 4 (train-) dataset and used to make predictions against the whole EAST-AFNET 4 (train-) dataset. **Plot A** shows model performances / discriminatory powers on EAST-AFNET 4 (train-) dataset as Area under the ROC curve (AUC) at two years. **Plot B** shows model performances / discriminatory powers on BBC-AF (test-) dataset as Area under the ROC curve (AUC) at two years.*

A

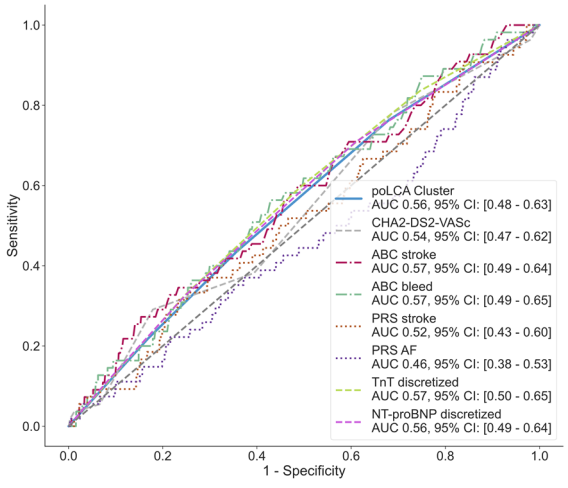


B

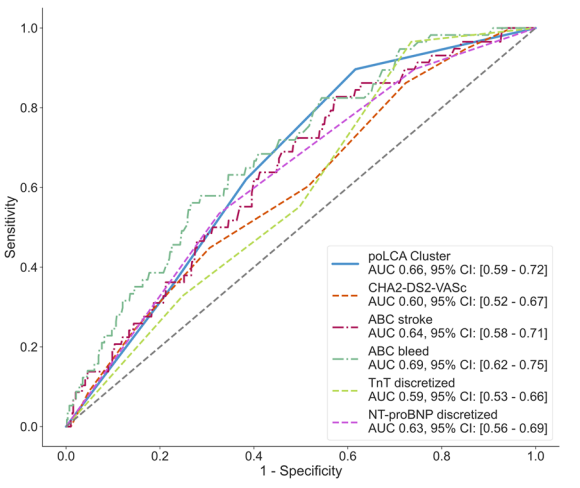


*Supplementary Figure 8: Receiver Operation Characteristics Curves for different Cox PH models **major bleeding event**. All models were fitted against the whole EAST-AFNET 4 (train-) dataset and used to make predictions against the whole EAST-AFNET 4 (train-) dataset. **Plot A** shows model performances / discriminatory powers on EAST-AFNET 4 (train-) dataset as Area under the ROC curve (AUC) at two years. **Plot B** shows model performances / discriminatory powers on BBC-AF (test-) dataset as Area under the ROC curve (AUC) at two years.*

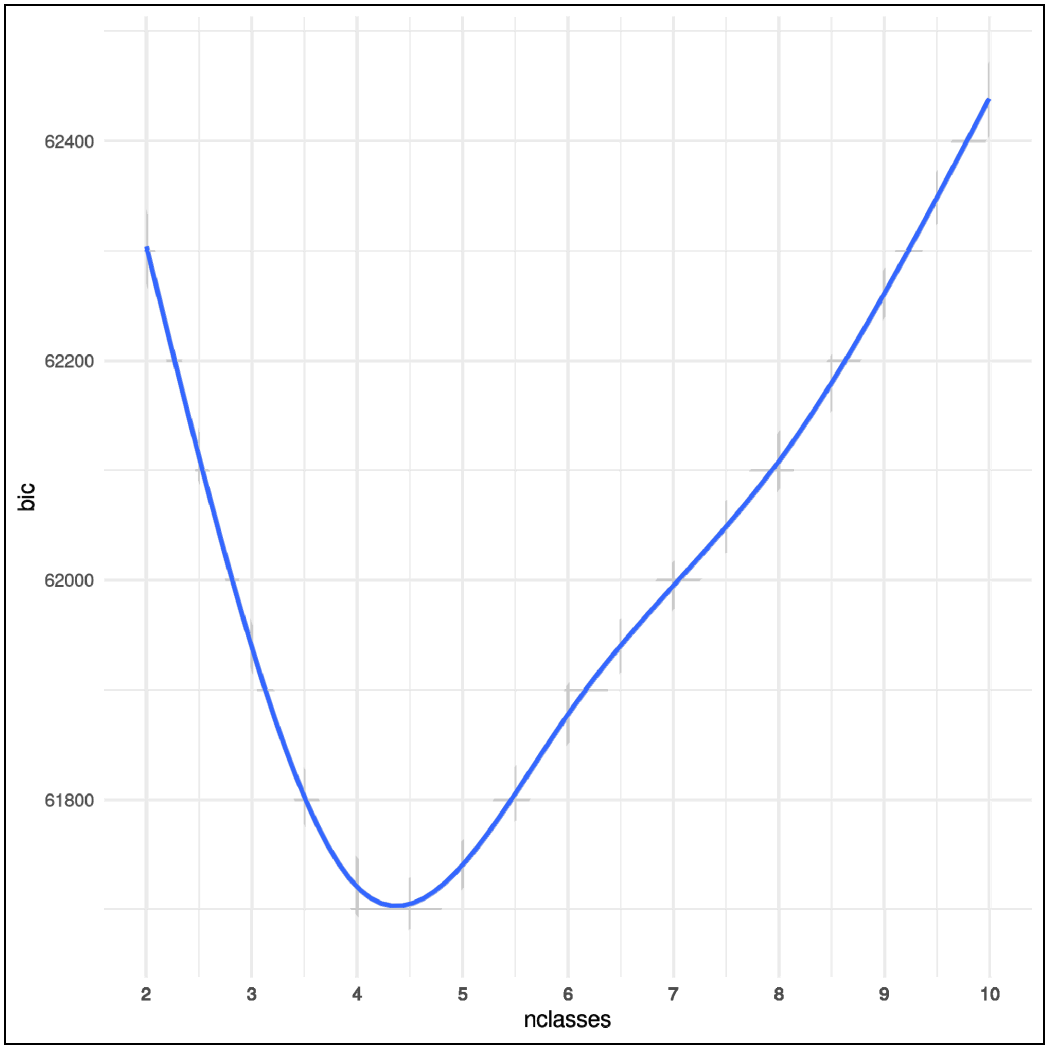
A



B



Supplementary Figure 9: Optimal number of cluster groups found by BIC for derivation cohort (EAST-AFNET4) and poLCA model.



Reduced biomolecule sets used to find new optimal poLCA clusters

We reduced the full set of biomolecule concentrations (n=13) to less biomolecules (8-12). For each reduced set [n=12 – n=8] we examined all possible biomarker combinations, calculated the optimal number of clusters and compared those found clusters with the original poLCA clusters for n=13 biomarkers by the adjusted rand index. We used the newly found clusters to predict the first primary outcome and report the c-index censored as well as the maximal hazard ratio of the subsequent Cox PH models. As we tested 2379 possible biomarker sets we are only reporting models with c-index above a certain threshold or adjusted rand index below a certain threshold. The idea is that possibly models could exist that do a very different partitioning of patients compared to the original clustering model and may yet yield better discriminatory power for assigning participants in low to high risk sub-groups.

Supplementary Table 10a: metrics for biomarker combinations with 12 markers. Showing all possible combinations.

Number biomarkers	Number of optimal clusters	Adjusted Rand Index	Biomarkers missing	c-index censored	max. hazard ratio
12	4	0.909	CA125	0.732	5.232
12	4	0.855	ESM1	0.730	4.875
12	4	0.827	FGF23	0.740	5.147
12	4	0.811	FABP3	0.732	5.274
12	4	0.799	TnT	0.733	5.539
12	4	0.729	NT-proBNP	0.736	5.208
12	4	0.711	ANGPT2	0.731	5.236
12	4	0.569	CRP	0.726	5.264
12	4	0.566	D-dimer	0.727	5.217
12	5	0.551	GDF 15	0.730	4.910
12	4	0.546	IL-6	0.723	5.186
12	4	0.531	IGFBP7	0.736	5.382
12	4	0.490	BMP10	0.731	5.291

Supplementary Table 10b: metrics for biomarker combinations with 11 markers. Showing 13 out of 78 combinations with c-index >0.75 or <0.45 adjusted Rand Index (descending order).

Number biomarkers	Number of optimal clusters	Adjusted Rand Index	Biomarkers missing	c-index censored	max. hazard ratio
11	4	0.803	CA125, FABP3	0.732	5.027
11	4	0.789	TnT, CA125	0.728	5.200
11	4	0.786	D-dimer, FABP3	0.730	5.105
11	4	0.772	D-dimer, FGF23	0.729	4.844
11	4	0.762	TnT, D-dimer	0.726	4.781
11	4	0.757	ESM1, FABP3	0.733	5.429
11	4	0.754	FABP3, ANGPT2	0.733	5.398
11	4	0.445	GDF 15, IGFBP7	0.721	4.807
11	4	0.439	BMP10, FGF23	0.729	5.187
11	4	0.434	FGF23, IGFBP7	0.721	5.120
11	3	0.415	NT-proBNP, IGFBP7	0.732	5.221
11	4	0.407	GDF 15, BMP10	0.723	5.284
11	4	0.393	BMP10, IGFBP7	0.732	5.544

Supplementary Table 10c: metrics for biomarker combinations with 10 markers. Showing 28 out of 286 combinations with c-index >0.70 or <0.38 adjusted Rand Index (descending order).

Number biomarkers	Number of optimal clusters	Adjusted Rand Index	Biomarkers missing	c-index censored	max. hazard ratio
10	4	0.754	D-dimer, CA125, FABP3	0.734	5.053
10	4	0.742	CA125, ESM1, FABP3	0.737	5.823
10	4	0.738	TnT, D-dimer, CA125	0.727	4.952
10	4	0.736	D-dimer, ESM1, FGF23	0.732	4.887
10	4	0.736	D-dimer, ESM1, FABP3	0.732	5.297
10	4	0.735	D-dimer, FABP3, ANGPT2	0.739	5.401
10	4	0.730	CA125, FABP3, ANGPT2	0.732	5.047
10	4	0.727	TnT, CA125, ESM1	0.731	5.199
10	4	0.721	TnT, D-dimer, ESM1	0.725	4.969
10	4	0.711	TnT, CA125, ANGPT2	0.731	4.868
10	4	0.711	D-dimer, FABP3, FGF23	0.733	4.983
10	4	0.703	NT-proBNP, D-dimer, CA125	0.729	4.978
10	4	0.701	TnT, D-dimer, ANGPT2	0.727	4.811
10	4	0.380	CA125, BMP10, IGFBP7	0.729	5.447
10	4	0.380	NT-proBNP, CA125, IGFBP7	0.736	5.789
10	4	0.377	IL-6, BMP10, IGFBP7	0.730	5.660
10	3	0.375	NT-proBNP, IGFBP7, ANGPT2	0.732	4.466
10	4	0.371	BMP10, ESM1, IGFBP7	0.740	6.101
10	3	0.370	BMP10, IGFBP7, ANGPT2	0.732	5.292

Number biomarkers	Number of optimal clusters	Adjusted Rand Index	Biomarkers missing	c-index censored	max. hazard ratio
10	5	0.367	GDF 15, ESM1, IGFBP7	0.722	5.692
10	3	0.365	NT-proBNP, BMP10, IGFBP7	0.734	5.055
10	5	0.363	CA125, ESM1, IGFBP7	0.734	6.841
10	4	0.362	GDF 15, BMP10, IGFBP7	0.731	5.716
10	3	0.362	IL-6, NT-proBNP, IGFBP7	0.728	5.084
10	4	0.361	FGF23, IGFBP7, ANGPT2	0.731	5.088
10	4	0.358	NT-proBNP, FABP3, IGFBP7	0.719	4.761
10	4	0.352	NT-proBNP, ESM1, IGFBP7	0.733	5.656
10	4	0.306	GDF 15, IGFBP7, ANGPT2	0.729	5.488

Supplementary Table 10d: metrics for biomarker combinations with 9 markers. Showing 21 out of 715 combinations with c-index >0.67 or <0.30 adjusted Rand Index (descending order).

Number biomarkers	Number of optimal clusters	Adjusted Rand Index	Biomarkers missing	c-index censored	max. hazard ratio
9	4	0.716	D-dimer, CA125, FABP3, ANGPT2	0.734	5.289
9	4	0.713	D-dimer, CA125, ESM1, FABP3	0.733	5.453
9	4	0.699	D-dimer, CA125, FABP3, FGF23	0.733	4.762
9	4	0.691	CA125, ESM1, FABP3, FGF23	0.738	5.530
9	4	0.686	NT-proBNP, D-dimer, CA125, FABP3	0.734	5.207
9	4	0.685	D-dimer, ESM1, FABP3, FGF23	0.738	5.437
9	4	0.684	D-dimer, ESM1, FABP3, ANGPT2	0.734	5.251
9	4	0.682	TnT, D-dimer, CA125, ANGPT2	0.729	5.198
9	4	0.675	CA125, ESM1, FABP3, ANGPT2	0.734	5.451
9	4	0.296	NT-proBNP, D-dimer, BMP10, IGFBP7	0.729	4.856
9	3	0.294	IL-6, NT-proBNP, IGFBP7, ANGPT2	0.730	5.515
9	4	0.294	NT-proBNP, BMP10, ESM1, IGFBP7	0.733	5.463
9	3	0.292	IL-6, NT-proBNP, GDF 15, IGFBP7	0.718	4.518
9	4	0.291	NT-proBNP, FABP3, IGFBP7, ANGPT2	0.718	4.789
9	4	0.285	BMP10, FABP3, IGFBP7, ANGPT2	0.735	5.903
9	5	0.283	NT-proBNP, GDF 15, CA125, IGFBP7	0.728	6.103
9	4	0.283	NT-proBNP, BMP10, FGF23, IGFBP7	0.730	5.271

Number biomarkers	Number of optimal clusters	Adjusted Rand Index	Biomarkers missing	c-index censored	max. hazard ratio
9	4	0.283	NT-proBNP, GDF 15, IGFBP7, ANGPT2	0.731	5.302
9	4	0.277	NT-proBNP, TnT, IGFBP7, ANGPT2	0.725	3.758
9	3	0.241	GDF 15, BMP10, IGFBP7, ANGPT2	0.734	3.510
9	4	0.237	NT-proBNP, GDF 15, BMP10, IGFBP7	0.733	5.659

Supplementary Table 10e: metrics for biomarker combinations with 8 markers. Showing 28 out of 1287 combinations with c-index >0.60 or <0.25 adjusted Rand Index (descending order).

Number biomarkers	Number of optimal clusters	Adjusted Rand Index	Biomarkers missing	c-index censored	max. hazard ratio
8	4	0.669	D-dimer, CA125, ESM1, FABP3, ANGPT2	0.731	5.371
8	4	0.664	D-dimer, CA125, FABP3, FGF23, ANGPT2	0.731	4.835
8	4	0.663	D-dimer, CA125, ESM1, FABP3, FGF23	0.739	5.627
8	4	0.640	TnT, D-dimer, CA125, ESM1, ANGPT2	0.726	4.735
8	4	0.636	GDF 15, D-dimer, CA125, FABP3, ANGPT2	0.731	5.035
8	4	0.630	NT-proBNP, D-dimer, CA125, FABP3, FGF23	0.735	4.644
8	4	0.630	D-dimer, ESM1, FABP3, FGF23, ANGPT2	0.731	4.739
8	4	0.623	TnT, D-dimer, CA125, ESM1, FABP3	0.727	5.017
8	4	0.622	NT-proBNP, D-dimer, CA125, ESM1, FABP3	0.726	5.238
8	4	0.610	GDF 15, D-dimer, ESM1, FABP3, ANGPT2	0.723	4.745
8	4	0.608	TnT, D-dimer, CA125, FGF23, ANGPT2	0.727	4.742
8	4	0.607	GDF 15, CA125, ESM1, FABP3, ANGPT2	0.724	4.696
8	4	0.603	TnT, D-dimer, ESM1, FGF23, ANGPT2	0.721	4.489
8	4	0.250	NT-proBNP, GDF 15, CA125, BMP10, IGFBP7	0.734	5.625
8	4	0.247	NT-proBNP, GDF 15, D-dimer, BMP10, IGFBP7	0.731	5.312
8	4	0.244	NT-proBNP, GDF 15, BMP10, IGFBP7, ANGPT2	0.728	4.851
8	4	0.244	NT-proBNP, BMP10, FABP3, FGF23, IGFBP7	0.722	4.851

Number biomarkers	Number of optimal clusters	Adjusted Rand Index	Biomarkers missing	c-index censored	max. hazard ratio
8	4	0.244	NT-proBNP, CA125, FABP3, IGFBP7, ANGPT2	0.728	5.265
8	3	0.243	NT-proBNP, TnT, D-dimer, BMP10, IGFBP7	0.731	6.210
8	4	0.243	NT-proBNP, GDF 15, FGF23, IGFBP7, ANGPT2	0.721	3.742
8	4	0.239	NT-proBNP, BMP10, ESM1, FABP3, IGFBP7	0.731	4.980
8	4	0.239	NT-proBNP, GDF 15, BMP10, ESM1, IGFBP7	0.734	5.413
8	4	0.234	NT-proBNP, BMP10, FABP3, IGFBP7, ANGPT2	0.725	4.942
8	4	0.230	NT-proBNP, GDF 15, BMP10, FGF23, IGFBP7	0.727	5.197
8	2	0.220	IL-6, NT-proBNP, GDF 15, FABP3, IGFBP7	0.720	2.844
8	3	0.189	NT-proBNP, TnT, BMP10, FABP3, IGFBP7	0.719	5.061
8	3	0.171	NT-proBNP, TnT, GDF 15, BMP10, IGFBP7	0.715	4.682
8	3	0.170	NT-proBNP, GDF 15, BMP10, FABP3, IGFBP7	0.718	4.911