ORIGINAL RESEARCH

Atherosclerotic Cardiovascular Disease Risk Prediction in Disaggregated Asian and Hispanic Subgroups Using Electronic Health Records

Fatima Rodriguez, MD, MPH; Sukyung Chung, PhD; Manuel R. Blum, MD; Adrien Coulet, PhD; Sanjay Basu, MD, PhD; Latha P. Palaniappan, MD, MS

***Background*—**Risk assessment is the cornerstone for atherosclerotic cardiovascular disease (ASCVD) treatment decisions. The Pooled Cohort Equations (PCE) have not been validated in disaggregated Asian or Hispanic populations, who have heterogeneous cardiovascular risk and outcomes.

***Methods and Results*—**We used electronic health record data from adults aged 40 to 79 years from a community-based, outpatient healthcare system in northern California between January 1, 2006 and December 31, 2015, without ASCVD and not on statins. We examined the calibration and discrimination of the PCE and recalibrated the equations for disaggregated race/ethnic subgroups. The cohort included 231 622 adults with a mean age of 53.1 (SD 9.7) years and 54.3% women. There were 56 130 Asian (Chinese, Asian Indian, Filipino, Japanese, Vietnamese, and other Asian) and 19 760 Hispanic (Mexican, Puerto Rican, and other Hispanic) patients. There were 2703 events (332 and 189 in Asian and Hispanic patients, respectively) during an average of 3.9 (SD 1.5) years of follow-up. The PCE overestimated risk for NHWs, African Americans, Asians, and Hispanics by 20% to 60%. The extent of overestimation of ASCVD risk varied by disaggregated racial/ethnic subgroups, with a predicted-to-observed ratio of ASCVD events ranging from 1.1 for Puerto Rican patients to 1.9 for Chinese patients. The PCE had adequate discrimination, although it varied significantly by race/ethnic subgroups (C-indices 0.66–0.83). Recalibration of the PCE did not significantly improve its performance.

***Conclusions*—**Using electronic health record data from a large, real-world population, we found that the PCE generally overestimated ASCVD risk, with marked heterogeneity by disaggregated Asian and Hispanic subgroups. **(*J Am Heart Assoc.* 2019;8:e011874. DOI: 10.1161/JAHA.118.011874.)**

Key Words: disparities • electronic health records • prevention • risk assessment

Hispanic and Asian populations are the 2 most rapidly growing minority groups in the United States and are expected to double in size by 2050 to 110 and 30 million, respectively.[1,2] Despite this projected growth, there is little data on the cardiovascular health of these groups.[3–7] When Hispanic and Asian individuals are studied, they are often aggregated into 1 group, masking important differences between distinct subpopulations. Prior work has documented marked heterogeneity in cardiovascular disease risk and mortality patterns in disaggregated Hispanic and Asian subgroups.[8–15]

Risk assessment is the foundation for guiding atherosclerotic cardiovascular disease (ASCVD) treatment decisions, especially for initiation of lipid-lowering therapy. The 2013 prevention guidelines released by the American Heart Association (AHA) and the American College of Cardiology (ACC) recommended statin therapy for primary prevention based on 10-year estimated ASCVD risk.[16] The recent 2018 update to these guidelines similarly relies on the Pooled Cohort Equations (PCE) to guide recommendations for risk stratification and statin treatment decisions.[17] This update highlights the uncertainty of the PCE performance in diverse race/ethnic groups such as Asians and Hispanics since the PCE was derived from non-Hispanic white (NHW) and African American populations.

From the Divisions of Cardiovascular Medicine (F.R., M.R.B.) and Primary Care and Population Health (S.C., L.P.P.), Stanford University School of Medicine, Stanford, CA; Palo Alto Foundation Research Institute, Palo Alto, CA (S.C.); Department of General Internal Medicine, Inselspital, Bern University Hospital, University of Bern, Switzerland (M.R.B.); Université de Lorraine, CNRS, Inria, LORIA, Nancy, France (A.C.); Stanford Center for Biomedical Informatics Research (A.C.) and Departments of Medicine and of Health Research and Policy, Centers for Primary Care and Outcomes Research and Center for Population Health Sciences (S.B.), Stanford University, Stanford, CA.

Accompanying Tables S1 through S3 are available at https://www.ahajournals.org/doi/suppl/10.1161/JAHA.118.011874

Correspondence to: Fatima Rodriguez, MD, MPH, Division of Cardiovascular Medicine, Stanford University, 870 Quarry Road, Falk CVRC, Stanford, CA 94305-5406. E-mail: frodrigu@stanford.edu

Received February 13, 2019; accepted June 4, 2019.

Clinical Perspective

What Is New?

- We found that the Pooled Cohort Equations generally overestimated risk among disaggregated Hispanic and Asian patients but that the degree of overestimation varied significantly by racial/ethnic subgroup.

What Are the Clinical Implications?

- It is reasonable to use the Pooled Cohort Equations among Hispanic and Asian patients since its performance is comparable with that of non-Hispanic whites.
- However, caution must be used in interpreting results by disaggregated patient racial/ethnic subgroups, since the degree of risk overestimation is different.
- Electronic health record data offer great promise in developing personalized, risk-specific cardiovascular risk prediction models.

The PCE computation requires sex, age, race (African American/NHW/other), total and high-density lipoprotein (HDL) cholesterol, systolic blood pressure, blood pressure treatment status, smoking status, and diabetes mellitus status.[2]

Whether the PCE adequately identifies risk for diverse populations remains controversial and several recent analyses of calibration and discrimination of the risk scores have yielded mixed results.[18–21] The ACC/AHA Work Group, which developed the risk estimator algorithm, acknowledges that it was designed for use only in men and women of NHW or African American descent, and that the risk estimator may not accurately predict risk in other racial/ethnic groups,[2] such as Hispanic and Asian populations, which together make up over a quarter of the US population.[1] Specifically, the Pooled Cohort Equations may overestimate risk in Mexican Americans and East Asians (Japanese, Korean, and Vietnamese) and underestimate risk in Puerto Ricans and South Asians.[2,16] However, this assumption has not been validated and the extent of over- or underestimation of ASCVD risk in specific populations remains unknown.

Methods and Results

The data used in this study will not be made publicly available.

Study Sample

The study sample was selected from electronic health records (EHR) of adults aged between 40 and 79 years in a large, community-based outpatient healthcare system in northern California between January 1, 2006 and December 31, 2015. To minimize missing information or incomplete event ascertainment, included patients were required to have at least 2 outpatient visits that were at least 1 year apart. The index date was defined as the first available cholesterol or blood pressure measurement date after an outpatient visit. Patients with pre-existing ASCVD, atrial fibrillation, or heart failure identified by the *International Classification of Diseases, Ninth Revision, Clinical Modification* (*ICD-9-CM*) coding scheme were excluded. As per the PCE derivation guideline, patients with atrial fibrillation at baseline were excluded because of the strong relationship between atrial fibrillation and stroke and the documented need for risk reducing therapies.[2]

We also excluded patients without complete information on cardiovascular risk factors necessary for ASCVD risk calculation or those who were on statins or other lipid-lowering therapies at baseline.

Race/ethnicity was self-reported (83.2%) or inferred (16.8%) based on the Social Security Record database as previously described.[22] The study sample included NHWs, African Americans, Asians, and Hispanics. Asians were disaggregated into Chinese, Asian Indian, Filipino, Japanese, Korean, Vietnamese, and other Asians. Based on US Census categories, Hispanics were disaggregated into Mexican, Puerto Rican, Cuban, and other Hispanics. Because of small sample size (n=223), Cubans were combined with other Hispanics.

ASCVD Events

In line with the ACC/AHA Work Group guidelines, we defined ASCVD events as the first fatal or non-fatal acute myocardial infarction or stroke.[2] Acute myocardial infarction was defined by *ICD-9-CM* codes 410.x0 and 410.x1.[20,23] Non-fatal incident ischemic and hemorrhagic stroke events were defined based on *ICD-9-CM* codes 433.x1, 434.x1 or 436.0.[20,24] Coronary heart disease followed by death within a year, or fatal or non-fatal ischemic or hemorrhagic stroke followed by death within a year were identified as ASCVD deaths. Death information was retrieved from the EHR and Social Security records.

Prediction of ASCVD Risk Based on Pooled Cohort Equation and Recalibration

The 5-year ASCVD predicted risk was calculated for adults between ages 40 to 79 using the published parameter estimates from the PCE.[2,20] We used the parameters developed for NHWs to estimate predicted values for Asian and Hispanic populations. The 5-year predicted risk values were adjusted to reflect varying follow-up time, assuming proportional hazard, which was validated in our data (global test of proportional χ^2 (11)=0.115 and plotting of Schoenfeld

residuals where slopes for all covariates were nearly flat). Patients were censored at the time of the first event.

The PCE was then recalibrated from the study population using a Cox proportional hazard model with the same set of covariates and specifications as the PCE, and using 10-fold cross-validation. We randomly partitioned the sample into 10 groups, and used a 9/10th (training set) to estimate coefficients and, by applying the coefficients to the remaining 1/10th (test set), computed predicted risk for the test set. After repeating this procedure with the other 9 training sets and 9 test sets, all the patients in the sample were assigned with a predicted 5-year risk values derived from the cross-validation. Since the PCE is defined for NHWs and African Americans, we used the PCE specifications for NHWs in recalibrating Asian and Hispanic subgroups. Since there is significant racial admixture among Hispanics, we also performed a sensitivity analysis using the PCE specifications for African Americans in recalibrating Hispanic subgroups.

Calibration and Discrimination of the PCE

We calculated the predicted incidence of ASCVD during the follow-up period based on the PCE and compared those with the observed events and drew statistical inference based on the paired *t*-test. Patients were then categorized into 4 groups according to their PCE 10-year predicted ASCVD risk which was computed using parameter estimates from the PCE: <5%, ≥5 and <7.5%, ≥7.5 and <10%, and ≥10%. For each risk subgroup, we calculated and plotted predicted ASCVD incidence during the follow-up period based on the PCE and compared those to the observed events to evaluate extent of over- or underestimation. The Hosmer-Lemeshow χ^2 statistic was used to estimate the calibration of the PCE estimator.

C-indices were computed for predicted risk based on the PCE and recalibrated equations, respectively, for each racial/ethnic group, and statistical difference was calculated by differences in the C-indices and 95% CI derived from 1000 bootstrapped samples. All analyses were performed using Stata statistical software version 13.1.

This study was approved by the Stanford Institutional Review Board for human subjects, which waived the need for patient consent.

The study sample assembly is shown in Figure 1. We restricted the analysis to patients aged between 40 and 79 years with at least 2 outpatient visits that were at least 1 year apart (n=273 585). We excluded 14 179 patients with pre-existing ASCVD, atrial fibrillation, or heart failure and 27 784 participants with missing risk factor information, missing race/ethnicity, or on lipid-lowering therapy at baseline. The final study population included a total of 231 622 patients with a mean follow-up of 3.9 years (SD 1.5).

The study population (Table 1) had a mean age of 53.1 (SD 9.7) and 54.3% were women. There were 56 130 Asian and 19 760 Hispanic patients. There was significant heterogeneity in ASCVD risk factors by disaggregated subgroup. Asian Indians (mean age 47.7) were significantly younger than NHWs and other Asian subgroups in the sample. Chinese patients had a diabetes mellitus prevalence of 6.5% compared with 17.6% for Filipinos. Within Hispanic subgroups, Mexicans had the highest prevalence of diabetes mellitus (16.5%), and

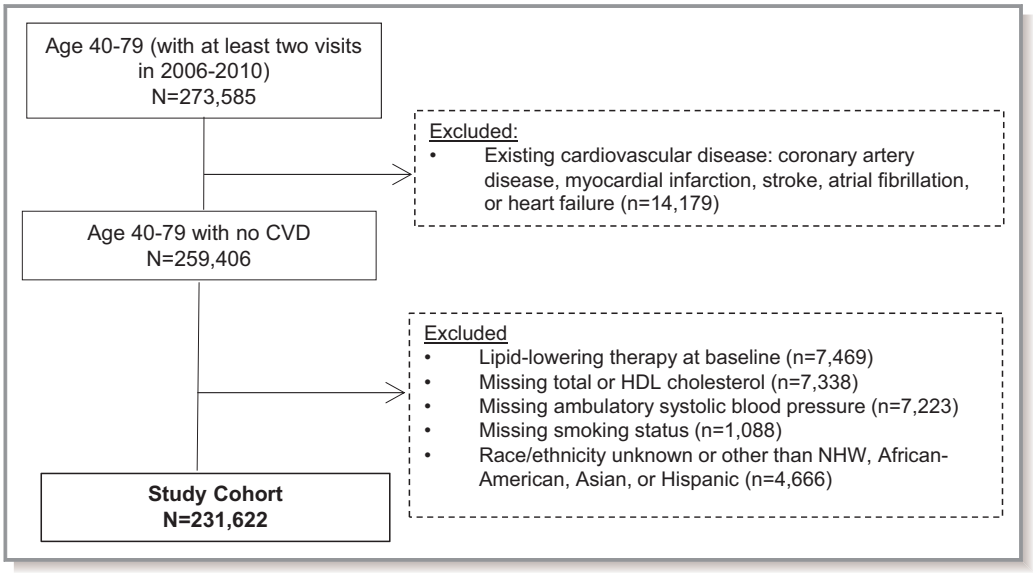

Figure 1. Study cohort. *Pre-existing cardiovascular disease was defined by the following *International Classification of Diseases, Ninth Revision, Clinical Modification* (*ICD-9-CM*) codes: Atrial fibrillation: 427.31; heart failure: 428*; coronary artery disease: 411*, 413*, 414*; myocardial infarction; 410*; and stroke: 430–434*, 436*. ASCVD indicates atherosclerotic cardiovascular disease; CVD, cardiovascular disease; HDL, high-density lipoprotein cholesterol; NHW, non-Hispanic white.

Puerto Ricans had the highest prevalence of smoking (8.1%) as compared with other Hispanics and NHWs.

There were a total 2703 events (332 and 189 in Asian and Hispanic patients, respectively) during the study period. The first occurrence of ASCVD by disaggregated race/ethnic group and event type is shown in Table S1. Asians and Hispanics had an aggregated event rate of 0.9% and 1.3%, respectively. These rates differed substantially by subgroup. The highest event rates were observed among Vietnamese (1.2%), Filipino (1.2%), and Puerto Rican patients (1.5%).

The differences between the PCE predicted and observed event rates by race/ethnicity are shown in Table 2. Overall, the PCE overestimated risk for NHWs, African Americans, Asians, and Hispanics by 20% to 60%. However, the extent of overestimation of ASCVD risk varied by disaggregated racial/ethnic subgroups, with a predicted-to-observed ratio of ASCVD events ranging from 1.1 for Puerto Rican patients to 1.9 for Chinese patients. The PCE slightly underestimated risk among Vietnamese patients, with an predicted-to-observed ratio of 0.9, although this difference was not statistically significant (P=0.655).

Observed versus predicted ASCVD risk by race/ethnicity and pre-specified 10-year risk categories is shown in Figure 2. Observed versus predicted ASCVD risk differed significantly within all race/ethnic groups. The PCE generally overestimated risk across race/ethnic groups. However, there was significant heterogeneity within disaggregated Hispanic and Asian subgroups across risk categories. Most of the overestimation of risk occurred in the higher risk categories. The PCE was better calibrated for Korean (χ^2=4.5, P=0.107), Puerto Rican (χ^2=1.9, P=0.40), and other Hispanic patients (χ^2=5.9, P=0.1) (Table S2).

The discrimination of predicted values from the PCE and recalibrated PCE by race/ethnicity is shown in Table 3. The C-index for NHWs was 0.77, 95% CI (0.76–0.78) and 0.74, 95% CI (0.66–0.81) for African Americans. Overall, the PCE had adequate discrimination for Hispanics and Asians with an overall C-index of 0.78, 95% CI (0.76–0.81) and 0.78; 95% CI (0.75–0.80), respectively. However, there was marked heterogeneity by racial/ethnic subgroup. The C-index ranged from 0.66 in Puerto Rican patients to 0.83 for Korean patients. Recalibration did not significantly improve the performance of the PCE in the study population. Using specifications for African Americans for the PCE recalibration among Hispanic subgroups, similarly failed to significantly improve its performance (Table S3).

Table 1. Characteristics of the Study Sample by Race/Ethnicity*

Race/Ethnic Subgroup	n	Follow-Up (y), Mean (SD)	Age (y), Mean (SD)	Female (%)	SBP (mm Hg) Mean (SD)	On Anti-Hypertensive Medication (%)†	HDL Cholesterol (mg/dL), Mean (SD)	Total Cholesterol, mg/dL, Mean (SD)	Type 2 Diabetes mellitus (%)	Current Smoker (%)
Overall	231 622	3.9 (1.5)	53.1 (9.7)	54.3%	123.0 (16.5)	3.8%	56.4 (16.5)	198.1 (36.5)	8.3%	5.1%
NHW (ref)	151 615	4 (1.5)	54.3 (9.6)	53.6%	124.1 (16.4)	4.0%	57.5 (16.9)	199.7 (36.7)	6.8%	5.6%
African American	4117	3.9 (1.5)*	52.4 (9.1)*	56.4%	128.5 (17.0)*	5.9%*	56.9 (16.3)	195.7 (38.1)*	14.8%*	9.6%*
Asian	56 130	3.9 (1.6)*	50.4 (9.5)*	54.9%	119.5 (16.3)*	3.0%*	54.7 (15.4)*	193.7 (35.2)*	10.0%*	3.1%*
Hispanic	19 760	3.8 (1.6)*	51.8 (9.3)*	57.3%*	124.2 (16.4)	4.3%	52.6 (15.1)*	199.0 (37.3)	14.3%*	5.8%
Chinese	23 171	4.0 (1.5)	51.1 (9.6)*	56.7%*	117.8 (16.0)*	2.4%*	56.5 (15.2)*	191.7 (34.0)*	6.5%	2.0%*
Asian Indian	13 815	3.8 (1.6)	47.7 (8.5)*	43.1%*	119.2 (15.5)*	3.0%*	47.9 (13.0)*	190.5 (34.7)*	13.4%*	3.0%*
Filipino	6220	4.0 (1.5)*	52.1 (9.4)*	63.4%*	126.3 (17.1)*	4.7%	55.8 (14.9)*	199.1 (37.7)	17.6%*	5.8%
Japanese	3825	4.1 (1.4)	53.7 (10.2)*	63.4%*	122.3 (16.8)*	2.9%*	62.0 (17.2)*	202.7 (36.0)*	10.4%*	3.0%*
Korean	1793	3.6 (1.6)	49.6 (9.9)*	61.5%*	117.9 (15.6)*	2.6%	57.5 (15.5)	193.8 (34.9)*	8.9%*	5.9%
Vietnamese	2093	3.9 (1.5)*	49.9 (8.4)*	58.4%*	118 (15.5)*	2.8%	56.5 (15.1)	200.9 (34.4)	7.6%	3.7%*
Other Asian	5213	3.4 (1.7)	50.8 (9.4)*	57.7%*	119.1 (16.5)*	3.2%	56.4 (15.9)*	195.4 (35.9)*	8.1%*	3.6%*
Mexican	8655	3.9 (1.5)*	51.6 (9.3)*	57.5%*	125.0 (16.4)*	4.3%	51.3 (14.6)*	199.2 (37.5)	16.5%*	5.7%
Puerto Rican	457	4.0 (1.5)*	51.4 (9.3)*	58.9%*	122.3 (15.3)*	5.0%	53.2 (14.8)*	199.1 (37.0)	13.3%*	8.1%
Other Hispanic	10 643	3.6 (1.7)	52.0 (9.4)*	57.1%*	123.5 (16.5)	4.2%	53.6 (15.5)*	198.9 (37.1)	12.5%*	5.7%

HDL indicates high-density lipoprotein; NHW, non-Hispanic white; SBP, systolic blood pressure.

*Difference between NHW and each subgroup is statistically significant at P<0.001 for all variables in BOLD. Statistical testing was based on t test for continuous variables and Fisher exact test for dichotomous variables.

†On antihypertensive medication when blood pressure was measured (if index blood pressure measure date falls into [Prescription start date −7 days, prescription end date +30 days]).

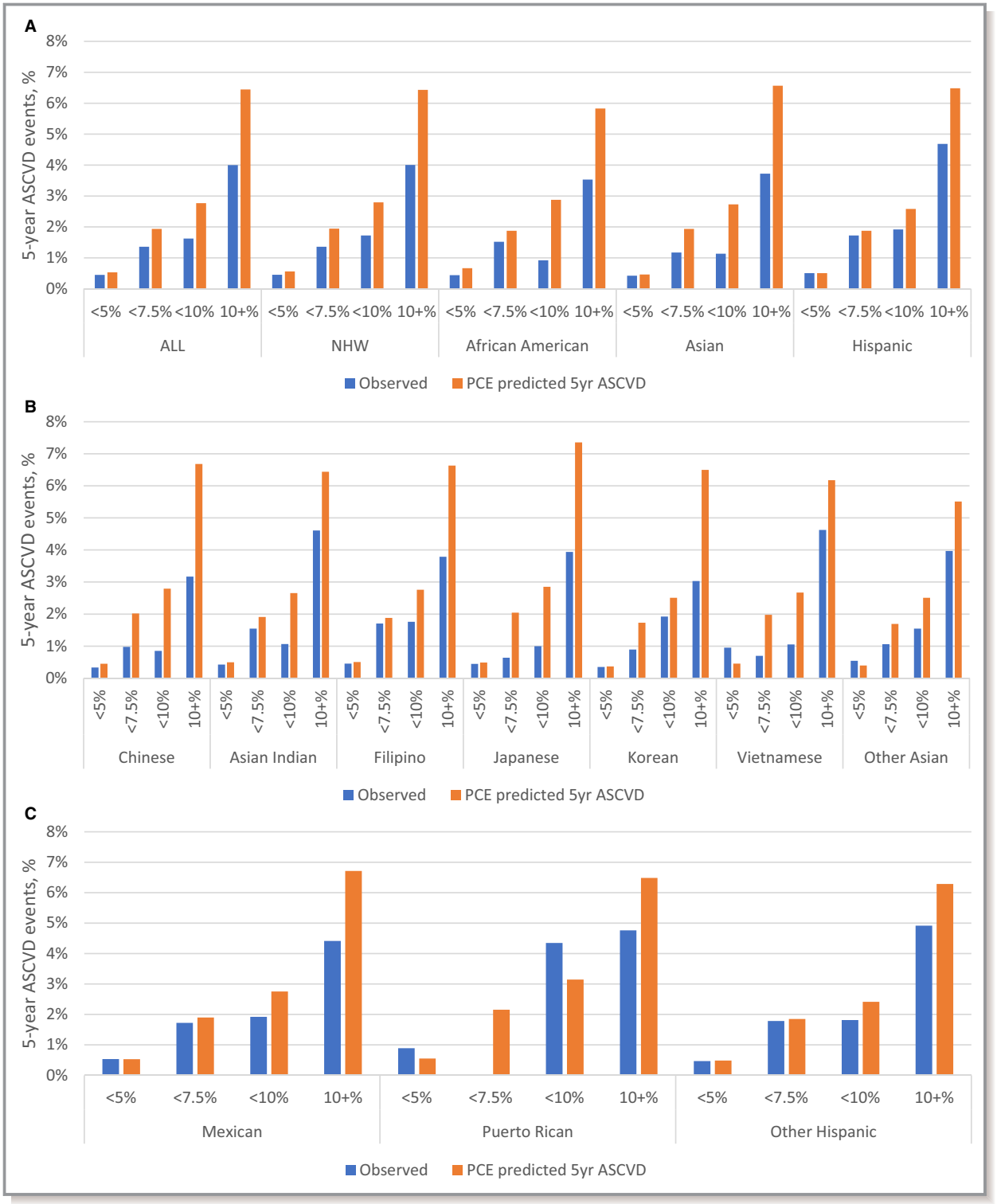

Figure 2. Comparison of 5-year observed vs predicted atherosclerotic cardiovascular risk by race/ethnicity and 10-year PCE risk categories. **A**, Overall population by major race/ethnic subgroups, (**B**) Asian subgroups, (**C**) Hispanic subgroups. ASCVD indicates atherosclerotic cardiovascular disease; PCE, Pooled Cohort Equation.

Discussion

Using EHR data from a large healthcare system in northern California, we found that the PCE generally overestimated ASCVD risk in a diverse cohort of patients. However, we document significant heterogeneity in the extent of risk overestimation among disaggregated Asian and Hispanic

Table 2. Pooled Cohort Equations Predicted vs Observed Rates of Atherosclerotic Cardiovascular Disease by Race/Ethnic Subgroup

	n	PCE-Predicted ASCVD	Observed ASCVD	Predicted-to-Observed Ratio	*P* Value*
All	231 622	1.70%	1.20%	1.5	<0.001
NHW	151 615	1.90%	1.20%	1.5	<0.001
African American	4117	2.10%	1.30%	1.6	<0.001
Asian	56 130	1.40%	0.90%	1.5	<0.001
Hispanic	19 760	1.60%	1.30%	1.2	<0.001
Chinese	23 171	1.40%	0.70%	1.9	<0.001
Asian Indian	13 815	1.20%	0.90%	1.3	<0.001
Filipino	6220	1.70%	1.20%	1.5	<0.001
Japanese	3825	1.90%	1.10%	1.8	<0.001
Korean	1793	1.20%	0.70%	1.6	0.024
Vietnamese	2093	1.10%	1.20%	0.9	0.655
Other Asian	5213	1.20%	1.00%	1.1	0.298
Mexican	8655	1.70%	1.30%	1.3	<0.001
Puerto Rican	457	1.60%	1.50%	1.1	0.887
Other Hispanic	10 643	1.60%	1.30%	1.2	0.024

ASCVD indicates atherosclerotic cardiovascular disease; NHW, non-Hispanic white; PCE, Pooled Cohort Equations.
*Difference between observed and PCE-predicted ASCVD risk of varying follow-up by subgroup based on paired *t* test.

subgroups, populations who were not included in the original PCE derivation cohorts. The PCE was better calibrated for Korean and Puerto Rican subgroups. Discrimination also differed by subgroups and recalibration failed to significantly improve the PCE performance across disaggregated subgroups.

ASCVD remains the leading cause of death for men and women in the United States and, importantly, ASCVD is largely preventable through lifestyle modifications and adequate risk factor management. Effective prevention requires an assessment of risk for selection of appropriate interventions. The current risk equations were developed to predict 10-year risk of a first ASCVD event, and were derived from 5 large community-based prospective cohorts of US NHW and African American populations. These risk calculators are widely used in routine to clinical practice to make decisions about statins, aspirin, and antihypertensives.[2,25,26] Several large studies have since been conducted to validate the PCE in real-world populations, with conflicting conclusions and resulting controversy.[18,20,27–30] A cohort of Kaiser Permanente northern California members in 2008 to 2013 concluded that the PCE substantially overestimated 5-year risk across diverse sociodemographic subgroups.[28] In a recent study of the Women's Health Initiative cohort of postmenopausal women aged 50 to 79 years, the PCE overestimated ASCVD risk across racial groups, although risks were better aligned with PCE predictions after inclusion of further Medicare-adjudicated ASCVD events.[27]

With these conflicting results, a recent study updated the data used to derive the PCE with newer cohorts and applied revised statistical models, which reduced risk overestimation overall and implausible risk estimates in African Americans in particular, but did not provide revised PCEs for racial groups other than African American or NHW adults.[31] Another study used a large cohort of primary care patients in New Zealand to recalibrate the PCE to reflect the low risk setting of a high-income country population, but similarly did not report revised equations for disaggregated race subgroups.[32] Our study is the first to specifically look at disaggregated Asian and Hispanic subgroups using a contemporary clinical cohort from EHR data.

Our findings have important real-world implications. Unlike most validation studies, which have relied on data from longitudinal cohort studies, our study population is strengthened by a diverse observational EHR-based patient sample of disaggregated Asians and Hispanics. This is significant because participants who elect to enroll in longitudinal cohort studies may be healthier than those enrolled in regular clinical care.[33] We found that data collected in routine clinical practice that are captured in the EHR can be used for ASCVD risk prediction, with similar performance to other studies.[20,27,29,34] Although Asian and Hispanic participants were not considered in the original PCE derivation cohorts, we found that, when studied in aggregate, the overall performance of the equations for these diverse patients was similar to that of NHWs.

Table 3. Discrimination of PCE and Recalibrated PCE by Race/Ethnic Subgroup

Race/Ethnic Subgroup	n	PCE	Recalibrated PCE*
		C-Index (95% CI)	C-Index (95% CI)
NHW	151 615	0.77 (0.76–0.78)	0.78 (0.77–0.79)
African American	4117	0.74 (0.66–0.81)	0.70 (0.64–0.77)
Asian	56 130	0.78 (0.75–0.80)	0.78 (0.76–0.80)
Hispanic	19 760	0.78 (0.76–0.81)	0.78 (0.75–0.81)
Chinese	23 171	0.78 (0.74–0.82)	0.79 (0.75–0.82)
Indian	13 815	0.78 (0.74–0.82)	0.78 (0.74–0.82)
Filipino	6220	0.77 (0.73–0.82)	0.78 (0.73–0.83)
Japanese	3825	0.79 (0.71–0.85)	0.77 (0.70–0.84)
Korean	1793	0.83 (0.74–0.91)	0.84 (0.75–0.92)
Vietnamese	2093	0.67 (0.55–0.77)	0.68 (0.56–0.78)
Other Asian	5213	0.77 (0.70–0.83)	0.76 (0.69–0.82)
Mexican	8655	0.78 (0.73–0.82)	0.77 (0.73–0.81)
Puerto Rican	457	0.66 (0.37–0.87)	0.63 (0.31–0.86)
Other Hispanic	10 643	0.79 (0.75–0.83)	0.78 (0.74–0.82)

PCE indicates Pooled Cohort Equations; NHW, non-Hispanic white.
*Differences in C-indices of PCE and recalibrated PCE in 1000 bootstrapped samples were not different from zero (ie, $|Z|<1.96$ or $P>0.05$) for all subgroups except for NHW ($Z=4.5$; $P<0.01$).

Importantly, we found that the PCE showed differences in calibration and discrimination by disaggregated race/ethnic subgroups. Disaggregated Asian and Hispanic subgroups differ significantly by cardiovascular risk and outcomes.[8,9,35,36] In fact, the revised cholesterol guidelines note this marked heterogeneity by race/ethnic groups and now identify ethnicity as a potential "ASCVD risk enhancer."[17] In our study, the PCE was better calibrated for Korean and Puerto Rican patients and had generally adequate discrimination across race/ethnic subgroups. As seen by differences in baseline characteristics, it is likely that differences in clinical disease phenotypes (eg, higher rates of diabetes mellitus in some subgroups) may impact the performance of the PCE among these subgroups. Although some Hispanic subgroups have African ancestry, using the parameters for African Americans did not improve the PCE's performance among Hispanics. Notably, recalibration of the PCE failed to improve discrimination across all groups. This suggests that novel risk prediction models should consider updated risk factors and relationships that are not currently present in the PCE.

Our results should be interpreted in light of several limitations. Because of low event numbers, some disaggregated racial/ethnic subgroups had to be excluded from the analysis (Cubans) and were combined with the "other" category. Importantly, the "other" Hispanic and Asian subgroups represent heterogeneous populations and our recalibrated equations thus should not be generalized to these subgroups. The study sample was comprised of a diverse population in northern California and may not reflect the general US population. Some racial groups may have been under- or overrepresented in our study population (ie, Mexicans among Hispanics and Chinese Americans among Asian subgroups) and the small numbers of patients in some subgroups may have affected our models' performance. Our study population is likely to be insured and of higher socioeconomic status and therefore our results cannot be extrapolated to all settings. Our data originate from the EHR, which has inherent limitations including that the data are collected for clinical purposes and not for research.[37] Race/ethnicity was self-reported in the majority of the study sample and inferred in the remainder based on previously validated methods using name lists to infer race/ethnic subgroup.[22] Such approaches may have led to misclassification of race/ethnicity. Incident ASCVD events that occurred outside of the health system may not be fully captured. This underreporting of events may have led to underestimation of observed event rates, although our results are well-aligned with other studies.[20,28,34] We attempted to minimize the potential missing information by requiring that patients have at least 2 outpatient visits and stayed in the healthcare system for at least 1 year and by matching EHR records to Social Security Records to ensure that deaths were appropriately captured.

Future risk prediction models should not only include Asians and Hispanics, the 2 largest and growing minority groups in the United States, but also disaggregate by major subgroups. Our study suggests that information routinely captured in clinical care may be useful in risk prediction. Using novel algorithms and continuing to refine the demographic and clinical variables considered will likely improve future risk prediction models for diverse populations.

In summary, we used EHR data from a health system in northern California to validate the PCE for estimating ASCVD risk among disaggregated Asian and Hispanic patients. We found that the PCE generally overestimated predicted risk among this diverse patient cohort, but provided adequate discrimination. These findings have important implications for the use of the EHR-based variables in ASCVD risk prediction across diverse populations.

Sources of Funding

Dr. Rodriguez received support from the McCormick Faculty Fellowship from Stanford University School of Medicine and the National Heart, Lung, and Blood Institute, National

Institutes of Health (1K01HL144607). Dr. Blum is supported by a grant from the Swiss National Science Foundation (P2BEP3_175289). Dr. Chung was supported by the National Heart, Lung, and Blood Institute, National Institutes of Health (R01HL126172). Dr. Palaniappan was supported by the National Institute on Minority Health and Health Disparities (R01 MD007012). Dr. Coulet is supported by The Widen Horizons program of the IDEX "Lorraine Université d'Excellence" (15-IDEX-0004) and the Pilot Grant Program of the Stanford Center for Clinical and Translational Research and Education.

Disclosures

None.

References

1. US Census Quick Facts. Available at: https://www.census.gov/quickfacts/fact/table/US/PST045217. Accessed August 24, 2018.
2. Goff DC Jr, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB Sr, Gibbons R, Greenland P, Lackland DT, Levy D, O'Donnell CJ, Robinson JG, Schwartz JS, Shero ST, Smith SC Jr, Sorlie P, Stone NJ, Wilson PW; American College of Cardiology/American Heart Association Task Force on Practice G. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014;129:S49–S73.
3. Mainous AG III, Majeed A, Koopman RJ, Baker R, Everett CJ, Tilley BC, Diaz VA. Acculturation and diabetes among Hispanics: evidence from the 1999–2002 National Health and Nutrition Examination Survey. *Public Health Rep*. 2006;121:60–66.
4. Davidson JA, Kannel WB, Lopez-Candales A, Morales L, Moreno PR, Ovalle F, Rodriguez CJ, Rodbard HW, Rosenson RS, Stern M. Avoiding the looming Latino/Hispanic cardiovascular health crisis: a call to action. *Ethn Dis*. 2007;17:568–573.
5. Lara M, Gamboa C, Kahramanian MI, Morales LS, Bautista DE. Acculturation and Latino health in the United States: a review of the literature and its sociopolitical context. *Annu Rev Public Health*. 2005;26:367–397.
6. Derby CA, Wildman RP, McGinn AP, Green RR, Polotsky AJ, Ram KT, Barnhart J, Weiss G, Santoro N. Cardiovascular risk factor variation within a Hispanic cohort: SWAN, the Study of Women's Health Across the Nation. *Ethn Dis*. 2010;20:396–402.
7. Palaniappan LP, Araneta MR, Assimes TL, Barrett-Connor EL, Carnethon MR, Criqui MH, Fung GL, Narayan KM, Patel H, Taylor-Piliae RE, Wilson PW, Wong ND. Call to action: cardiovascular disease in Asian Americans: a science advisory from the American Heart Association. *Circulation*. 2010;122:1242–1252.
8. Rodriguez F, Hastings KG, Boothroyd DB, Echeverria S, Lopez L, Cullen M, Harrington RA, Palaniappan LP. Disaggregation of cause-specific cardiovascular disease mortality among Hispanic subgroups. *JAMA Cardiol*. 2017;2:240–247.
9. Frank AT, Zhao B, Jose PO, Azar KM, Fortmann SP, Palaniappan LP. Racial/ethnic differences in dyslipidemia patterns. *Circulation*. 2014;129:570–579.
10. Hastings KG, Eggleston K, Boothroyd D, Kapphahn KI, Cullen MR, Barry M, Palaniappan LP. Mortality outcomes for Chinese and Japanese immigrants in the USA and countries of origin (Hong Kong, Japan): a comparative analysis using national mortality records from 2003 to 2011. *BMJ Open*. 2016;6:e012201.
11. Hastings KG, Jose PO, Kapphahn KI, Frank AT, Goldstein BA, Thompson CA, Eggleston K, Cullen MR, Palaniappan LP. Leading causes of death among Asian American subgroups (2003–2011). *PLoS One*. 2015;10:e0124341.
12. Nguyen MT, Hu J, Hastings KG, Daza EJ, Cullen MR, Orloff LA, Palaniappan LP. Thyroid cancer mortality is higher in Filipinos in the United States: an analysis using national mortality records from 2003 through 2012. *Cancer*. 2017;123:4860–4867.
13. Pu J, Romanelli R, Zhao B, Azar KM, Hastings KG, Nimbal V, Fortmann SP, Palaniappan LP. Dyslipidemia in special ethnic populations. *Endocrinol Metab Clin North Am*. 2016;45:205–216.
14. Palaniappan L, Wang Y, Fortmann SP. Coronary heart disease mortality for six ethnic groups in California, 1990–2000. *Ann Epidemiol*. 2004;14:499–506.
15. Wang EJ, Wong EC, Dixit AA, Fortmann SP, Linde RB, Palaniappan LP. Type 2 diabetes: identifying high risk Asian American subgroups in a clinical population. *Diabetes Res Clin Pract*. 2011;93:248–254.
16. Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, Goldberg AC, Gordon D, Levy D, Lloyd-Jones DM, McBride P, Schwartz JS, Shero ST, Smith SC Jr, Watson K, Wilson PW. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014;129:S1–S45.
17. Grundy SM, Stone NJ, Bailey AL, Beam C, Birtcher KK, Blumenthal RS, Braun LT, de Ferranti S, Faiella-Tommasino J, Forman DE, Goldberg R, Heidenreich PA, Hlatky MA, Jones DW, Lloyd-Jones D, Lopez-Pajares N, Ndumele CE, Orringer CE, Peralta CA, Saseen JJ, Smith SC Jr, Sperling L, Virani SS, Yeboah J. 2018 (AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA guideline on the management of blood cholesterol: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation*. 2019;139:e1082–e1143. [epub ahed of print]
18. DeFilippis AP, Young R, Carrubba CJ, McEvoy JW, Budoff MJ, Blumenthal RS, Kronmal RA, McClelland RL, Nasir K, Blaha MJ. An analysis of calibration and discrimination among multiple cardiovascular risk scores in a modern multiethnic cohort. *Ann Intern Med*. 2015;162:266–275.
19. Karmali KN, Goff DC Jr, Ning H, Lloyd-Jones DM. A systematic examination of the 2013 ACC/AHA pooled cohort risk assessment tool for atherosclerotic cardiovascular disease. *J Am Coll Cardiol*. 2014;64:959–968.
20. Muntner P, Colantonio LD, Cushman M, Goff DC Jr, Howard G, Howard VJ, Kissela B, Levitan EB, Lloyd-Jones DM, Safford MM. Validation of the atherosclerotic cardiovascular disease Pooled Cohort risk equations. *JAMA*. 2014;311:1406–1415.
21. Lee CH, Woo YC, Lam JK, Fong CH, Cheung BM, Lam KS, Tan KC. Validation of the Pooled Cohort equations in a long-term cohort study of Hong Kong Chinese. *J Clin Lipidol*. 2015;9:640–646.e2.
22. Wong EC, Palaniappan LP, Lauderdale DS. Using name lists to infer Asian racial/ethnic subgroups in the healthcare setting. *Med Care*. 2010;48:540–546.
23. Yeh RW, Sidney S, Chandra M, Sorel M, Selby JV, Go AS. Population trends in the incidence and outcomes of acute myocardial infarction. *N Engl J Med*. 2010;362:2155–2165.
24. Go AS, Hylek EM, Chang Y, Phillips KA, Henault LE, Capra AM, Jensvold NG, Selby JV, Singer DE. Anticoagulation therapy for stroke prevention in atrial fibrillation: how well do randomized trials translate into clinical practice? *JAMA*. 2003;290:2685–2692.
25. Bibbins-Domingo K; Force USPST. Aspirin use for the primary prevention of cardiovascular disease and colorectal cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2016;164:836–845.
26. Whelton PK, Carey RM, Aronow WS, Casey DE Jr, Collins KJ, Dennison Himmelfarb C, DePalma SM, Gidding S, Jamerson KA, Jones DW, MacLaughlin EJ, Muntner P, Ovbiagele B, Smith SC Jr, Spencer CC, Stafford RS, Taler SJ, Thomas RJ, Williams KA Sr, Williamson JD, Wright JT Jr. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation*. 2018;138:e426–e483.
27. Mora S, Wenger NK, Cook NR, Liu J, Howard BV, Limacher MC, Liu S, Margolis KL, Martin LW, Paynter NP, Ridker PM, Robinson JG, Rossouw JE, Safford MM, Manson JE. Evaluation of the Pooled Cohort risk equations for cardiovascular risk prediction in a multiethnic cohort from the Women's Health Initiative. *JAMA Intern Med*. 2018;178:1231–1240.
28. Rana JS, Tabada GH, Solomon MD, Lo JC, Jaffe MG, Sung SH, Ballantyne CM, Go AS. Accuracy of the atherosclerotic cardiovascular risk equation in a large contemporary, multiethnic population. *J Am Coll Cardiol*. 2016;67:2118–2130.
29. DeFilippis AP, Young R, McEvoy JW, Michos ED, Sandfort V, Kronmal RA, McClelland RL, Blaha MJ. Risk score overestimation: the impact of individual cardiovascular risk factors and preventive therapies on the performance of the American Heart Association-American College of Cardiology-Atherosclerotic Cardiovascular Disease risk score in a modern multi-ethnic cohort. *Eur Heart J*. 2017;38:598–608.
30. Cook NR, Ridker PM. Calibration of the Pooled Cohort Equations for atherosclerotic cardiovascular disease: an update. *Ann Intern Med*. 2016;165:786–794.

31. Yadlowsky S, Hayward RA, Sussman JB, McClelland RL, Min YI, Basu S. Clinical implications of revised Pooled Cohort Equations for estimating atherosclerotic cardiovascular disease risk. *Ann Intern Med*. 2018;169:20–29.
32. Pylypchuk R, Wells S, Kerr A, Poppe K, Riddell T, Harwood M, Exeter D, Mehta S, Grey C, Wu BP, Metcalf P, Warren J, Harrison J, Marshall R, Jackson R. Cardiovascular disease risk prediction equations in 400 000 primary care patients in New Zealand: a derivation and validation study. *Lancet*. 2018;391:1897–1907.
33. Galea S, Tracy M. Participation rates in epidemiologic studies. *Ann Epidemiol*. 2007;17:643–653.
34. Wolfson J, Vock DM, Bandyopadhyay S, Kottke T, Vazquez-Benitez G, Johnson P, Adomavicius G, O'Connor PJ. Use and customization of risk scores for predicting cardiovascular events using electronic health record data. *J Am Heart Assoc*. 2017;6:e003670. DOI: 10.1161/JAHA.116.003670.
35. Jose PO, Frank AT, Kapphahn KI, Goldstein BA, Eggleston K, Hastings KG, Cullen MR, Palaniappan LP. Cardiovascular disease mortality in Asian Americans. *J Am Coll Cardiol*. 2014;64:2486–2494.
36. Daviglus ML, Talavera GA, Aviles-Santa ML, Allison M, Cai J, Criqui MH, Gellman M, Giachello AL, Gouskova N, Kaplan RC, LaVange L, Penedo F, Perreira K, Pirzada A, Schneiderman N, Wassertheil-Smoller S, Sorlie PD, Stamler J. Prevalence of major cardiovascular risk factors and cardiovascular diseases among Hispanic/Latino individuals of diverse backgrounds in the United States. *JAMA*. 2012;308:1775–1784.
37. Hersh WR, Weiner MG, Embi PJ, Logan JR, Payne PR, Bernstam EV, Lehmann HP, Hripcsak G, Hartzog TH, Cimino JJ, Saltz JH. Caveats for the use of operational electronic health record data in comparative effectiveness research. *Med Care*. 2013;51:S30–S37.

SUPPLEMENTAL MATERIAL

Table S1. ASCVD Incidence by race/ethnicity*.

	N	Total ASCVD Events	Total ASCVD	MI	Ischemic stroke	Hemorrhagic stroke	Fatal CAD
All	231,622	2,703	1.20%	0.39%	0.54%	0.18%	0.06%
NHW	151,615	1,894	1.20%	0.41%	0.58%	0.18%	0.08%
African American	4,117	53	1.30%	0.41%	0.80%	0.05%	0.02%
Asian	56,130	494	0.90%	0.30%	0.37%	0.18%	0.02%
Hispanic	19,760	262	1.30%	0.43%	0.61%	0.26%	0.04%
Chinese	23,171	166	0.70%	0.21%	0.34%	0.17%	0.00%
Indian	13,815	123	0.90%	0.41%	0.29%	0.17%	0.02%
Filipino	6,220	72	1.20%	0.50%	0.43%	0.13%	0.10%
Japanese	3,825	41	1.10%	0.24%	0.58%	0.24%	0.03%
Korean	1793	13	0.70%	0.17%	0.33%	0.22%	0.00%
Vietnamese	2,093	26	1.20%	0.43%	0.57%	0.19%	0.05%
Other Asian	5,213	53	1.00%	0.29%	0.46%	0.25%	0.02%
Mexican	8,655	115	1.30%	0.40%	0.70%	0.22%	0.00%
Puerto Rican	457	7	1.50%	0.44%	0.66%	0.44%	0.00%
Other Hispanic	10,643	140	1.30%	0.44%	0.53%	0.28%	0.07%

*During study period with average follow-up time of 3.9 years
ASCVD = atherosclerotic cardiovascular disease; MI = myocardial infarction; CAD = coronary artery disease
The table shows the incidence of ASCVD events during follow-up. Only the first occurrence of an ASCVD event was counted.

Table S2. Observed and PCE Predicted ASCVD Risk by PCE-predicted 10-year Risk Category and Race/ethnicity.

	PCE 10yr risk category	Observed	PCE predicted 5yr ASCVD	Ch-sq	P-Val
ALL	<5%	0.4%	0.5%		
	<7.5%	1.4%	1.9%		
	<10%	1.6%	2.8%		
	10+%	4.0%	6.4%	490.5	<0.001
NHW	<5%	0.5%	0.6%		
	<7.5%	1.4%	1.9%		
	<10%	1.7%	2.8%		
	10+%	4.0%	6.4%	351.4	<0.001
African American	<5%	0.4%	0.7%		
	<7.5%	1.5%	1.9%		
	<10%	0.9%	2.9%		
	10+%	3.5%	5.8%	14.9	<0.001
Asian	<5%	0.4%	0.5%		
	<7.5%	1.2%	1.9%		
	<10%	1.1%	2.7%		
	10+%	3.7%	6.6%	120.0	<0.001
Hispanic	<5%	0.5%	0.5%		
	<7.5%	1.7%	1.9%		
	<10%	1.9%	2.6%		
	10+%	4.7%	6.5%	18.0	<0.001
Chinese	<5%	0.3%	0.5%		
	<7.5%	1.0%	2.0%		
	<10%	0.9%	2.8%		
	10+%	3.2%	6.7%	78.6	<0.001
Asian Indian	<5%	0.4%	0.5%		
	<7.5%	1.5%	1.9%		
	<10%	1.1%	2.7%		
	10+%	4.6%	6.4%	13.1	<0.001
Filipino	<5%	0.5%	0.5%		
	<7.5%	1.7%	1.9%		

	<10%	1.8%	2.8%		
	10+%	3.8%	6.6%	14.3	<0.001
Japanese	<5%	0.4%	0.5%		
	<7.5%	0.6%	2.0%		
	<10%	1.0%	2.9%		
	10+%	3.9%	7.4%	16.5	<0.001
Korean	<5%	0.3%	0.4%		
	<7.5%	0.9%	1.7%		
	<10%	1.9%	2.5%		
	10+%	3.0%	6.5%	4.5	0.107
Vietnamese	<5%	1.0%	0.5%		
	<7.5%	0.7%	2.0%		
	<10%	1.1%	2.7%		
	10+%	4.6%	6.2%	12.0	0.002
Other Asian	<5%	0.5%	0.4%		
	<7.5%	1.1%	1.7%		
	<10%	1.5%	2.5%		
	10+%	4.0%	5.5%	6.7	0.036
Mexican	<5%	0.5%	0.5%		
	<7.5%	1.7%	1.9%		
	<10%	1.9%	2.8%		
	10+%	4.4%	6.7%	12.8	0.002
Puerto Rican	<5%	0.9%	0.5%		
	<7.5%	0.0%	2.2%		
	<10%	4.3%	3.1%		
	10+%	4.8%	6.5%	1.9	0.4
Other Hispanic	<5%	0.5%	0.5%		
	<7.5%	1.8%	1.8%		
	<10%	1.8%	2.4%		
	10+%	4.9%	6.3%	5.9	0.1

Table S3. Pooled Cohort Equation Discrimination using African American Parameters among Hispanics.

Race/Ethnic Subgroup	PCE	Recalibrated PCE	p-value for comparison
	C-index (95% CI)	C-index (95% CI)	
All Hispanics	0.75 (0.71, 0.78)	0.78 (0.75, 0.80)	0.043
Mexican	0.74 (0.69, 0.80)	0.78 (0.73, 0.82)	0.095
Puerto Rican	0.67 (0.41, 0.88)	0.67 (0.43, 0.89)	0.960
Other Hispanic	0.76 (0.71, 0.80)	0.78 (0.74, 0.82)	0.160