SUPPLEMENTAL MATERIAL

Supplemental Table 1. Subject characteristics from Electronic Health Record data used to construct risk models in this study

	EHR (n = 42,058)	Framingham Original Cohort* (n = 8,491)	Pooled Cohort+ (n = 24,626)
Sex			
Male	17,481 (41%)	3,969 (47%)	10,745 (44%)
Female	24,577 (58%)	4,522 (53%)	13881 (56%)
Race			
White	30,667 (73%)	N/A	20,338 (83%)
Black	2,908 (6.9%)	N/A	4,288 (17%)
Other or Not Reported	8,483 (20.2%)	N/A	0 (0%)
Age (years)	52 (9)	49 (11)	56 (6)
SBP (mmHg)	123 (13)	128 (19)	126 (18)
HDL (mg/dL)	48 (12)	52 (14)	52 (15)
Total Cholesterol (mg/dL)	194 (28)	214 (42)	215 (41)
Smoking			
Never/Former	35,557 (84%)	5,545 (65%)	18,316 (74%)
Current	6,501 (15%)	2,946 (35%)	6,310 (26%)
Diabetes			
No	38,157 (91%)	8,063 (95%)	22,390 (91%)
Yes	3,901 (9%)	428 (5%)	2,236 (9%)
Taking BP-lowering medications			
No	29,284 (70%)	7,557(89%)	19,416 (79%)
Yes	12,774 (30%)	934 (11%)	5,210 (21%)

Framingham Original Cohort (used to construct the Original Framingham Risk Score (FRS)), and Pooled Cohort data (used to construct the Original Pooled Cohort Equations (PCE)). Entries are given as N (%) or Mean (SD).
EHR = Electronic Health Record

* Summarized from Table 1 at http://circ.ahajournals.org/content/117/6/743/tab-figures-data
+ Summarized from Table 2 at
http://circ.ahajournals.org/highwire/filestream/47233/field_highwire_adjunct_files/0/Risk_Assessment_Full_Work_Group_Report.doc

Supplemental Table 2. Original and Refitted coefficients and baseline risk for Framingham Risk Score Cox regression models

	Original FRS		Refitted FRS	
Variables	*Males*	*Females*	*Males*	*Females*
	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$	$\hat{\beta}$
Log(Age)	3.06	2.33	**3.75**	**4.44**
Log(Total Cholesterol)	2.12	1.21	**0.23**	**-0.07**
Log(HDL)	-0.93	-0.71	**-0.54**	**-0.47**
Log(Untreated SBP)	1.93	2.76	2.15	**2.13**
Log(Treated SBP)	2.00	2.82	2.20	**2.20**
Smoker (Yes/No)	0.65	0.53	0.56	**0.71**
Diabetes (Yes/No)	0.57	0.69	0.45	**0.51**
Average $\sum\widehat{\boldsymbol{\beta}}_{\boldsymbol{i}}\boldsymbol{x}_{\boldsymbol{i}}$	23.98	26.19	24.56	25.80
Baseline risk	0.049	0.021	0.046	0.026

FRS = Framingham Risk Score

Bolded coefficients in the Refitted FRS are those which differ significantly ($p < 0.05$) from the corresponding Original FRS coefficient.

Supplemental Table 3: Coefficients for Original and Refitted Pooled Cohort Equations models (White females)

Covariates	Original PCE $\hat{\beta}$	Refitted PCE $\hat{\beta}$
Log(Age)	–29.80	-24.49
Log(Age), Squared	4.88	4.22
Log(Total Cholesterol)	13.54	8.3
Log(Total Cholesterol) x Log(Age)	–3.11	-1.83
Log(HDL)	–13.58	-4.87
Log(HDL) x Log(Age)	3.15	1.07
Log(Treated SBP)	2.02	2.23
Log(Untreated SBP)	1.96	2.16
Smoker (Yes/No)	7.57	3.8
Smoker x Log(Age)	–1.67	-0.75
Diabetes (Yes/No)	0.66	0.42
Average $\sum\widehat{\boldsymbol{\beta}}_i\boldsymbol{x}_i$	-29.18	-16.95
Baseline risk	0.011	0.013

PCE = Pooled Cohort Equations

Supplemental Table 4: Coefficients for Original and Refitted Pooled Cohort Equations models (White males)

Covariates	Original PCE $\hat{\beta}$	Refitted PCE $\hat{\beta}$
Log(Age)	12.344	8.15
Log(Total Cholesterol)	11.853	6.09
Log(Total Cholesterol) x Log(Age)	–2.664	-1.23
Log(HDL)	–7.990	-3.5
Log(HDL) x Log(Age)	1.769	0.67
Log(Treated SBP)	1.797	1.67
Log(Untreated SBP)	1.764	1.6
Smoker (Yes/No)	7.837	4.28
Smoker x Log(Age)	–1.795	-0.9
Diabetes (Yes/No)	0.658	0.35
Average $\sum\widehat{\boldsymbol{\beta}}_i\boldsymbol{x}_i$	61.18	43.50
Baseline risk	0.0374	0.027

PCE = Pooled Cohort Equations

Supplemental Table 5: Coefficients for Original and Refitted Pooled Cohort Equations models (Black females)

Covariates	Original PCE $\hat{\beta}$	Refitted PCE $\hat{\beta}$
Log(Age)	17.114	2.74
Log(Total Cholesterol)	0.940	0.1
Log(HDL)	–18.920	0.05
Log(HDL) x Log(Age)	4.475	NA*
Log(Treated SBP)	29.291	2.63
Log(Treated SBP) x Log(Age)	–6.432	NA*
Log(Untreated SBP)	27.820	2.51
Log(Untreated SBP) x Log(Age)	–6.087	NA*
Smoker (Yes/No)	0.691	-0.75
Diabetes (Yes/No)	0.874	0.41
Average $\sum\hat{\boldsymbol{\beta}}_i\boldsymbol{x}_i$	86.61	23.71
Baseline risk	0.018	0.014

PCE = Pooled Cohort Equations

* Terms omitted from Cox model due to full model not converging when fit to given data.

Supplemental Table 6: Coefficients for Original and Refitted Pooled Cohort Equations models (Black males)

Covariates	Original PCE $\hat{\beta}$	Refitted PCE $\hat{\beta}$
Log(Age)	2.469	4.4
Log(Total Cholesterol)	0.302	-1.77
Log(HDL)	–0.307	-0.6
Log(Treated SBP)	1.916	5.62
Log(Untreated SBP)	1.809	5.65
Smoker (Yes/No)	0.549	0.53
Diabetes (Yes/No)	0.645	-0.04
Average $\sum\hat{\boldsymbol{\beta}}_i\boldsymbol{x}_i$	19.54	33.08
Baseline risk	0.043	0.020

PCE = Pooled Cohort Equations

Supplemental Table 7. Calibration and Discrimination of Original Framingham Risk Score and Pooled Cohort Equations models, for subpopulations defined by statin use and race.

	FRS*		PCE+	
	Original	**Refitted**	**Original**	**Refitted**
Non-statin users (N = 35,281^)				
Calibration statistic (p-value)	9.2 (0.03)	4.7 (0.2)	26.4 (<0.001)	8.4 (0.04)
C-index (95% CI)	0.742 (0.724-0.759)	0.758 (0.740-0.776)	0.754 (0.730-0.778)	0.753 (0.729-0.777)
Whites and blacks (N = 33,742^)				
Calibration statistic (p-value)	9.3 (0.03)	5.1 (0.17)	37.0 (<0.001)	15.7 (0.001)
C-index (95% CI)	0.738 (0.722-0.754)	0.751 (0.735-0.767)	0.745 (0.723-0.767)	0.744 (0.721-0.766)
Blacks only (N = 2,875^)				
Calibration statistic (p-value)	8.4 (0.04)	9.2 (0.03)	2.8 (0.42)	20.3 (<0.001)
C-index (95% CI)	0.699 (0.640-0.758)	0.703 (0.644-0.762)	0.725 (0.639-0.810)	0.696 (0.610-0.782)

PCE = Pooled Cohort Equations, FRS = Framingham Risk Score

* = Using FRS definition of CV events

+ = Using PCE definition of CV events

^ = Number of individuals in the test set used to evaluate predictions. Due to the 50/50 training test split, this is also the number of individuals among the N = 42,058 in the training set (distinct from the test set) used to refit the FRS and PCE models.

Supplemental Table 8. Results of a logistic regression evaluating the association between baseline characteristics and whether end of follow-up was due to disenrollment, among patients with less than 5 years of follow-up.

	Odds Ratio	P-value
(Intercept)	0.000	<0.01
Year of Enrollment	1.589	<0.01
Race (Other)	1.393	0.16
Race (White)	4.590	<0.01
Hispanic	1.190	0.74
Female	0.862	0.18
Age	1.008	0.19
SBP	1.001	0.79
BMI	0.994	0.54
HDL	1.005	0.29
TC	1.000	0.82
Current Smoker	0.992	0.91
Has Diabetes	0.891	0.53
Taking SBP Meds	1.054	0.67
Taking a Statin	1.203	0.20

SBP = Systolic Blood Pressure, BMI = Body Mass Index, HDL = High Density Lipoprotein, TC = Total Cholesterol