Supplemental Material

Sharma A, Zheng Y, Ezekowitz JA, et al. Cluster analysis of cardiovascular phenotypes in patients with type 2 diabetes and established atherosclerotic cardiovascular disease: a potential approach to precision medicine.

Supplemental Table 1. Observed Clinical Outcomes by Cluster in TECOS

	Cluster I		Cluster II		Cluster III		Cluster IV	
N	6001 (40.9%)		3490 (23.7%)		2672 (18.2%)		2508 (17.1%)	
	n (%)	Events/100 pt-years	n (%)	Events/100 pt-years	n (%)	Events/100 pt-years	n (%)	Events/100 pt-years
Primary composite CV outcome: CV death, nonfatal MI, nonfatal stroke, or hospitalization for unstable angina	695 (11.6)	4.13	299 (8.6)	2.89	274 (10.3)	3.75	422 (16.8)	6.38
Secondary composite CV outcome: CV death, nonfatal MI, or nonfatal stroke	594 (9.9)	3.49	265 (7.6)	2.55	253 (9.5)	3.44	379 (15.1)	5.67
CV death	215 (3.6)	1.19	161 (4.6)	1.48	129 (4.8)	1.66	241 (9.6)	3.34
Fatal or non-fatal MI	345 (5.7)	2	78 (2.2)	0.74	69 (2.6)	0.92	124 (4.9)	1.82
Fatal or non-fatal stroke	128 (2.1)	0.73	51 (1.5)	0.48	90 (3.4)	1.21	92 (3.7)	1.34
Hospitalization for unstable angina	124 (2.1)	0.71	42 (1.2)	0.4	25 (0.9)	0.33	54 (2.2)	0.78
All-cause mortality	347 (5.8)	1.91	214 (6.1)	1.96	191 (7.1)	2.46	332 (13.2)	4.6
Hospitalization for HF	173 (2.9)	0.99	49 (1.4)	0.46	55 (2.1)	0.73	180 (7.2)	2.67
Composite of cardiovascular death, non-fatal MI, non-fatal stroke, hospitalization for unstable angina, or hospitalization for HF	787 (13.1)	4.72	324 (9.3)	3.15	308 (11.5)	4.24	516 (20.6)	8.05
Other CV death (besides fatal MI or fatal stroke) or hospitalization for HF	337 (5.6)	1.88	192 (5.5)	1.77	169 (6.3)	2.19	345 (13.8)	6.38

CV: Cardiovascular; HF: heart failure; MI: myocardial infarction

Supplemental Table 2. Baseline Characteristics by Cluster in EXSCEL

	Cluster I	Cluster II	Cluster III	Cluster IV
N	7727	4413	599	2013
Age (years)[a]	62.5 (9.3)	60.2 (9.9)	60.5 (8.4)	63.6 (8.5)
Women	2643 (34.2)	1979 (44.8)	258 (43.1)	723 (35.9)
Race				
White	7655 (99.1)	1228 (27.8)	509 (85.0)	1783 (88.6)
Black	9 (0.1)	773 (17.5)	25 (4.2)	71 (3.5)
Asian	12 (0.2)	1318 (29.9)	52 (8.7)	70 (3.5)
Other	46 (0.6)	1094(24.8)	13 (2.1)	89 (4.4)
Region				
Asia Pacific and Other	303 (3.9)	1103 (25.0)	55 (9.2)	68 (3.4)
Eastern Europe	2562 (33.2)	26 (0.6)	273 (45.6)	1154 (57.3)
Latin America	173 (2.2)	2260 (51.2)	80 (13.4)	214 (10.6)
North America	2509 (32.5)	801 (18.2)	55 (9.2)	343 (17.0)
Western Europe	2180 (28.2)	223 (5.1)	136 (22.7)	234 (11.6)
Ethnicity				
Hispanic/Latino	57 (0.7)	2643 (59.9)	82 (13.7)	244 (12.1)
Duration of diabetes (years)[b]	13.3 (8.2)	12.9 (8.3)	11.8 (7.2)	13.2 (8.6)
HbA$_{1c}$ (%)	8.0 (0.9)	8.2 (1.0)	8.3 (0.9)	8.2 (1.0)
HbA$_{1c}$ (mmol/mol)	64.4 (10.2)	66.1 (10.8)	67.1 (10.1)	65.8 (10.6)
Body mass index (kg/m^2)	33.8 (6.2)	30.0 (5.9)	34.3 (6.3)	33.6 (6.4)
Systolic blood pressure (mmHg)	135.8 (16.2)	135.0 (18.4)	136.7 (15.7)	135.0 (16.1)
Diastolic blood pressure (mmHg)	77.8 (10.3)	78.4 (10.2)	79.4 (9.8)	78.4 (10.5)
eGFR (mL/min/1.73 m^2)[c]	78.8 (23.4)	79.6 (25.4)	82.3 (24.3)	73.1 (23.4)
Urinary albumin:creatinine ratio (g/mol)	1.4 (0.5, 4.7)	2.2 (0.5, 7.9)	1.7 (0.4, 6.0)	1.8 (0.0, 7.0)
LDL-cholesterol (mmol/L)	2.4 (1.7)	2.5 (1.1)	2.7 (1.0)	2.5 (1.1)
HDL-cholesterol (mmol/L)	1.1 (0.8)	1.1 (0.4)	1.1 (0.3)	1.1 (0.3)
Triglycerides (mmol/L)	2.3 (5.1)	2.1 (1.4)	2.3 (1.3)	2.2 (1.4)
Triglycerides (mg/dL)	204(452)	186(124)	204(115)	195(124)
Prior Myocardial infarction	2302 (29.8)	1147 (26.0)	142 (23.7)	1088 (54.0)
Prior $\geq$50% coronary stenosis	2635 (34.1)	1173 (26.6)	130 (21.8)	854 (42.4)
Prior PCI	2433 (31.5)	1130 (25.6)	115 (19.2)	715 (35.5)
Prior CABG	1344 (17.4)	509 (11.5)	66 (11.0)	447 (22.2)
Prior cerebrovascular disease	1493 (19.3)	640 (14.5)	129 (21.6)	452 (22.5)
Prior peripheral arterial disease	672 (8.7)	181 (4.1)	98 (16.4)	258 (12.8)
Prior heart failure	158 (2.0)	32 (0.7)	186 (31.1)	2013 (100.0)

NYHA class 3 or higher	101 (63.9)	31 (96.9)	24 (12.9)	160 (7.9)
Current smoker	1010 (13.1)	444 (10.1)	72 (12.0)	195 (9.7)
Diabetic neuropathy	2409 (31.2)	1165 (26.4)	301 (50.3)	839 (41.7)
Retinopathy	1275 (16.5)	554 (12.6)	144 (24.1)	543 (27.0)

Abbreviations: CABG, coronary artery bypass graft; eGFR, estimated glomerular filtration rate; HbA$_{1c}$, glycated hemoglobin; HDL, high-density lipoprotein; LDL, low-density lipoprotein; PCI, percutaneous coronary intervention.

**Supplemental Table 3. Baseline Glucose-Lowering and Cardiac-Related Medication Use by Cluster in EXSCEL**

	Cluster I	Cluster II	Cluster III	Cluster IV
N	7727	4413	599	2013
Metformin	5903 (76.4)	3505 (79.4)	468 (78.1)	1418 (70.4)
Sulfonylurea	2626 (34.0)	1845 (41.8)	201 (33.6)	729 (36.2)
Biguanides	5904 (76.4)	3505 (79.4)	468 (78.1)	1418 (70.4)
Thiazolidinedione	369 (4.8)	156 (3.5)	18 (3.0)	36 (1.8)
Insulin	3577 (46.3)	1960 (44.4)	307 (51.3)	992 (49.3)
Aspirin	4910 (63.5)	2597 (58.8)	336 (56.1)	1537 (76.4)
Beta blocker	4495 (58.2)	1857 (42.1)	325 (54.3)	1534 (76.2)
ACE inhibitor or ARB	6143 (79.5)	3073 (69.6)	477 (79.6)	1705 (84.7)
Calcium channel blocker	2523 (32.7)	1290 (29.2)	232 (38.7)	665 (33.0)
Diuretic	3532 (45.7)	1351 (30.6)	323 (53.9)	1237 (61.5)
Statin	5952 (77.0)	2892 (65.5)	394 (65.8)	1607 (79.8)
Other lipid-lowering medication	6234 (80.7)	3080 (69.8)	415 (69.3)	1641 (81.5)

Abbreviations: ACE, angiotensin-converting enzyme; ARB, angiotensin-receptor blocker.

Supplemental Table 4. Observed Clinical Outcomes by Cluster in the EXSCEL Trial

	Cluster I		Cluster II		Cluster III		Cluster IV	
N	7727		4413		599		2013	
	n (%)	Events/100 pt-years	n (%)	Events/100 pt-years	n (%)	Events/100 pt-years	n (%)	Events/100 pt-years
MACE (CVD/MI/stroke)	985 (12.7)	3.93	331 (7.5)	2.63	64 (10.7)	3.45	364 (18.1)	6.10
Fatal or non-fatal MI	606 (7.8)	2.39	161 (3.6)	1.27	28 (4.7)	1.49	181 (9.0)	3.00
CV death	347 (4.5)	1.25	149 (3.4)	1.10	31 (5.2)	1.56	196 (9.7)	3.01
Hospitalization for acute coronary syndrome	700 (9.1)	2.79	223 (5.1)	1.77	37 (6.2)	1.98	212 (10.5)	3.55
All-cause death	544 (7.0)	1.96	225 (5.1)	1.66	54 (9.0)	2.73	268 (13.3)	4.11
Fatal stroke	32 (0.4)	0.12	6 (0.1)	0.05	2 (0.3)	0.10	13 (0.6)	0.21
Fatal MI	17 (0.2)	0.06	3 (0.1)	0.02	3 (0.5)	0.16	10 (0.5)	0.16
Hospitalization for HF	242 (3.1)	0.93	55 (1.2)	0.43	18 (3.0)	0.95	135 (6.7)	2.20
Hypoglycemia requiring assistance	202 (2.6)	0.72	162 (3.7)	1.18	20 (3.3)	1.00	82 (4.1)	1.25

Supplemental Figure 1. Association Between Cluster and Individual Cardiovascular Clinical Outcomes
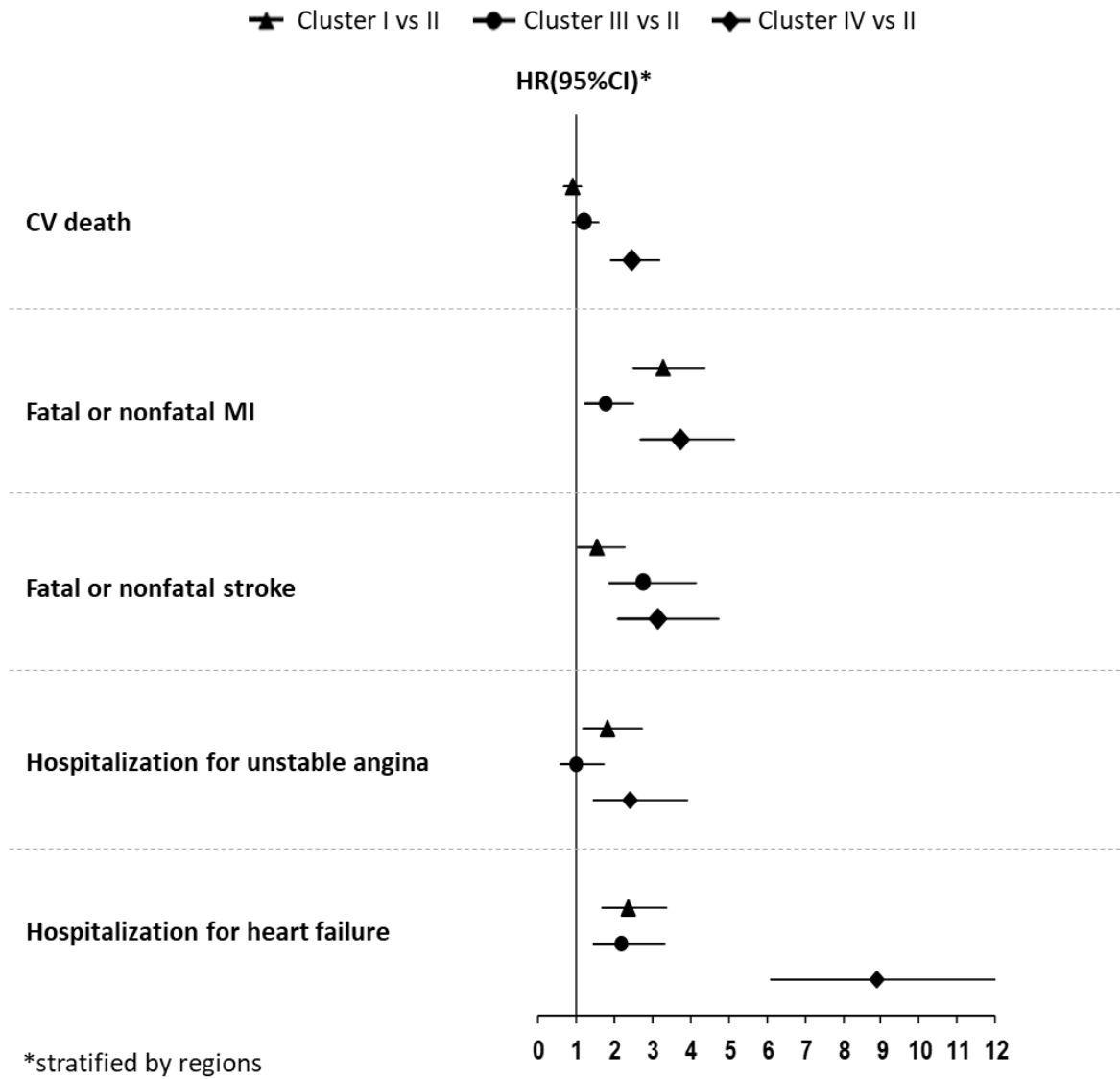
Supplemental Figure 2. Interaction Between Study Treatment and Cluster on Clinical Outcomes
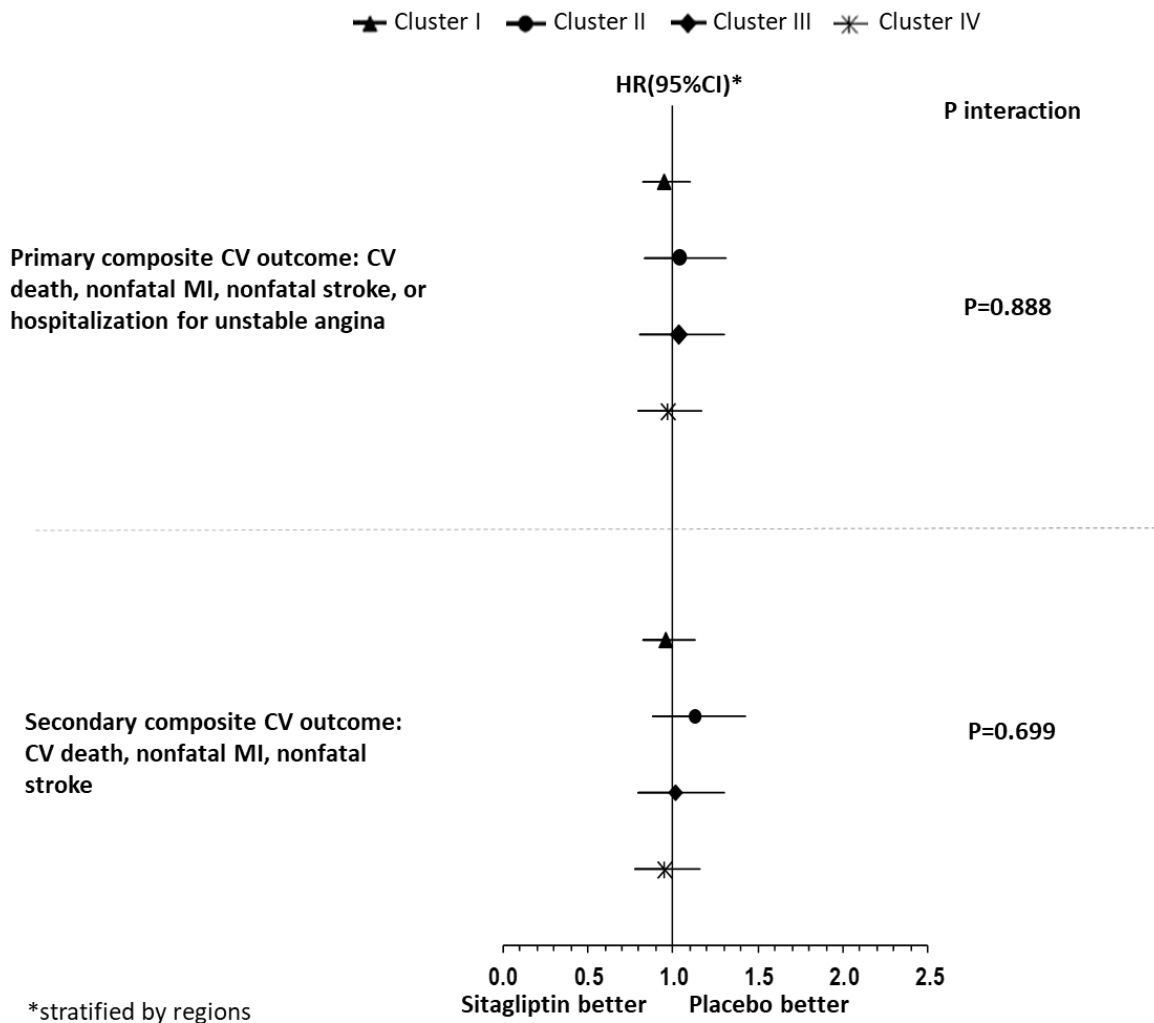
Supplemental Appendix 1. Variable Cluster Analysis in TECOS

Variable clustering was done for continuous and categorical variables separately. There were 14 continuous variables and 26 categorical variables considered at the beginning of analysis (see table below for details). Using SAS PROC VARCLUS, variables are aggregated into several dimensions by maximizing correlation within dimensions and minimizing correlations between dimensions. Second eigenvalue less than 1 of each dimension during each iteration was used to stop splitting dimensions. For continuous variables, 4 dimensions (clusters) were eventually identified. For categorical variables, 10 clusters were determined. Interpretation of variable dimension reduction is not critical in our analysis especially for binary variables because our goal is to use lower dimension of variables to best describe the data, to create a summary score for each variable cluster and for each patient. Then the summary score was used for the next analysis: patient clustering. The same 14 continuous variables and 26 categorical variables were selected for validation analysis in EXSCEL.

| Continuous variables	age, BMI, weight, systolic BP at baseline, diastolic BP at baseline, heart rate, HbA1c, hemoglobin, HDL cholesterol, triglycerides, log-transformed UACR, LDL cholesterol, duration of diabetes, and eGFR
Categorical variables	sex, Hispanic or Latino ethnicity, race, NYHA class, prior coronary artery disease, prior MI, prior percutaneous coronary intervention, prior peripheral arterial disease, prior coronary artery bypass graft prior cerebrovascular disease, heart failure, chronic obstructive pulmonary disease; chronic liver disease, cancer, hypertension, alcohol abuse, complication of amputation, foot ulcer, retinopathy, blindness, diabetic neuropathy, currently smoking, depression.

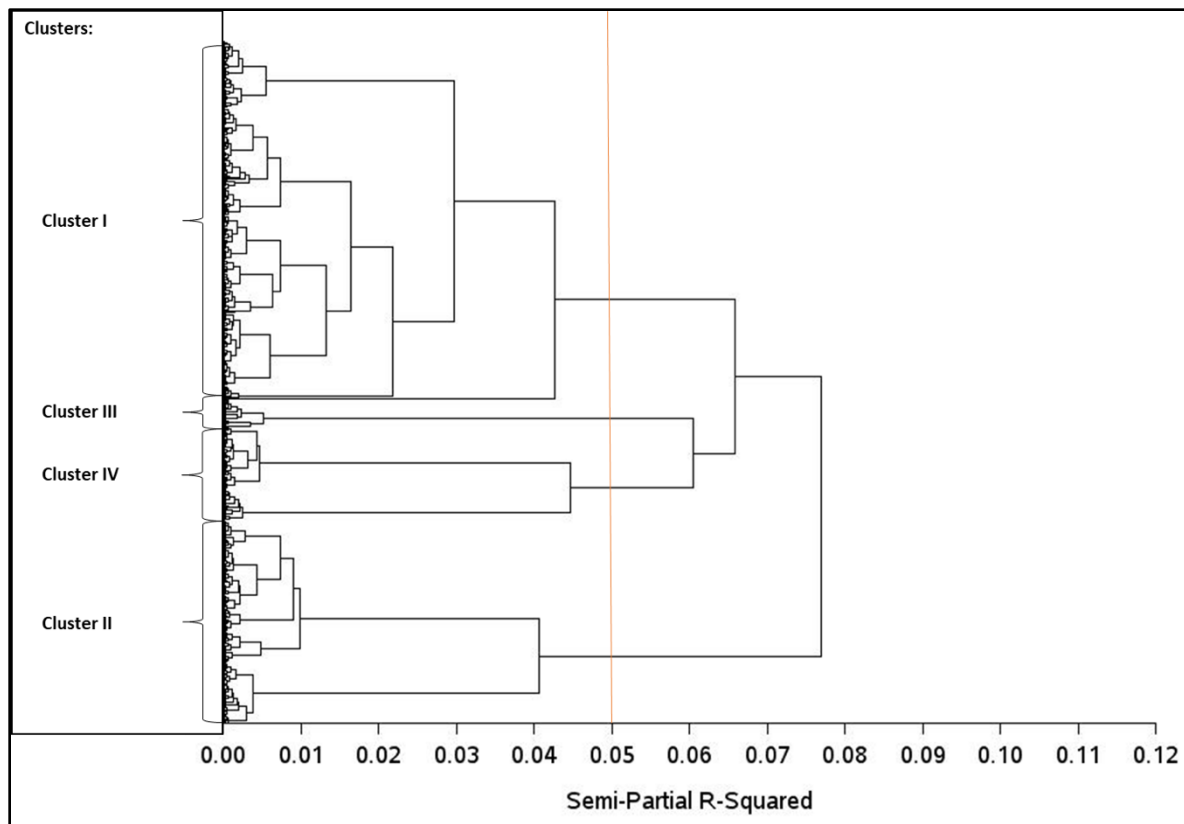**Supplemental Appendix 2. Patient Cluster Analysis in TECOS**

The analysis started with each patient as a cluster and then iteratively merged clusters together until all patients were merged into a single cluster according to the distance between clusters. Ward's minimum variance method was used to define the distance. In brief, the distance between two clusters K and L was defined as

$$D_{K,L} = \frac{\sum_j \left(x_{K,j} - x_{L,j}\right)^2}{\frac{1}{n_K} + \frac{1}{n_L}}$$

where j indexes the cluster's 13 variable scores. $x_{K,j}$ is the value of the jth standardized variable score for cluster K and n_K is the number of original patients in cluster K at that stage. At each iteration, the distance was calculated between every possible combination of two clusters. Two clusters with the smallest distance were merged and then new distance was calculated. Once the patient was classified into one cluster they stayed in that cluster for the whole analysis. For each iteration, cubic-clustering criterion (CCC) were calculated. The CCC metrics were used to determine number of clusters. In the published literature, there are no standard criteria to decide the number of clusters, but a local maximum was used as a guideline. Figures below plotted these metrics against the number of clusters (only up to 40) using SAS PROC CLUSTER. It seems 5 clusters separated the population best. A tree diagram is also provided below according to the semi-partial R2 at each iteration, which roughly measures loss of homogeneity within clusters caused by their merger. Thus, small values of the semi-partial R2 indicate that two similar clusters have been merged whereas large values indicate the merger of two heterogeneous clusters. The 4-cluster model retained a semipartial R2 of 0.05 and formed much clearer patterns of patient clusters than the 5-cluster model. Therefore the 4-cluster model was presented in the manuscript.
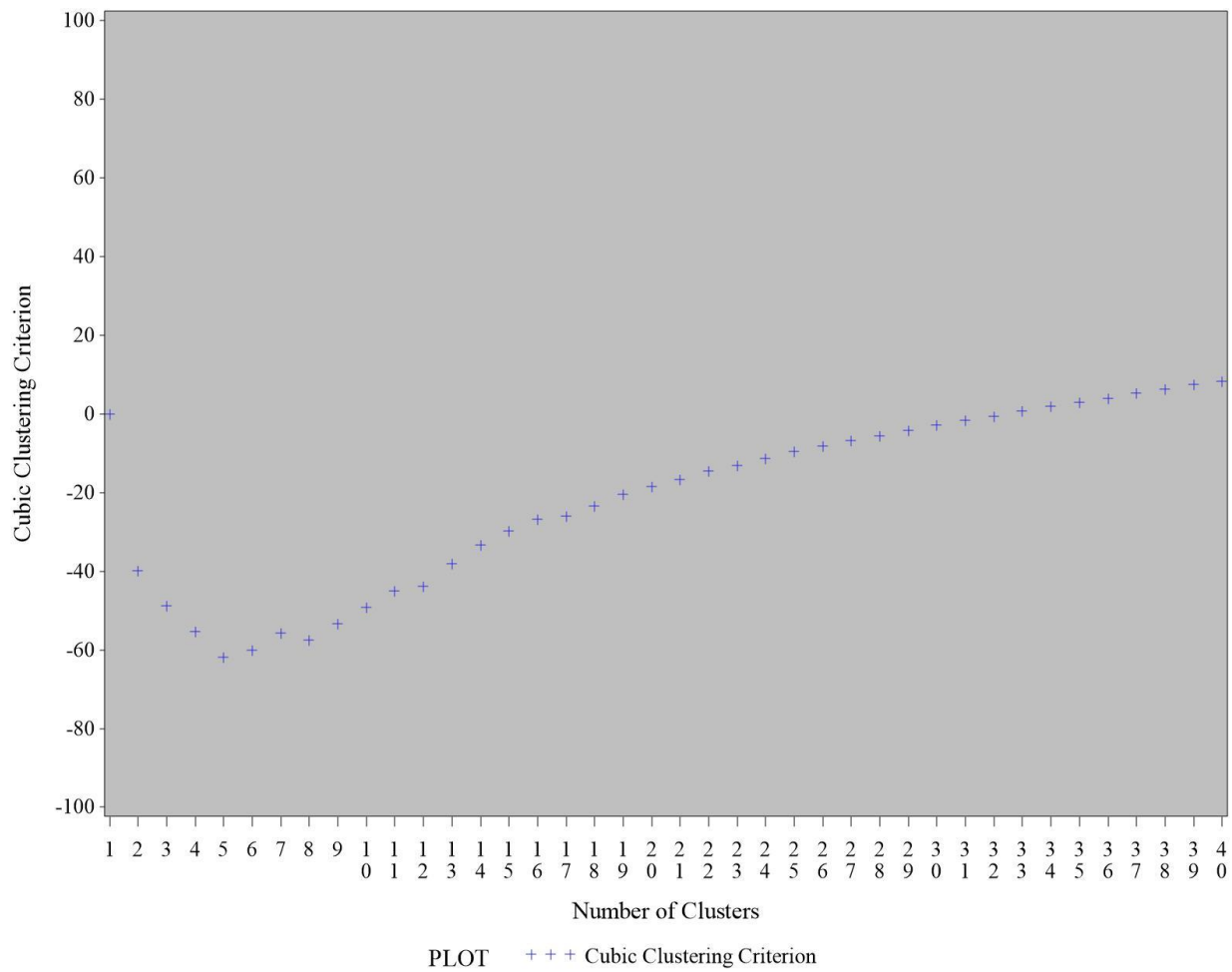
A semipartial R^2 of 0.05, which was the same value used in the TECOS cluster analysis, was used to determine the final number of clusters in the validation analysis in EXSCEL. 4 patient clusters were identified (see tree diagram below).
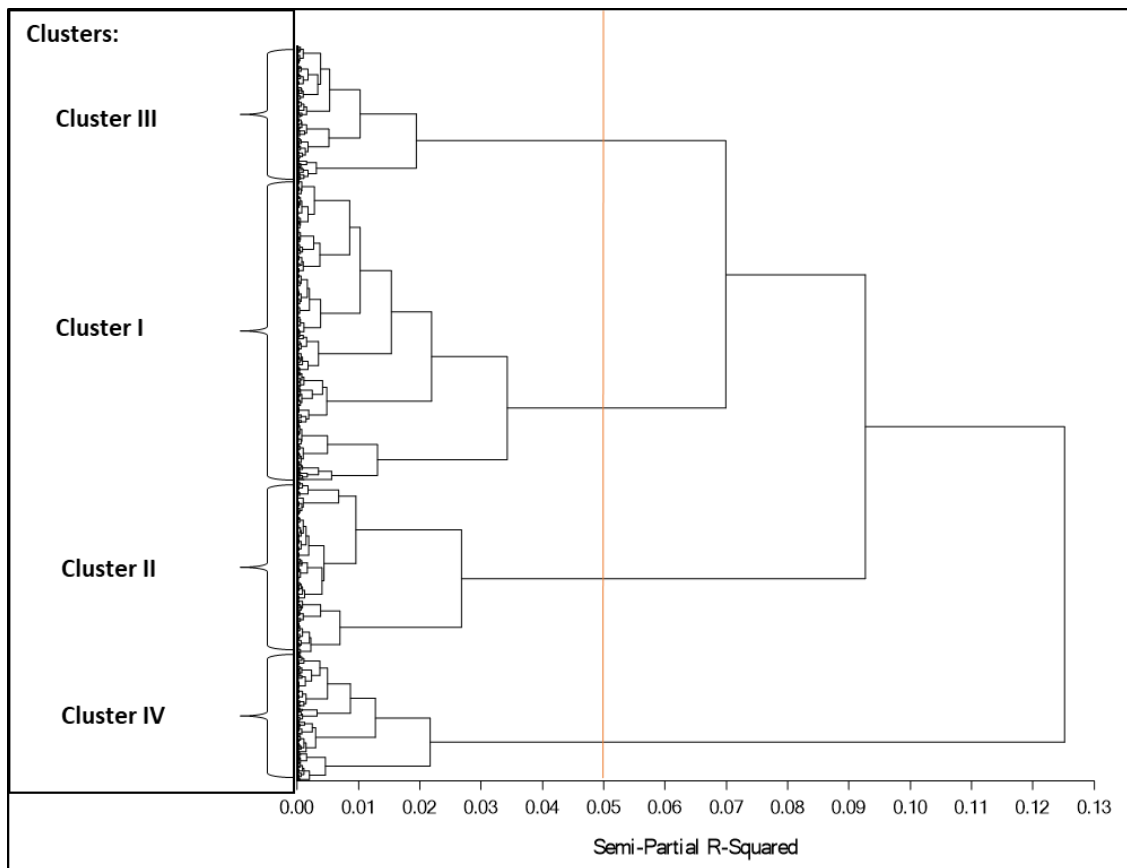
Supplemental Appendix 3. 5-Cluster Model Cluster Analysis in TECOS

Supplemental Appendix 3-Table 1. Baseline Characteristics by 5 Clusters

	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V
n	3285	3440	3791	1616	2539
Age (years)[a]	68.0 (7.4)	65.6 (7.9)	63.6 (7.9)	62.4 (7.0)	66.8 (8.1)
Women	679 (20.7%)	1516 (44.1 %)	883 (23.3%)	313 (19.4%)	906 (35.7%)
Race					
White	3086 (93.9%)	2729 (79.3%)	555 (14.6%)	1444 (89.4%)	2143 (84.4%)
Black	81 (2.5%)	199 (5.8%)	100 (2.6%)	26 (1.6%)	41 (1.6%)
Asian	22 (0.7%)	201 (5.8%)	2870 (75.7%)	46 (2.8%)	126 (5.0%)
Other	96 (2.9%)	311 (9.0%)	266 (7.0%)	100 (6.2%)	229 (9.0%)
Region					
Asia Pacific and Other	648 (19.7%)	501 (14.6%)	2917 (76.9%)	283 (17.5%)	216 (8.5%)
Eastern Europe	643 (19.6%)	1212 (35.2%)	81 (2.1%)	560 (34.7%)	1469 (57.9%)
Latin America	207 (6.3%)	569 (16.5%)	248 (6.5%)	173 (10.7%)	274 (10.8%)
North America	1102 (33.5%)	576 (16.7%)	289 (7.6%)	333 (20.6%)	294 (11.6%)
Western Europe	685 (20.9%)	582 (16.9%)	256 (6.8%)	267 (16.5%)	286 (11.3%)
Ethnicity					
Hispanic/Latino	302 (9.2%)	660 (19.2%)	302 (8.0%)	209 (12.9%)	325 (12.8%)
Duration of diabetes (years)[b]	13.3 (8.6)	12.0 (8.3)	10.9 (7.6)	9.1 (6.3)	11.4 (8.3)
HbA1c (%)	7.2 (0.5)	7.2 (0.5)	7.3 (0.5)	7.2 (0.5)	7.2 (0.5)
HbA1c (mmol/mol)	55.3 (6.6)	55.8 (7.4)	56.5 (6.6)	55.8 (7.0)	55.5 (7.0)
Body mass index (kg/m^2)	31.1 (5.4)	30.9 (5.6)	26.8 (4.3)	32.5 (5.5)	31.7 (5.6)
Systolic blood pressure (mmHg)	133.9 (16.5)	137.7 (17.5)	132.2 (16.7)	137.3 (16.6)	135.7 (17.1)
Diastolic blood pressure (mmHg)	74.0 (10.5)	77.9 (10.5)	76.9 (10.0)	80.2 (9.8)	78.8 (10.4)
eGFR (mL/min/1.73 m^2)§	71.6 (19.4)	75.2 (22.0)	76.5 (20.3)	82.1 (21.6)	71.7 (21.3)
eGFR <50 mL/min/1.73 m^{2c}	358 (11.0)	311 (9.1)	282 (7.5)	68 (4.2)	352 (14.0)
Urinary albumin:creatinine ratio (mg/g)	11.0 (4.4, 31.8)	11.8 (3.5, 40.8)	7.6 (3.3, 28.0)	13.0 (4.4, 37.8)	17.8 (4.7, 55.0)
Total cholesterol (mg/dL)	149.8 (34.1)	176.7 (48.7)	157.7 (39.1)	181.2 (49.9)	173.8 (49.7)
LDL-cholesterol (mg/dL)	80.9 (95.3)	97.8 (40.1)	88.1 (34.1)	96.3 (41.0)	97.4 (40.7)
HDL-cholesterol (mg/dL)	43.5 (12.3)	45.4 (13.9)	42.9 (10.8)	40.3 (12.2)	43.8 (13.0)
Triglycerides (mg/dL)	142.1 (62.1)	166.0 (85.7)	140.3 (66.0)	262.0 (178.2)	169.8 (90.5)
Triglycerides (mmol/L)	1.6 (0.7)	1.9 (1.0)	1.6 (0.7)	2.9 (2.0)	1.9 (1.0)
Prior atherosclerotic coronary disease	3244 (98.8%)	991 (28.8%)	3007 (79.3%)	1505 (93.1%)	2116 (83.3%)
Myocardial infarction	1764 (53.7%)	554 (16.1%)	1454 (38.4%)	934 (57.8%)	1549 (61.0%)
≥50% coronary stenosis	2437 (74.2%)	523 (15.2%)	2380 (62.8%)	1011 (62.6%)	1336 (52.6%)
Prior PCI	1859 (57.6%)	393 (11.5%)	1560 (41.8%)	857 (53.6%)	1045 (41.7%)
CABG	1302 (39.6%)	207 (6.0%)	1023 (27.0%)	440 (27.2%)	692 (27.3%)
Prior cerebrovascular disease	439 (13.4%)	1678 (48.8%)	625 (16.5%)	205 (12.7%)	641 (25.2%)
Prior peripheral arterial disease	238 (7.2%)	1376 (40.0%)	436 (11.5%)	70 (4.3%)	313 (12.3%)
Prior heart failure	5 (0.2%)	79 (2.3%)	13 (0.3%)	8 (0.5%)	2538 (100.0%)
NYHA class 3 or higher	3 (60.0%)	34 (43.0%)	7 (53.8%)	4 (50.0%)	325 (12.8%)
Cigarette smoking					
Current smoker	122 (3.7%)	536 (15.6%)	311 (8.2%)	438 (27.1%)	271 (10.7%)
Prior smoker	2015 (61.3%)	1120 (32.6%)	1215 (32.0%)	491 (30.4%)	1003 (39.5%)
Never smoked	1148 (34.9%)	1784 (51.9%)	2265 (59.7%)	687 (42.5%)	1265 (49.8%)
Diabetic neuropathy	769 (23.4%)	1231 (35.8%)	436 (11.5%)	141 (8.7%)	777 (30.6%)
Retinopathy	490 (14.9%)	597 (17.4%)	282 (7.4%)	55 (3.4%)	440 (17.3%)

Results for continuous variables are mean (SD) or median (Q1, Q3), and categorical variables are n (%). Urinary albumin:creatinine ratio data available for only 5148 patients (n= 2606 for sitagliptin, n=2542 for placebo). SI conversion factors: urine albumin creatinine ratio (mg/g to g/mol), multiply by 0.1131; total cholesterol, LDL-cholesterol and HDL-cholesterol (mg/dL to mmol/L), multiply by 0.0259; triglycerides (mg/dL to mmol/L), multiply by 0.0113.

Abbreviations: CABG, coronary artery bypass graft; eGFR, estimated glomerular filtration rate; HbA$_{1c}$, glycated hemoglobin; HDL, high-density lipoprotein; LDL, low-density lipoprotein; PCI, percutaneous coronary intervention.

[a]Age missing among patients in Lithuania as birth date could not be provided.

[b]Duration = (year of randomization – year of diagnosis) + 1.

[c]MDRD formula used to calculate eGFR. Site-reported values are presented.

Supplemental Appendix 3-Table 2. Antihyperglycemic and Cardiac-Related Medication Use at Baseline by 5 Clusters

|	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V
n	3285	3440	3791	1616	2539
Metformin	2656 (80.9%)	2749 (79.9%)	3318 (87.5%)	1388 (85.9%)	1855 (73.1%)
Sulfonylurea	1247 (38.0%)	1431 (41.6%)	2235 (59.0%)	640 (39.6%)	1092 (43.0%)
Thiazolidinedione	137 (4.2%)	75 (2.2%)	101 (2.7%)	44 (2.7%)	39 (1.5%)
Insulin	966 (29.4%)	957 (27.8%)	440 (11.6%)	327 (20.2%)	718 (28.3%)
More than two agents above	1956 (59.5%)	1927 (56.0%)	2669 (70.4%)	900 (55.7%)	1368 (53.9%)
Beta blocker	2439 (74.2%)	1552 (45.1%)	2135 (56.3%)	1205 (74.6%)	1991 (78.4%)
ACE inhibitor or ARB	2812 (85.6%)	2645 (76.9%)	2572 (67.8%)	1365 (84.5%)	2161 (85.1%)
Calcium channel blocker	1243 (37.8%)	1242 (36.1%)	1055 (27.8%)	539 (33.4%)	882 (34.7%)
Diuretic	1461 (44.5%)	1554 (45.2%)	853 (22.5%)	653 (40.4%)	1499 (59.0%)
Thiazide	885 (60.6%)	989 (63.6%)	540 (63.3%)	416 (63.7%)	634 (42.3%)
Aspirin	2767 (84.2%)	2285 (66.4%)	3185 (84.0%)	1367 (84.6%)	1914 (75.4%)
Other antiplatelet	729 (22.2%)	478 (13.9%)	1221 (32.2%)	307 (19.0%)	452 (17.8%)
Statin	2926 (89.1%)	2332 (67.8%)	3183 (84.0%)	1327 (82.1%)	1951 (76.8%)
Ezetimibe	262 (8.0%)	160 (4.7%)	146 (3.9%)	115 (7.1%)	78 (3.1%)
Nitrates	786 (23.9%)	358 (10.4%)	766 (20.2%)	302 (18.7%)	601 (23.7%)

Results are n (%).

Abbreviations: ACE , angiotensin-converting enzyme; ARB, angiotensin-receptor blocker.
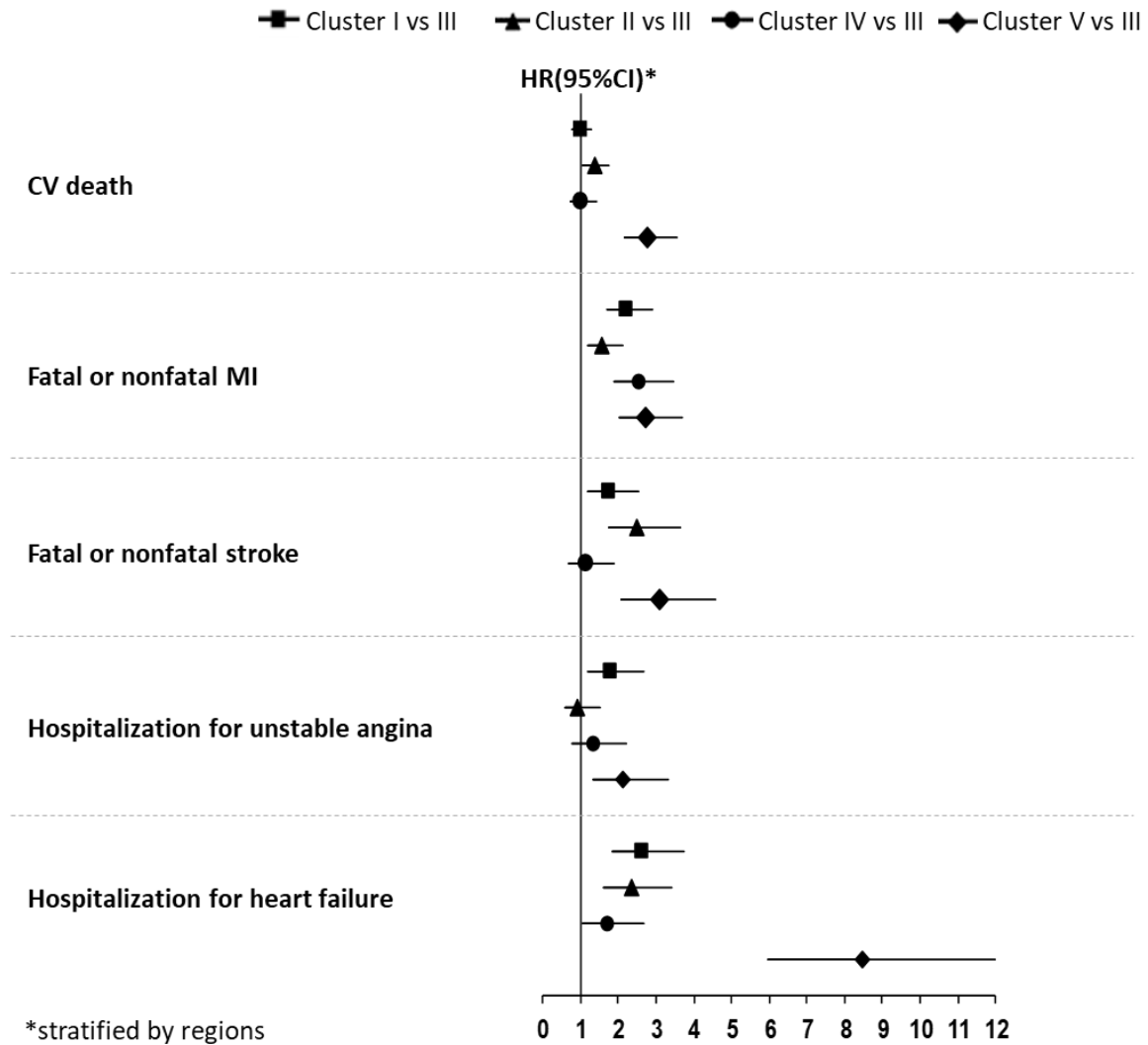
Supplemental Appendix 3-Figure 1. Kaplan-Meier estimated cumulative incidence of cardiovascular death, nonfatal myocardial infarction, nonfatal stroke, or hospitalization for unstable angina endpoint by 5 clusters
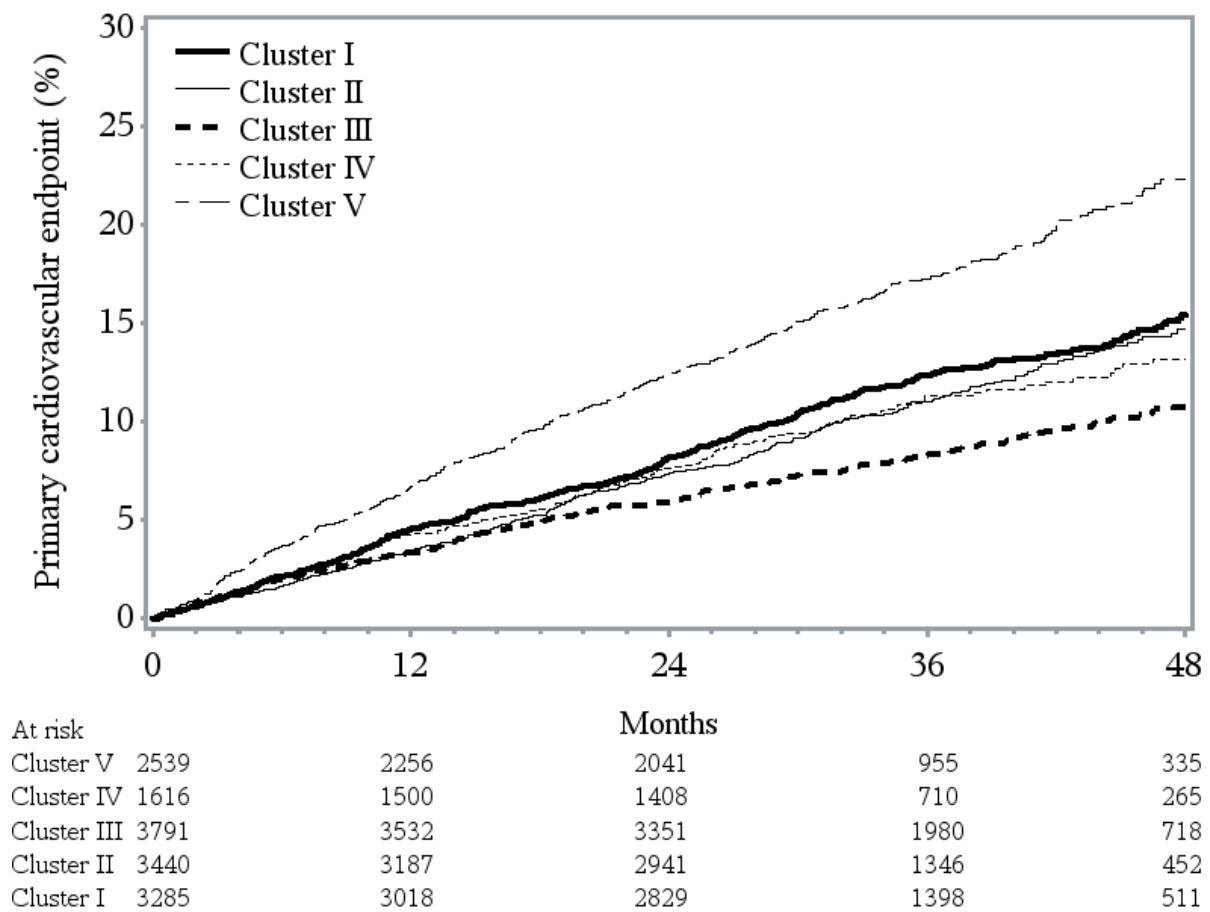
Supplemental Appendix 3-Figure 2. Association between 5 clusters and clinical outcomes (cluster III as reference category)
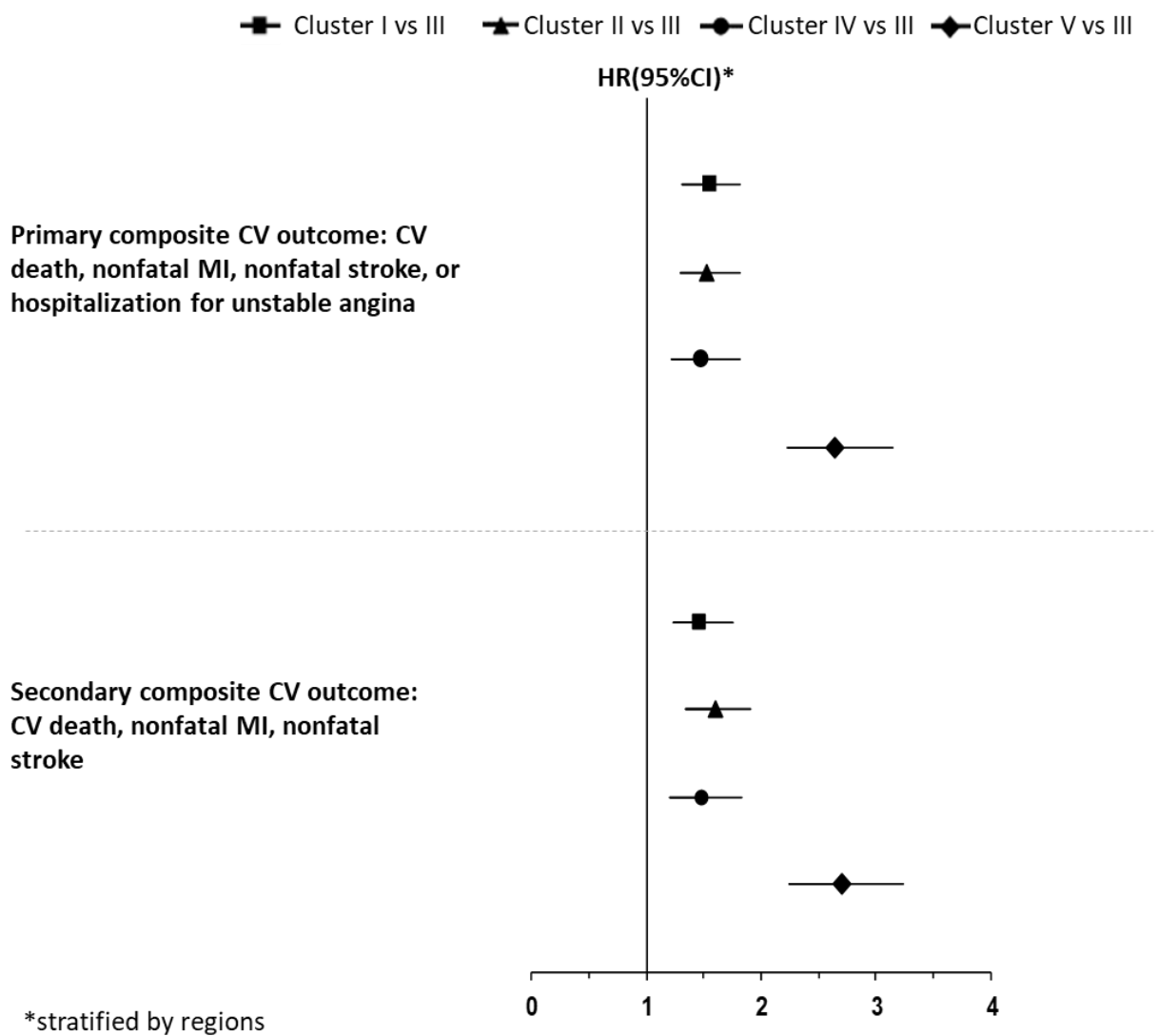
Supplemental Appendix 3-Figure 3. Association between 5 clusters and individual cardiovascular clinical outcomes

Supplemental Appendix 3-Figure 4. Interaction Between Study Treatment and 5 Clusters on Clinical Outcomes