Supplemental Online Content

Duncan MS, Greevy RA, Tindle HA, et al. Inclusion of smoking data in cardiovascular disease risk estimation. *JAMA Cardiol*. Published online December 8, 2021. doi:10.1001/jamacardio.2021.4990

This supplemental material has been provided by the authors to give readers additional information about their work.

eMethods

ASCVD Risk Factor Definitions

To assess the statistical significance of adding smoking related variables to the 2013 PCE, our risk factor definitions were the same as those previously published by Goff et al.[1] in sex-specific models (age, total cholesterol, HDL cholesterol, SBP, use of antihypertensive medication, and diabetes mellitus).

Blood was drawn from participants at each examination cycle following an overnight fast of at least 10 hours. Biospecimens were stored at -20° (pre-1990 exams) to -80° C (post-1990 exams) until they were assayed. Total and HDL cholesterol were directly measured using standardized assays.[2] SBP was averaged based on two physician readings. Use of antihypertensive medication was self-reported at examination cycles 2-7 and verified by study staff upon review of provided medication at examination cycle 8. Participants were classified as having diabetes mellitus based on fasting blood glucose >126 mg/dL or use of medications for the treatment of diabetes mellitus.

Quantification of smoking status and intensity have been previously described.[3,4] Briefly, at the baseline examination, participants were categorized as "current," "former," or "never" smokers based on their responses to questions regarding current and prior smoking habits. From responses given for age at which the participant starting smoking, usual cigarettes smoked per day in the past, age at quitting (former smokers), and current number of cigarettes smoked per day (current smokers), we calculated ever smokers' pack years and years since quitting (YSQ) for former smokers. Never smokers were assigned a pack-year value of 0. For this analysis our smoking measures included current/former/never smoking status, pack-years, and YSQ.

YSQ is a conditionally relevant predictor in that it is intuitive for current smokers to have a value of 0 (i.e., they have not yet quit) and for former smokers' YSQ value to be greater than 0, but there is no relevant value for never smokers. Furthermore, we previously observed that YSQ was associated with ASCVD risk among heavy ever smokers ($\geq$20 pack-years).[4] Thus, we used the two-part predictor method described by Dziak and Henry[5] to adjust for this conditionally relevant predictor in heavy ever smokers only. The two variables that represent the effect of YSQ among heavy ever smokers are an indicator of when YSQ is relevant (heavy ever smokers), and a mean-centered YSQ value (YSQ*), where the mean is calculated in heavy ever smokers only. Never and non-heavy ever smokers' value of YSQ* was then set to 0 so that they would not impact (i.e., exert statistical leverage on) the estimation of the effect of YSQ* among heavy ever smokers.

Reasoning for modifications to PCE

We first fit the sex-specific 2013 PCE in our data, re-estimating both the beta coefficients and baseline hazard in our sample for fair model comparison. We then used the same predictors as the 2013 PCE with three changes: 1) inclusion of continuous variables on their natural scale (i.e., not logarithmically transformed); 2) up to third order polynomials on continuous variables (and their interactions) to account for non-linearity; and 3) adjustment for antihypertensive use and measured SBP rather than including separate terms for treated/untreated SBP. Justifications for these alterations are provided below.

Natural Scale and Polynomials

We made this change because using log-transformed predictors, as were used in the 2013 PCE, assumes a logarithmic association between the risk factor and log-hazard of ASCVD and therefore imposes a specific "shape" on the association that may or may not hold. By including continuous variables on their natural scale and allowing for up to 3rd order polynomials on these continuous risk factors, we avoid imposing a specific association on these relationships and allow the model to estimate an association that more accurately reflects the relationship between risk factors and log hazards of ASCVD that appears in our data without making assumptions on what those associations "should" look like. Age, total cholesterol, HDL cholesterol, pack-years smoked, and years since quitting were modeled with a second order polynomial (quadratic function); SBP was modeled with a third order polynomial (cubic function).

Inclusion of SBP in the model

In the PCE, SBP is represented by two terms – untreated SBP and treated SBP. For those receiving antihypertensives, their untreated SBP is set to 0; similarly, those who are not on antihypertensives have their treated SBP set to 0. This approach works and accurately estimates the effects of treated/untreated SBP if untreated SBP is centered on its mean among the untreated and treated SBP is centered on its mean in the treated[5]. However, the mean value that SBP is centered on will be sample-dependent. In our approach, which includes measured SBP and a

binary indicator of antihypertensive treatment, which are interacted with one another, we avoid the issue of sample-dependency and are still able to capture the effect of treated vs. untreated SBP on ASCVD risk.

Supplemental References

1. Goff DC, Lloyd-jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk. *J Am Coll Cardiol*. 2014;63(25):2886. doi:10.1016/j.jacc.2014.02.606

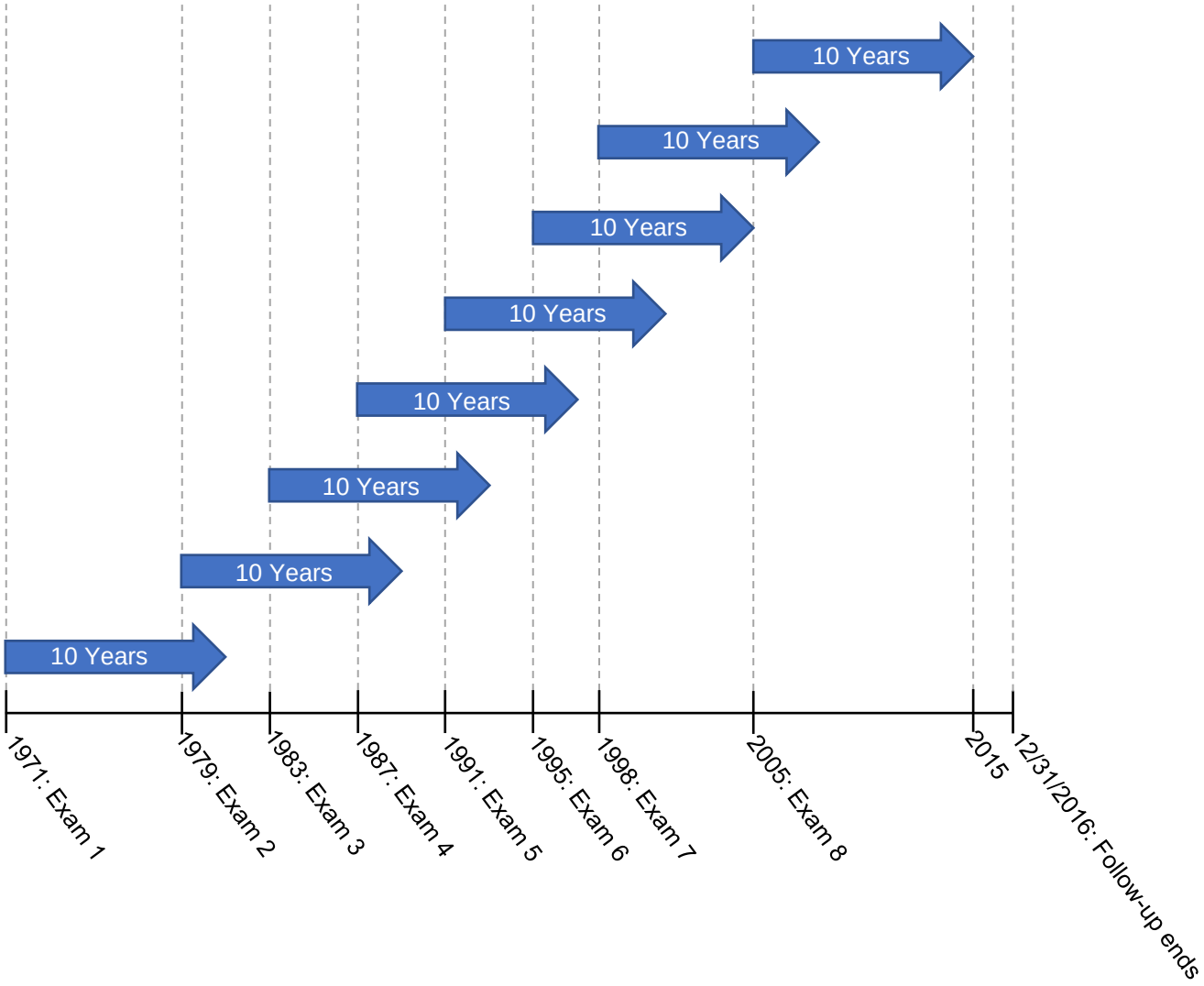
2. Cobain MR, Pencina MJ, D'Agostino RB, Vasan RS. Lifetime risk for developing dyslipidemia: the Framingham Offspring Study. *Am J Med*. 2007;120(7):623-630. doi:10.1016/j.amjmed.2006.12.015

3. Tindle HA, Stevenson Duncan M, Greevy RA, et al. Lifetime Smoking History and Risk of Lung Cancer: Results From the Framingham Heart Study. *J Natl Cancer Inst*. 2018;110(11). doi:10.1093/jnci/djy041

4. Duncan MS, Freiberg MS, Jr RAG, Kundu S, Vasan RS, Tindle HA. Association of Smoking Cessation With Subsequent Risk of Cardiovascular Disease. *J Am Med Assoc*. 2019;322(7):642-650. doi:10.1001/jama.2019.10298

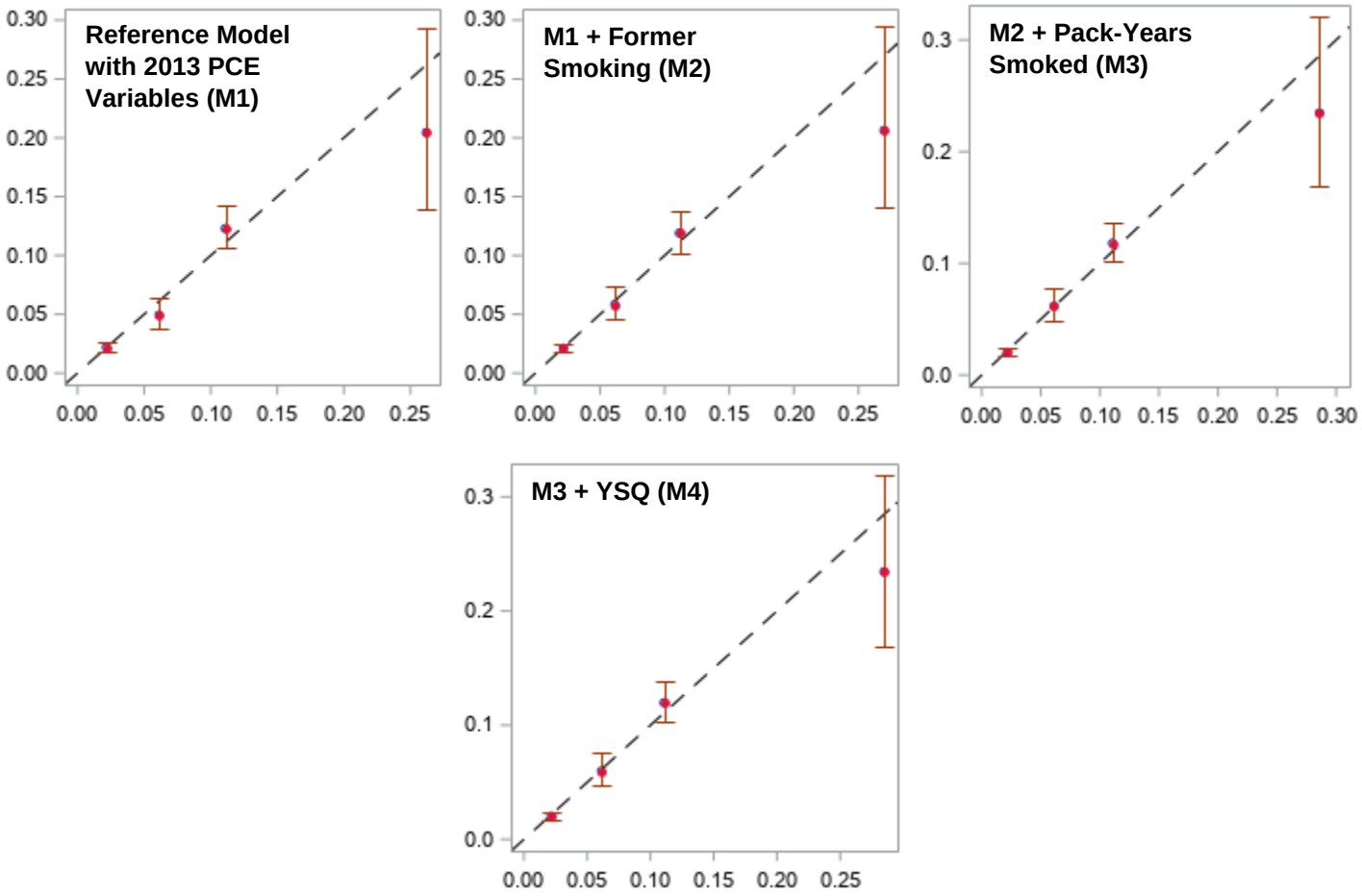
5. Dziak JJ, Henry KL. Two-Part Predictors in Regression Models. *Multivariate Behav Res*. 2017;52(5):551-561. doi:10.1080/00273171.2017.1333404

eFigure 1. Timeline for FHS Offspring Examinations



Timing of FHS Offspring examinations and corresponding 10-year follow-up periods for hypothetical FHS Offspring participant

eFigure 2. Calibration Plots in Men: Observed Probability (y-axis) vs. Predicted Probability (x-axis)

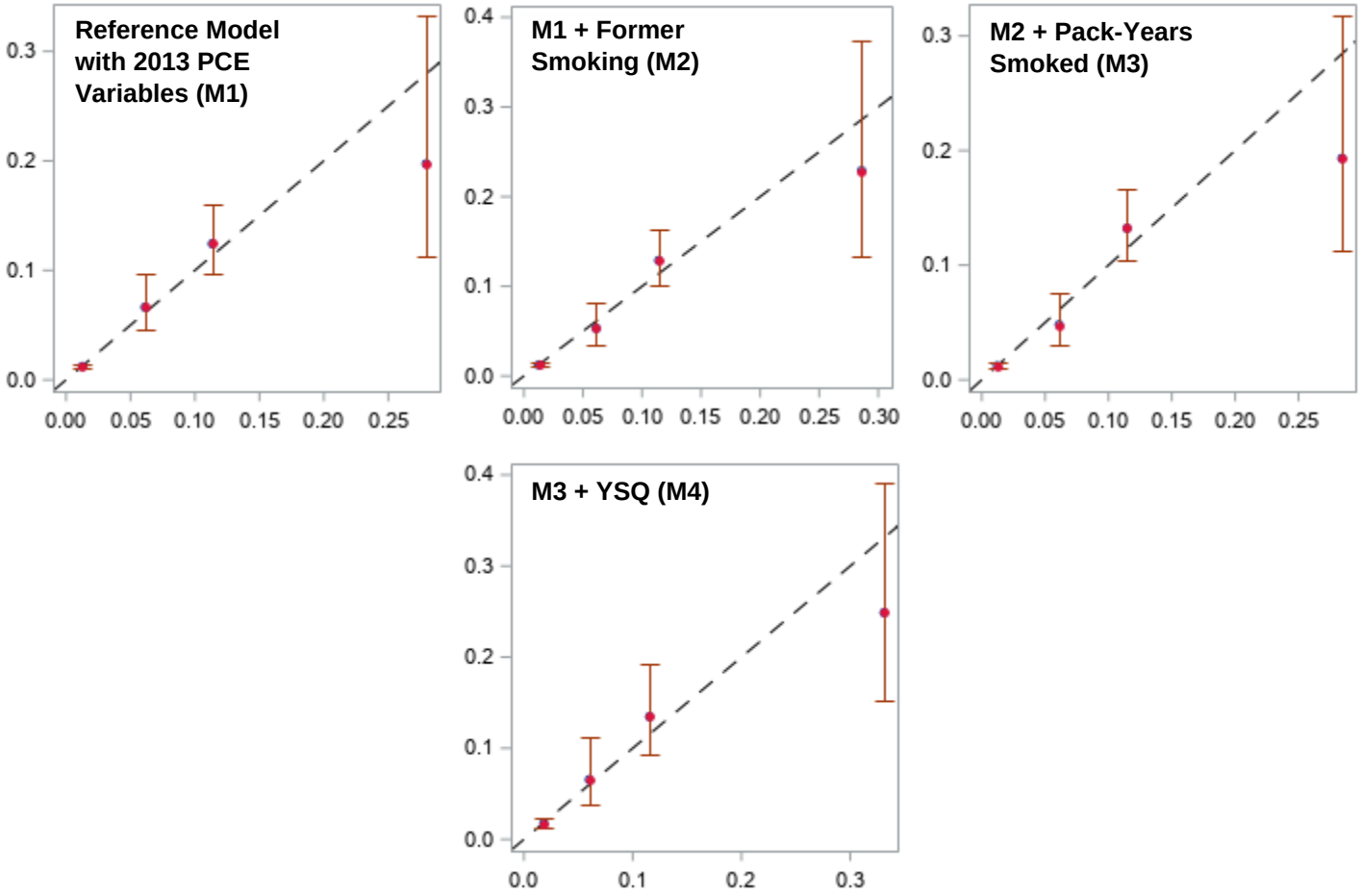


Abbreviations: CFN, current/former/never; M, model; PCE, pooled cohort equations; YSQ, years since quitting

Predictors in the reference model (M1) were as follows: age, total cholesterol, HDL cholesterol, SBP, antihypertensive use, current smoking

10-year ASCVD risk cutpoints of <5%, 5-7.49%, 7.5-19.9%, and ≥20% were used to be consistent with the 2013 PCE, which defined these groups as low, borderline, intermediate, and high risk, respectively

eFigure 3. Calibration Plots in Women: Observed Probability (y-axis) vs. Predicted Probability (x-axis)

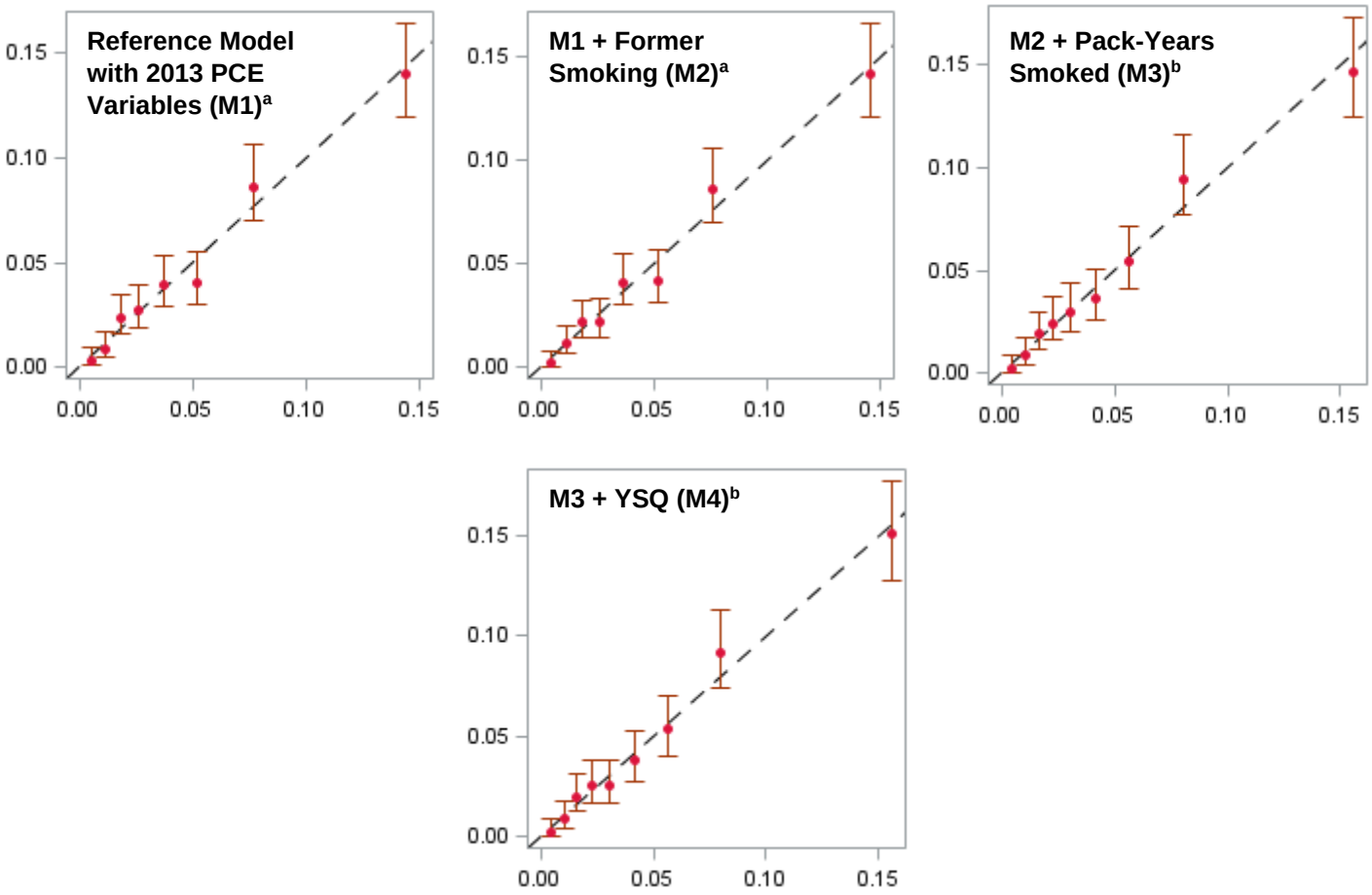


Abbreviations: CFN, current/former/never; M, model; PCE, pooled cohort equations; YSQ, years since quitting

Predictors in the reference model (M1) were as follows: age, total cholesterol, HDL cholesterol, SBP, antihypertensive use, current smoking

10-year ASCVD risk cutpoints of <5%, 5-7.49%, 7.5-19.9%, and ≥20% were used to be consistent with the 2013 PCE, which defined these groups as low, borderline, intermediate, and high risk, respectively

eFigure 4. Observed vs. Predicted Plots in Men by Deciles: Observed Probability (y-axis) vs. Predicted Probability (x-axis)



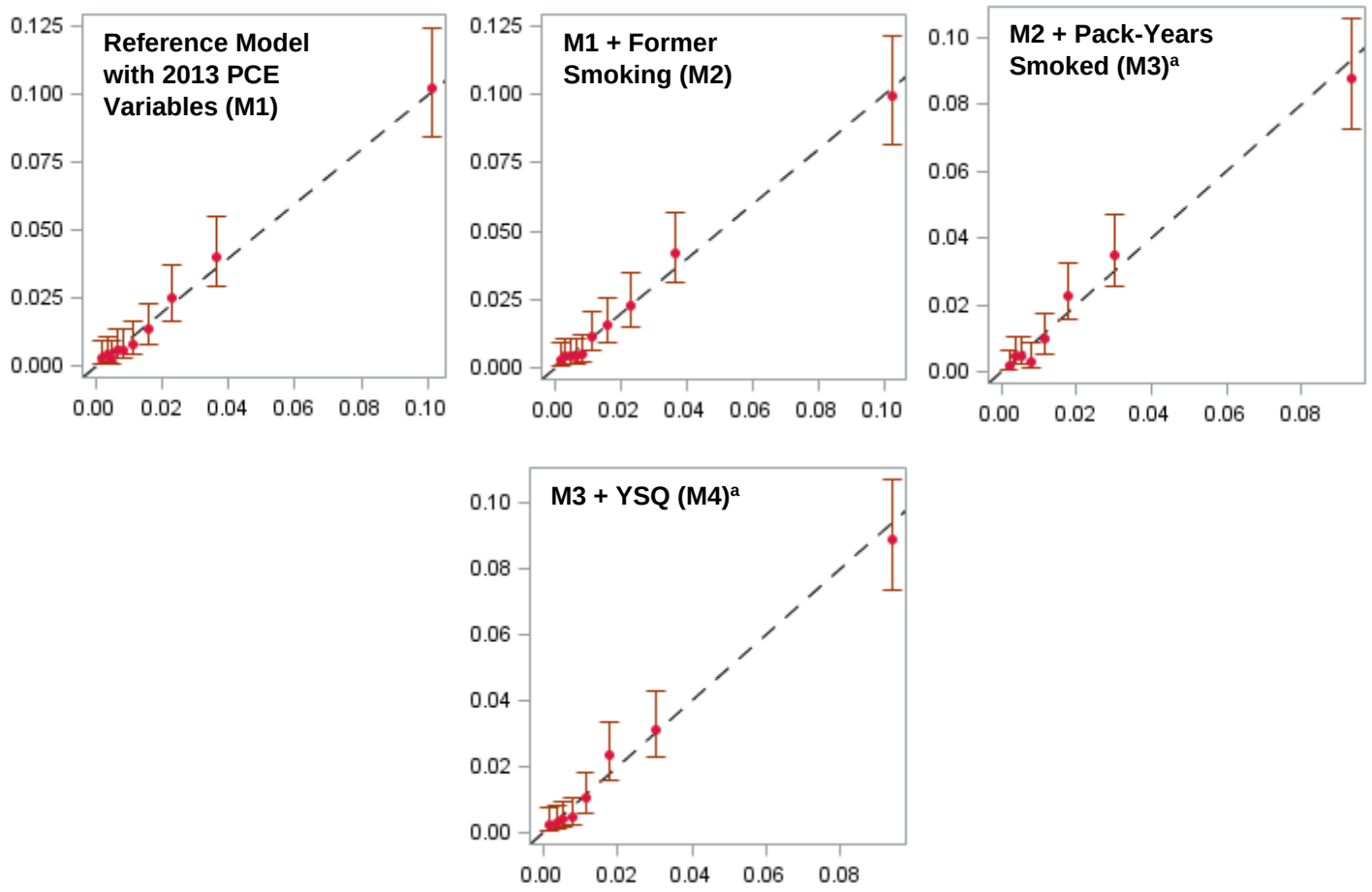
Abbreviations: M, model; PCE, pooled cohort equations; YSQ, years since quitting

Predictors in the reference model (M1) were as follows: age, total cholesterol, HDL cholesterol, SBP, antihypertensive use, current smoking

a Predicted vs. observed probabilities for model 1 are shown by 8 quantiles rather than 10 due to low number of events in some groups, precluding the ability to calculate fit statistics

b Predicted vs. observed probabilities for model 1 are shown by 9 quantiles rather than 10 due to low number of events in some groups, precluding the ability to calculate fit statistics

eFigure 5. Observed vs. Predicted Plots in Women by Deciles: Observed Probability (y-axis) vs. Predicted Probability (x-axis)

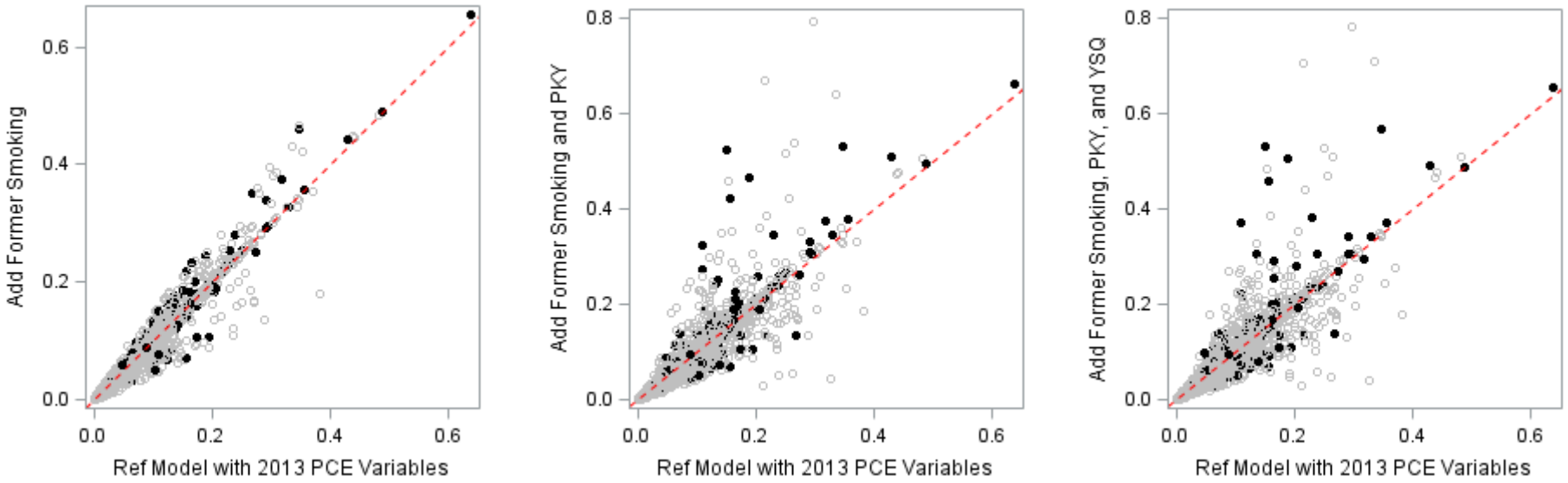


Abbreviations: M, model; PCE, pooled cohort equations; YSQ, years since quitting

Predictors in the reference model (M1) were as follows: age, total cholesterol, HDL cholesterol, SBP, antihypertensive use, current smoking

a Predicted vs. observed probabilities for models 4 and 5 are displayed by octiles rather than deciles because there were too few events in the groups when deciles were used

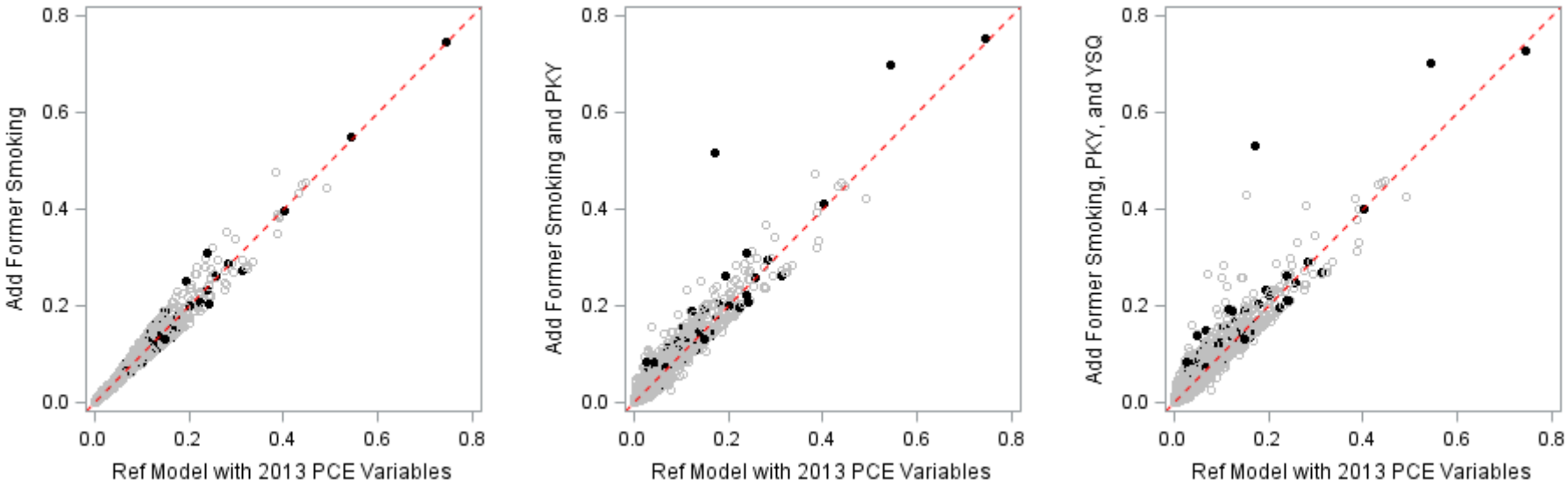
eFigure 6. Predicted Probability of ASCVD in Men: Models with Additional Smoking Data (y-axis) vs. Reference Model with 2013 PCE Variables (x-axis)



Abbreviations: PCE, pooled cohort equations; PKY, pack-years; YSQ, years since quitting

Filled circles indicate individuals who had an ASCVD event, open circles are nonevents

eFigure 7. Predicted Probability of ASCVD in Women: Models with Additional Smoking Data (y-axis) vs. Reference Model with 2013 PCE Variables (x-axis)



Abbreviations: PCE, pooled cohort equations; PKY, pack-years; YSQ, years since quitting

Filled circles indicate individuals who had an ASCVD event, open circles are nonevents

eTable 1. Reclassification of Risk in Heavy Ever Smoking (≥20 Pack-Years) Men

Model with 2013 PCE Variables plus former smoking status, pack-years smoked, years since quitting

Non-Events

2013 PCE[a]	<5%	[5%, 7.5%)	[7.5%, 20%)	≥20%
<5%	1557	68	9	0
5-7.49%	258	404	119	1
7.5-19.9%	25	150	705	55
≥20%	5	4	25	54
Totals:	467	252	2720	

Events

2013 PCE[a]	<5%	[5%, 7.5%)	[7.5%, 20%)	≥20%
<5%	39	5	0	0
5-7.49%	5	19	10	0
7.5-19.9%	2	11	99	10
≥20%	0	0	2	19
Totals:	25	20	176	

Predicted risk under 2013 PCE in rows; predicted risk under our final model in columns

Green fill indicates a movement in the correct direction (lower risk for non-events, higher risk for events); Red fill indicates a move in the incorrect direction (higher risk for non-events, lower risk for events); Gray fill indicates no reclassification

Among non-events, 467 (13.6%) were correctly reclassified as lower risk and 252 (7.3%) were incorrectly reclassified as higher risk under our final model compared to the 2013 PCE

Among events, 25 (11.3%) were correctly reclassified as higher risk and 20 (9.0%) were incorrectly reclassified as lower risk under our final model compared to the 2013 PCE

a 2013 PCE with baseline hazard and coefficients re-estimated in the current sample for fair model comparison. Otherwise, results would appear artificially favorable since the Model with 2013 PCE Variables plus 3-level Smoking Status, Pack-Years, and Years Since Quitting was developed entirely in this sample and would necessarily outperform the 2013 PCE which were developed in an external sample

eTable 2. Reclassification of Risk in Heavy Ever Smoking (≥20 Pack-Years) Women

Model with 2013 PCE Variables plus former smoking status, pack-years smoked, years since quitting

Non-Events

2013 PCE[a]	<5%	[5%, 7.5%)	[7.5%, 20%)	≥20%
<5%	2094	57	22	0
5-7.49%	97	70	32	0
7.5-19.9%	28	59	130	22
≥20%	0	1	16	24
Totals:	201	133	2318	

Events

2013 PCE[a]	<5%	[5%, 7.5%)	[7.5%, 20%)	≥20%
<5%	32	4	3	0
5-7.49%	5	4	4	0
7.5-19.9%	1	4	18	5
≥20%	0	0	1	8
Totals:	16	11	62	

Predicted risk under 2013 PCE in rows; predicted risk under our final model in columns

Green fill indicates a movement in the correct direction (lower risk for non-events, higher risk for events); Red fill indicates a move in the incorrect direction (higher risk for non-events, lower risk for events); Gray fill indicates no reclassification

Among non-events, 201 (7.6%) were correctly reclassified as lower risk and 133 (5.0%) were incorrectly reclassified as higher risk under our final model compared to the 2013 PCE

Among events, 16 (18.0%) were correctly reclassified as higher risk and 11 (12.4%) were incorrectly reclassified as lower risk under our final model compared to the 2013 PCE

a 2013 PCE with baseline hazard and coefficients re-estimated in the current sample for fair model comparison. Otherwise, results would appear artificially favorable since the Model with 2013 PCE Variables plus 3-level Smoking Status, Pack-Years, and Years Since Quitting was developed entirely in this sample and would necessarily outperform the 2013 PCE which were developed in an external sample

eTable 3. Validation Sample Model Results

Model No.	Description	-2 Log L	df	Δdf	LR χ^2	LR p-value	R^2	c-statistic	Δ Harrell's c [95% CI]	NRI(>0) [95% CI]	Relative IDI [95% CI]
Men (n=2715, 91 Incident ASCVDs)											
1	Model with 2013 PCE Variables[a]	1325.69	23	--	--	--	7.53%	0.7573	--	--	--
2	M1 + Former Smoking	1324.11	26	3	1.58	0.66	7.67%	0.7601	0.003 [-0.004, 0.010]	0.044 [-0.190, 0.259]	0.024 [-0.026, 0.082]
3	M2 + Pack-Years	1319.66	31	8	6.03	0.64	8.06%	0.7674	0.010 [-0.001, 0.022]	0.164 [-0.062, 0.372]	0.089 [-0.0002, 0.015]
4	M3 + YSQ	1315.30	36	13	10.39	0.66	8.44%	0.7695	0.012 [-0.002, 0.029]	0.289 [0.077, 0.512]	0.141 [0.033, 0.257]
Women (n=3318, 47 Incident ASCVDs)											
1	Model with 2013 PCE Variables[a]	661.695	23	--	--	--	12.94%	0.8016	--	--	--
2	M1 + Former Smoking	659.416	26	3	2.28	0.52	13.27%	0.8110	0.009 [-0.006, 0.026]	0.062 [-0.254, 0.344]	0.005 [-0.017, 0.029]
3	M2 + Pack-Years	657.545	31	8	4.15	0.84	13.55%	0.8116	0.010 [-0.010, 0.031]	0.189 [-0.108, 0.486]	0.041 [-0.034, 0.122]
4	M3 + YSQ	651.706	36	13	9.99	0.69	14.39%	0.8167	0.015 [-0.015, 0.047]	0.248 [-0.049, 0.509]	0.140 [0.034, 0.270]

Abbreviations: current/former/never (CFN); confidence interval (CI); degrees of freedom (df); integrated discrimination improvement (IDI); likelihood ratio (LR); continuous net reclassification improvement (NRI(>0)); years since quitting smoking (YSQ)

a This is the reference for Models 5 when calculating the likelihood ratio test, Δ Harrell's c-statistic, NRI(>0), and relative IDI