



2020: Appendix to White Paper 23-21

A Generalized Method for the Creation and Evaluation of Polygenic Scores: Details for Each Report

This appendix contains key information for models released in 2020 developed under the methods outlined in [White Paper 23-21](#). See the appendix for [2021 here](#), [2022 here](#), and [2023 here](#).

2020 Contents:

1. [LDL Cholesterol](#)
2. [Atrial Fibrillation](#)
3. [Coronary Artery Disease](#)
4. [High Blood Pressure](#)
5. [Migraine](#)
6. [Uterine Fibroids](#)
7. [Obstructive Sleep Apnea](#)
8. [Restless Legs Syndrome](#)
9. [Gout](#)
10. [Nonalcoholic Fatty Liver Disease](#)
11. [References](#)

1. LDL Cholesterol

James Ashenhurst, PhD, lead analyst

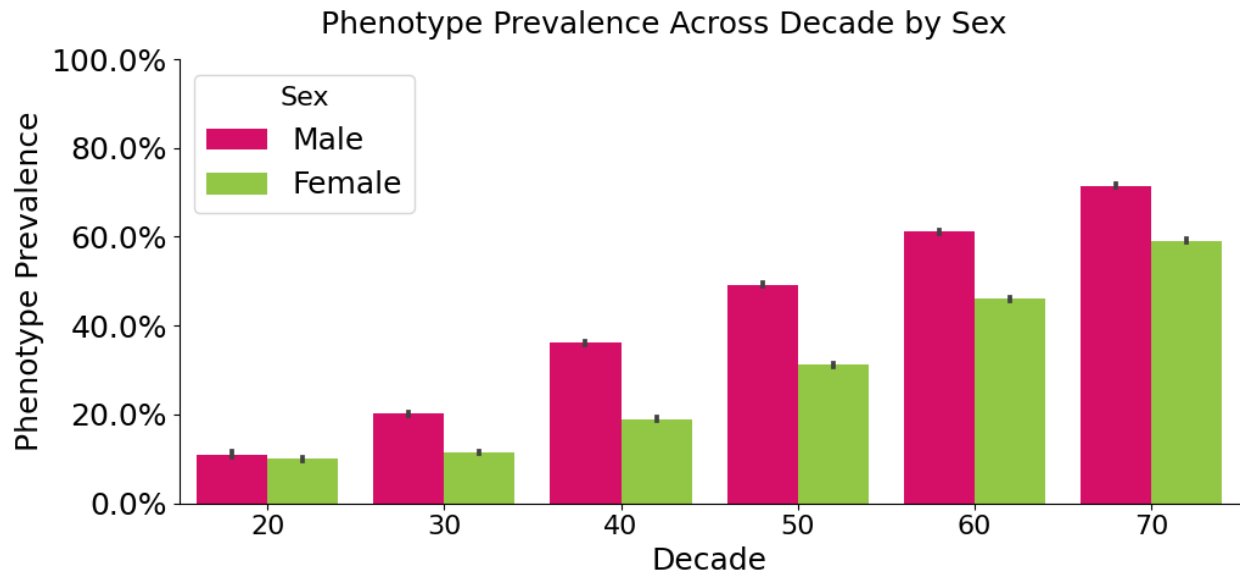
Report added May, 2020

Phenotype definition

Cases and controls were defined in two stages of logic. In the first stage, questions about recent and highest ever LDL cholesterol (LDL-C) levels were combined into a single phenotype representing ever having reported LDL-C above 160 mg/dL. Individuals who answered 160 mg/dL or above for either LDL question were counted as cases. Those who answered below 160 mg/dL for both questions were counted as controls, as were participants who answered below 160 mg/dL for one question but who did not answer the other. In the second step, prescription medication information was used to infer that a participant ever had high LDL-C. Specifically, among those with self-reported LDL-C lab values data, controls were changed to cases if they indicated a history of being prescribed medication to lower their cholesterol. This step was taken to account for the fact that, for those individuals, self-reported values may have been concurrent with medical management of LDL-C (i.e., lowered by medication), and thus individuals without high LDL-C at the time of self-report may still have had a history of high LDL-C.

Model description

The final model selected for LDL-C consisted of 2,950 genetic variants, age, sex, age*sex, age², and age²*sex interaction terms, as well as the first five principal components (PCs) derived from the European ancestry population. Genomic PCs are not used when determining an individual customer's report result. This one model is applied to all customers, with personalized binary qualitative results and estimated likelihoods based on genetic ancestry and birth sex.

Figure 1-1: High LDL-C phenotype distribution in the European ancestry training set**Table 1-1:** High LDL-C participant cohort descriptives

Platform	Ancestry Group	Sample Use	N	Age mean (SD)	Sex (% female)	High LDL-C prevalence (%)
V1 to V5	European	GWAS	617,165	56.2 (13.8)	54.60%	41.99%
V5	European	Training the European Model	511,469	55.0 (13.9)	55.50%	40.60%
		Testing	56,749	55.1 (14.0)	55.24%	40.94%
	Sub-Saharan African/African American	Testing	18,710	50.1 (13.5)	59.02%	40.94%
	East/Southeast Asian	Testing	18,357	44.7 (14.2)	57.51%	27.07%
	Hispanic/Latino	Testing	72,806	47.8 (14.0)	56.46%	33.86%
	South Asian	Testing	6,128	44.3 (13.0)	37.73%	34.48%
	Northern African/Western Asian	Testing	5,267	49.4 (14.7)	40.38%	38.47%

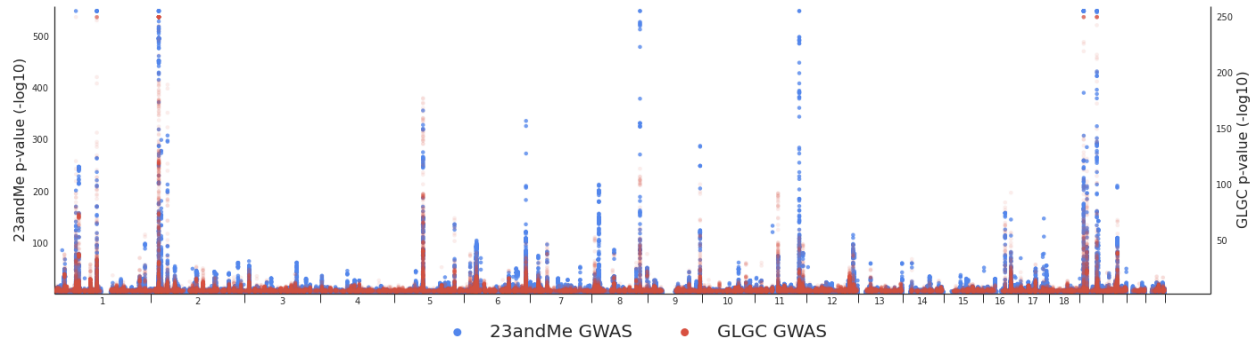


Figure 1-2: Manhattan plot of 23andMe and Willer et al., (2013) GWAS summary statistics for LDL-C. The p -values for each variant shared between the 23andMe (blue) and Global Lipids Genetics Consortium (GLGC; red) genome-wide association studies (GWASs) are depicted. Chromosomal location is represented on the horizontal axis, and the negative log of the p -value is represented on the vertical axis (scaled separately for each analysis).

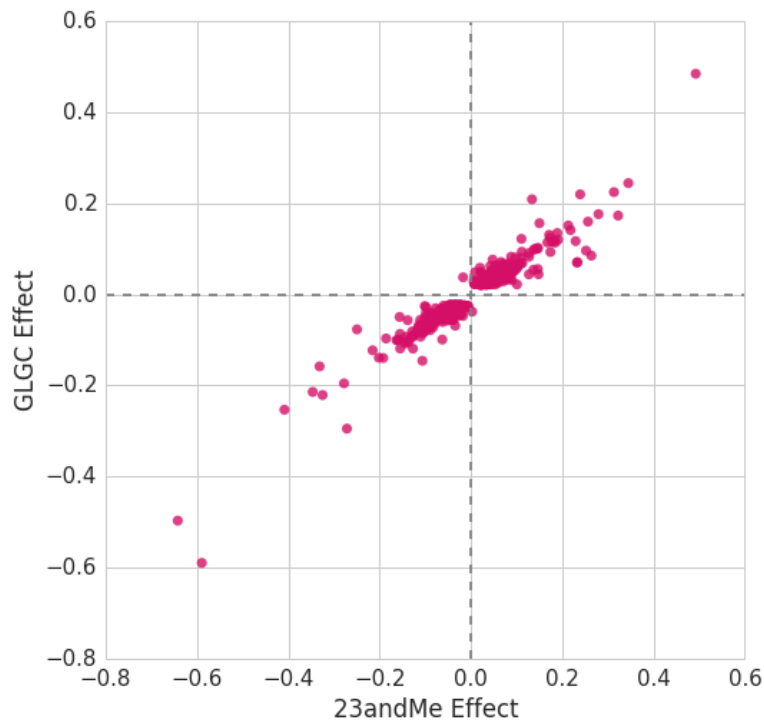


Figure 1-3: Scatter plot showing the estimated effect sizes for 23andMe (change in log-odds per unit predictor change) and Global Lipids Genetics Consortium (GLGC; linear betas; Willer et al., 2013) genome-wide significant hits shared between the two GWAS for LDL cholesterol.

Table 1-2: High LDL-C model performance characteristics

Ancestry Group (test sets)	Full Model AUROC	Genetics Only AUROC	Odds Ratio top 5% versus average (95%CIs)	Odds Ratio top 5% versus bottom 5% (95%CIs)
European	0.7770	0.6456	2.81 (2.58 to 3.07)	10.24 (9.02 to 11.63)
Sub-Saharan African/African American	0.7312	0.5985	1.91 (1.67 to 2.23)	4.10 (3.34 to 5.05)
East/Southeast Asian	0.7635	0.5888	1.91 (1.64 to 2.22)	4.30 (3.43 to 5.39)
Hispanic/Latino	0.7561	0.6179	2.31 (2.15 to 2.49)	5.87 (5.27 to 6.55)
South Asian	0.7828	0.6222	2.69 (2.08 to 3.47)	7.75 (5.29 to 11.37)
Northern African/Western Asian	0.7776	0.6188	2.81 (2.13 to 3.72)	7.49 (5.04 to 11.14)

Table 1-3: High LDL-C qualitative result characteristics

Ancestry Group (test sets)	Odds Ratio for Result Threshold	Percent Above Threshold	Likelihood Ratio of "Increased" Result (95% CIs)
European	1.5	22.79%	1.97 (1.91 to 2.03)
Sub-Saharan African/African American	1.5	12.32%	1.69 (1.56 to 1.82)
East/Southeast Asian	1.5	10.37%	1.63 (1.50 to 1.78)
Hispanic/Latino	1.5	17.19%	1.80 (1.74 to 1.86)
South Asian	1.5	18.29%	1.82 (1.64 to 2.02)
Northern African/Western Asian	1.5	17.47%	1.85 (1.64 to 2.08)

Baseline prevalences used for likelihood estimate

Ancestry- and sex-specific baseline prevalences of ever having had high LDL cholesterol were derived from the 2017 data release of the Behavioral Risk Factor Surveillance System (BRFSS; Centers for Disease Control and Prevention [CDC], 2017). The specific calculated variable (coded _RFCHOL1) represents the concept: adults who have had their cholesterol checked and have been told by a doctor, nurse, or other health professional that it was high. The ancestry variable used (coded _RACE) included the categories White non-Hispanic, Black non-Hispanic, Asian non-Hispanic, and Hispanic. Analysis was restricted only to those between the ages of 70 and 79, to capture this decade of age (coded _AGEG5YR). The baselines used for each sex and ancestry combination and how they map to each 23andMe ancestry group are shown in Table 4.

Table 1-4: High LDL-C baseline prevalences in BRFSS data between ages 70 to 79

Group	Matched 23andMe Population(s)	Sex	N	Prevalence	95% CI
White Non-Hispanic	European, Northern African/Western Asian, Other	Male	23,256	55.02%	54.38% to 55.66%
		Female	33,369	54.22%	53.69% to 54.76%
Black Non-Hispanic	Sub-Saharan African/African American	Male	1,335	52.58%	49.91% to 55.26%
		Female	2,795	53.42%	51.57% to 55.27%
Asian Non-Hispanic	East/Southeast Asian, South Asian	Male	327	46.18%	40.77% to 51.58%
		Female	386	51.30%	46.31% to 56.28%
Hispanic	Hispanic/Latino	Male	951	46.58%	43.41% to 49.75%
		Female	1,619	51.95%	49.51% to 54.38%

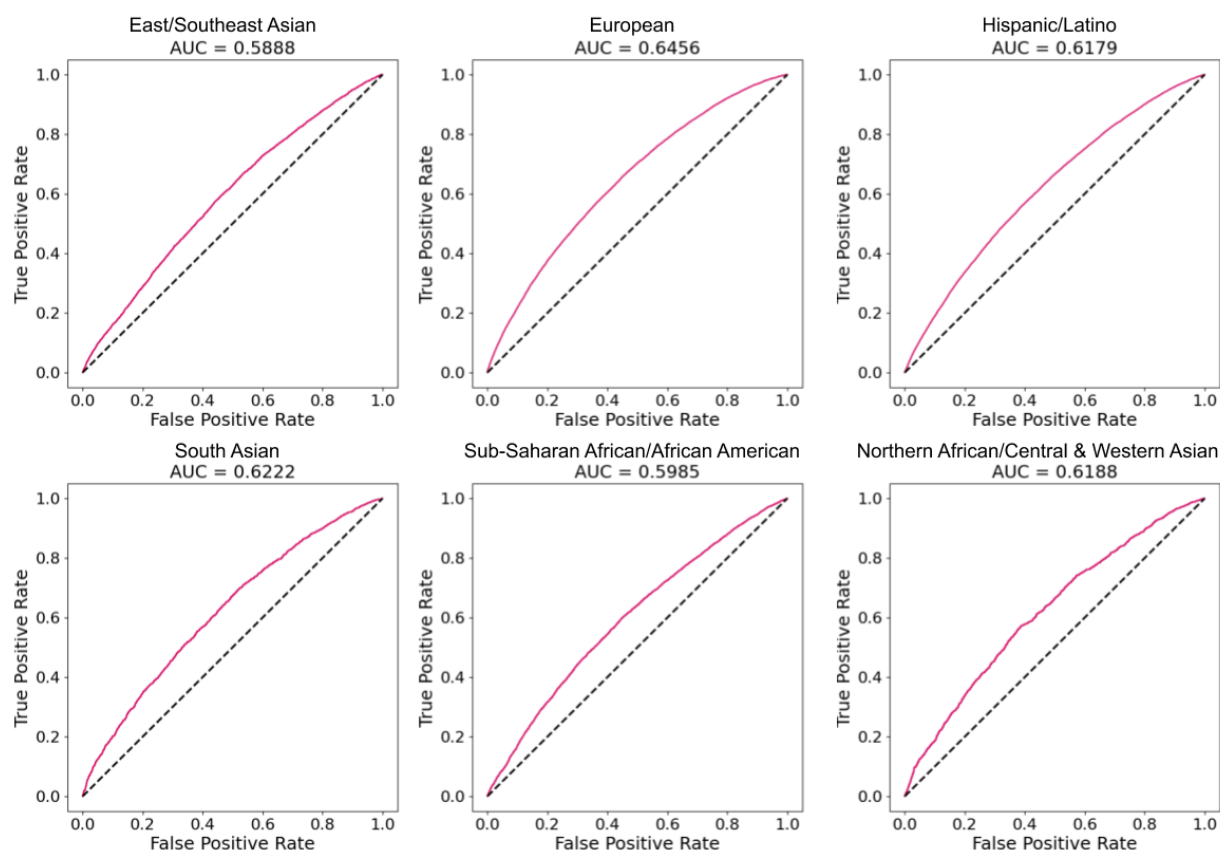
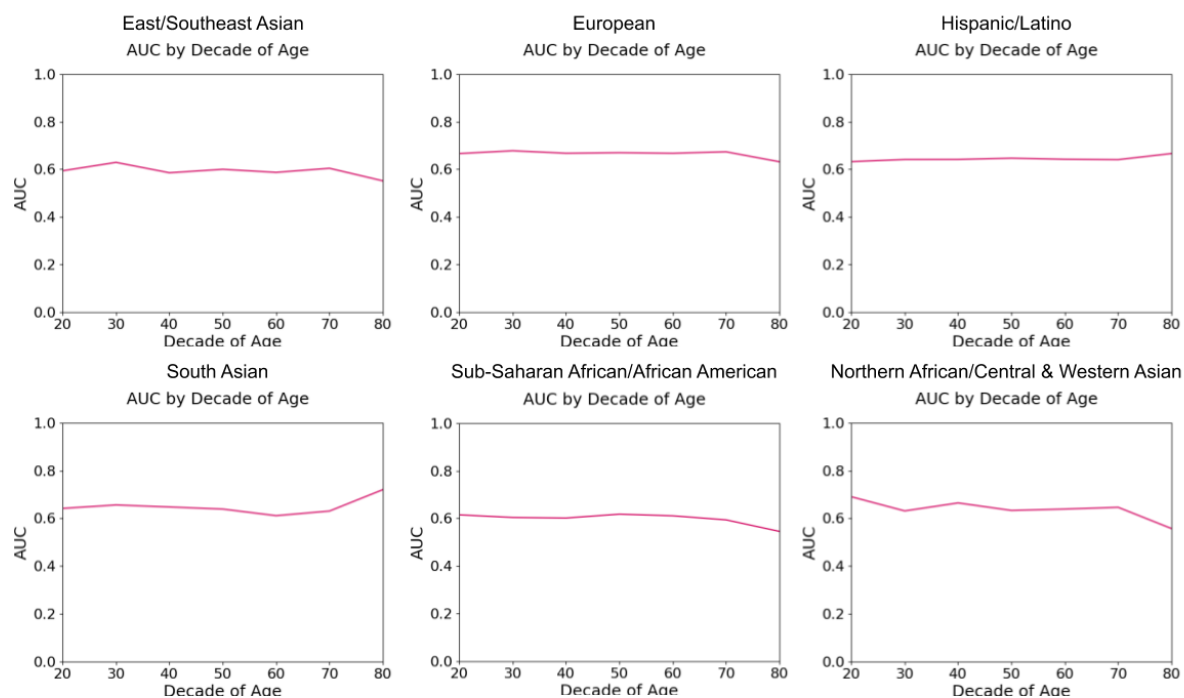
Figure 1-4: High LDL-C AUROC across ancestry-specific test sets**Figure 1-5:** High LDL-C AUROC within each decade of age across ancestry-specific test sets

Figure 1-6: High LDL-C case/control standardized PGS distributions across ancestry-specific test sets

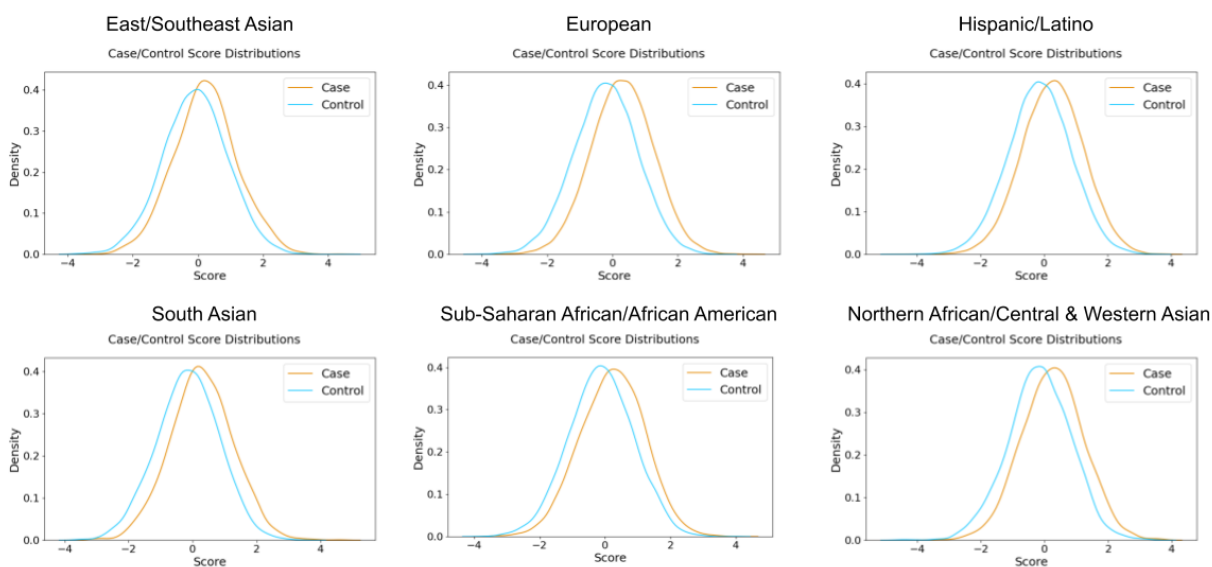
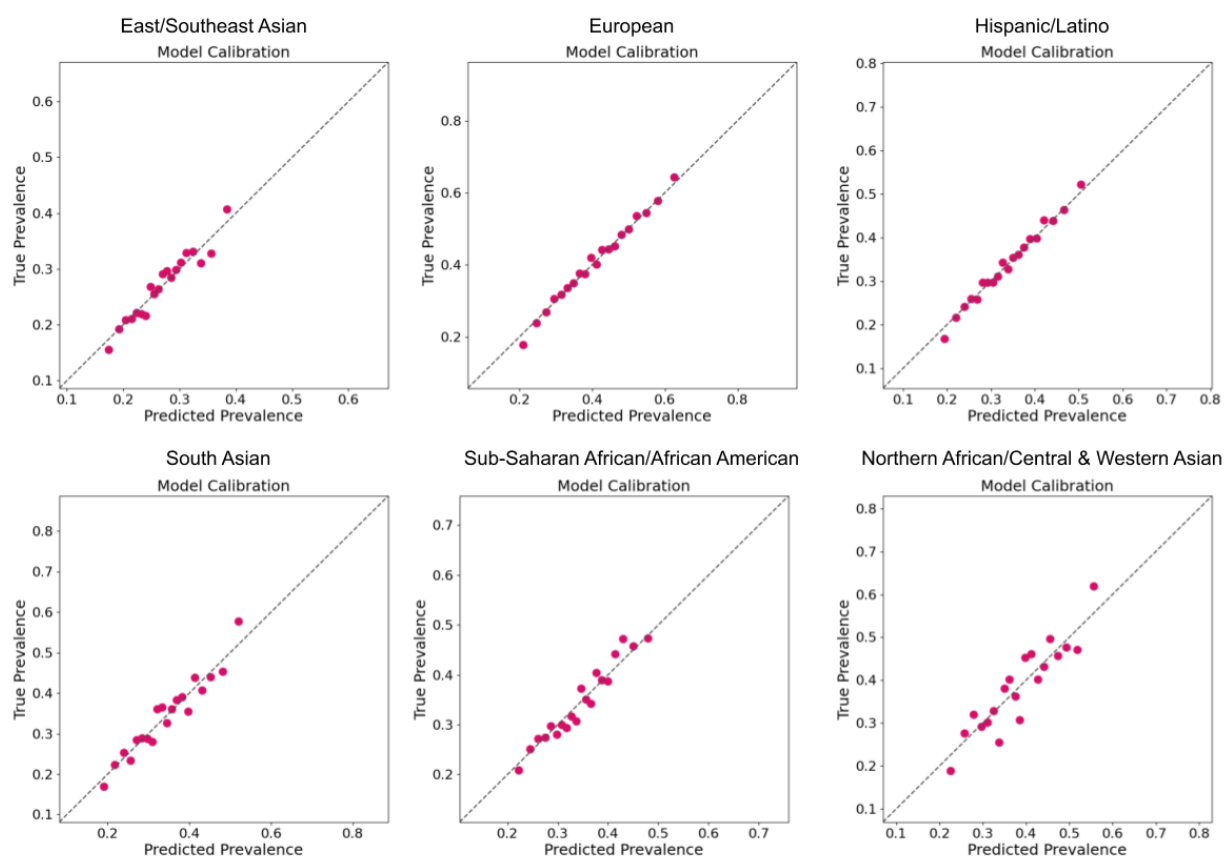


Figure 1-7: High LDL-C Platt-scaled calibration plots across ancestry-specific test sets



External Validation

As described in the generalized method White Paper, we conducted an external analysis on our LDL PGS model in the Atherosclerosis Risk in Communities (ARIC) study cohort. To determine LDL status, we combined coded phenotypes [ldlsiu02](#), [cholmdcode01](#), and [cholmdcode02](#). These phenotypes report measured LDL levels, whether the participant is taking medications to treat high cholesterol, and whether the participant is taking secondary medications that might affect cholesterol levels. Any participant who reported taking cholesterol lowering medication or had LDL > 160mg/dL was considered a case. All other participants were considered controls. Any participant taking secondary medication that might affect cholesterol levels was removed from the analysis.

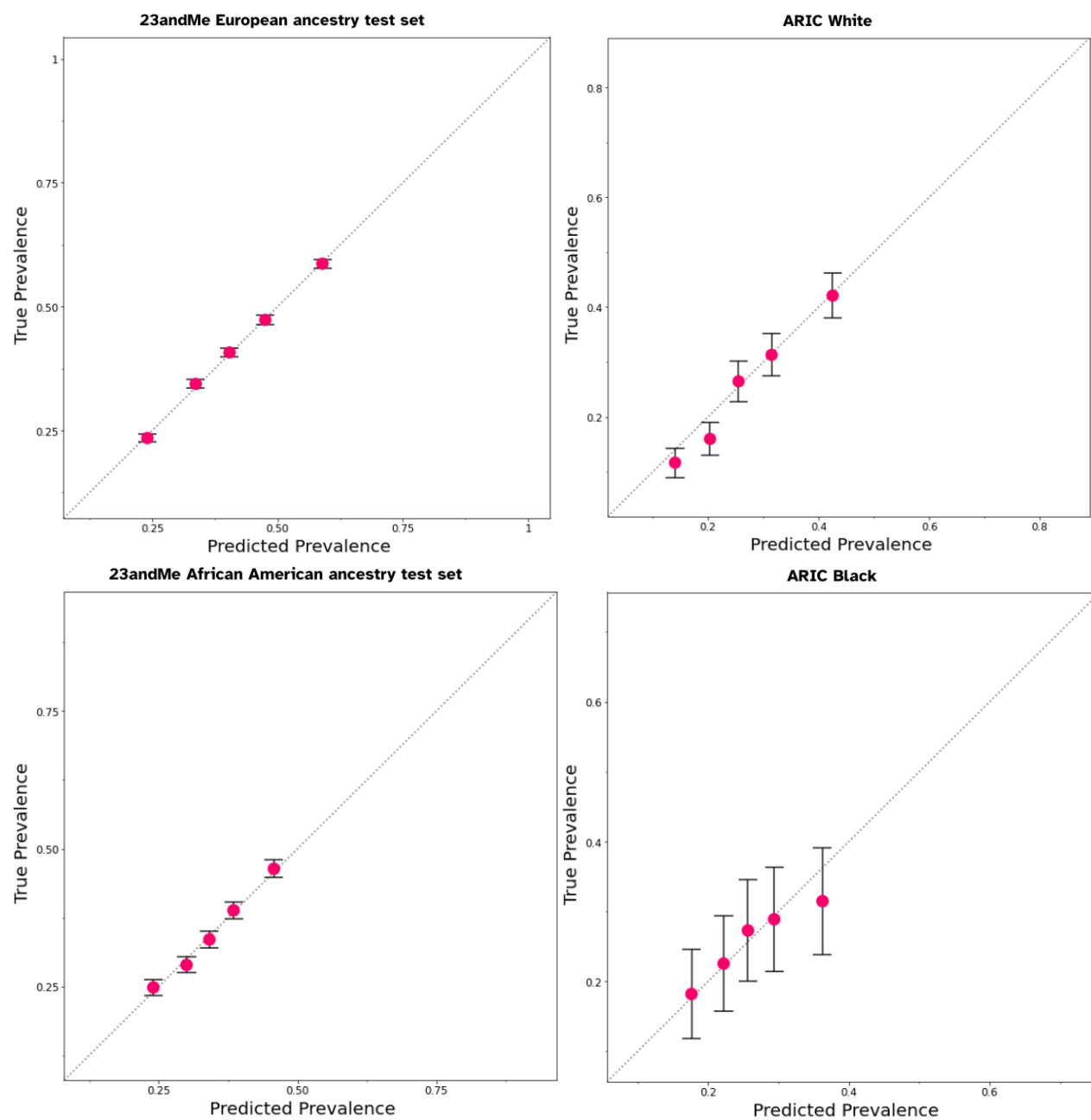
Table 1-5: ARIC cohort statistics for High LDL Cholesterol

Cohort	N	Mean max age (SD)	Sex (% female)	High LDL Prevalence (%)
ARIC white	2,779	53.8 (5.6)	51.85%	25.51%
ARIC Black	713	52.9 (5.7)	56.52%	25.67%

Table 1-6: High LDL Cholesterol PGS performance in 23andMe and ARIC cohorts: PGS stratifies risk similarly in 23andMe and external cohorts.

Cohort	Genetics Only AUC (95%CIs)	Odds Ratio top 20% versus average (95%CIs)	Odds Ratio top 20% versus bottom 20% (95%CIs)
23andMe European ancestry test set	0.6456 (0.6409 to 0.6503)	2.06 (1.96 to 2.18)	4.62 (4.37 to 4.90)
ARIC White	0.6716 (0.6475 to 0.6957)	2.02 (1.57 to 2.60)	5.49 (4.03 to 7.47)
23andMe African American ancestry test set	0.5985 (0.5898 to 0.6072)	1.72 (1.56 to 1.88)	2.62 (2.38 to 2.89)
ARIC Black	0.5735 (0.5245 to 0.6226)	1.22 (0.74 to 2.04)	2.07 (1.19 to 3.59)

Figure 1-7: Calibration of High LDL PGS in ARIC cohorts: PGS scores remain well calibrated in both ARIC cohorts.



2. Atrial Fibrillation (Afib)

James Ashenhurst, PhD & Jianan Zhan, PhD, lead analysts

Report added May, 2020

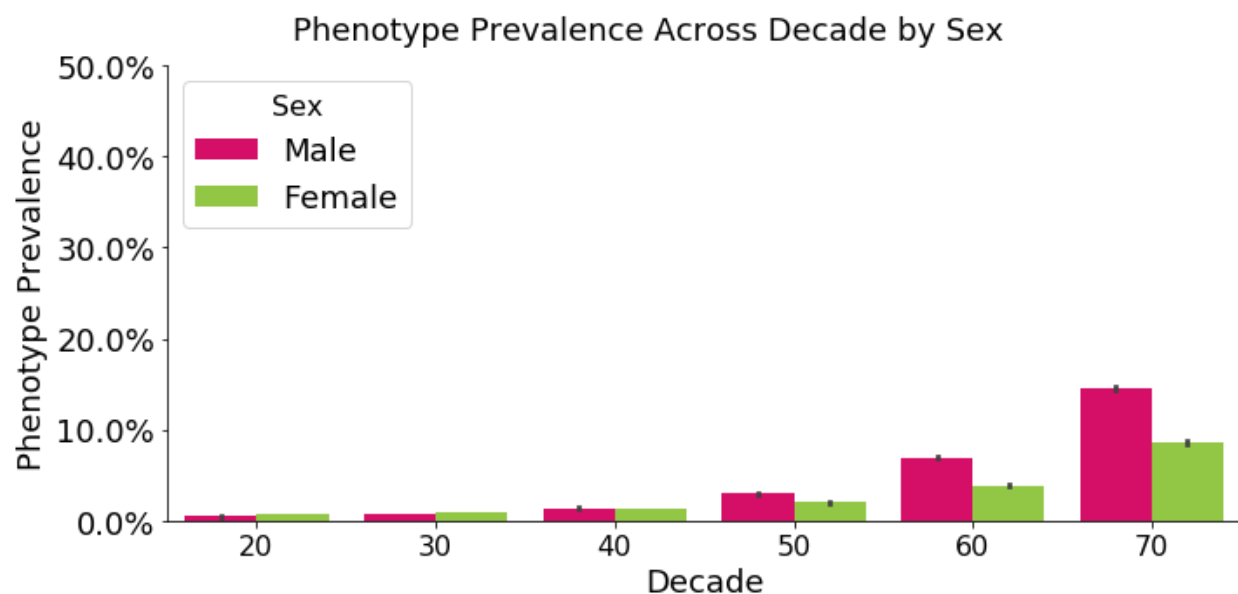
Phenotype definition

Cases and controls were defined based on self-reported responses to questions about current or past diagnosis or treatment for atrial fibrillation or flutter. All question phrasings focused on receiving a diagnosis of atrial fibrillation or flutter, whereas some included undergoing treatment for atrial fibrillation. Although some phrasings had minor differences, the general phrasing was: “Have you ever been diagnosed with atrial fibrillation or flutter?”

Participants answering affirmatively were treated as cases, while those who report no history of diagnosis or treatment for atrial fibrillation or flutter were counted as controls.

Model description

The final model selected for Afib consisted of 2,946 genetic variants, age, sex, age*sex, age², and age²*sex interaction terms, as well as the first five principal components (PCs) derived from the European ancestry population. Genomic PCs are not used when determining an individual customer's report result. This one model is applied to all customers, with personalized binary qualitative results and estimated likelihoods based on genetic ancestry and birth sex.

Figure 2-1: Afib phenotype distribution in the European training set**Table 2-1:** Afib participant cohort descriptives

Platform	Ancestry Group	Sample Use	N	Age mean (SD)	Sex (% female)	Afib prevalence (%)
V1 to V5	European	GWAS	1,746,252	50.2 (15.8)	55.29%	3.40%
V5	European	Training the European Model	1,331,117	47.9 (15.9)	57.11%	3.09%
		Testing	332,610	47.9 (15.9)	57.00%	3.10%
	Sub-Saharan African/African American	Testing	85,385	41.4 (14.5)	58.49%	1.36%
	East/Southeast Asian	Testing	77,120	36.9 (13.0)	60.66%	0.59%
	Hispanic/Latino	Testing	311,117	39.3 (13.8)	57.92%	1.10%
	South Asian	Testing	18,978	38.3 (12.4)	43.10%	0.38%
	Northern African/Western Asian	Testing	17,072	41.9 (14.8)	44.86%	0.95%

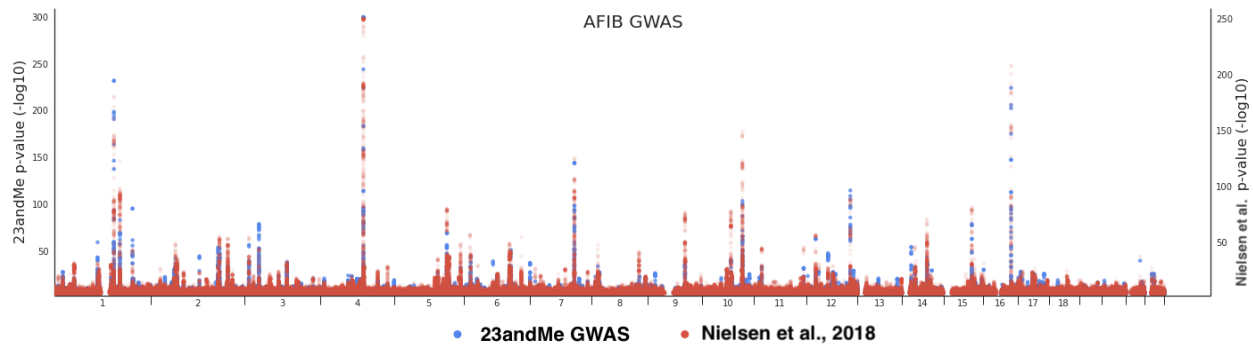


Figure 2-2: Manhattan plot of 23andMe (blue) and Nielsen et al., 2018 (red) GWAS summary statistics for atrial fibrillation. The p -values for each variant shared between two GWAS are depicted. Chromosomal location is represented on the horizontal axis, and the negative log of the p -value is represented on the vertical axis (scaled separately for each analysis).

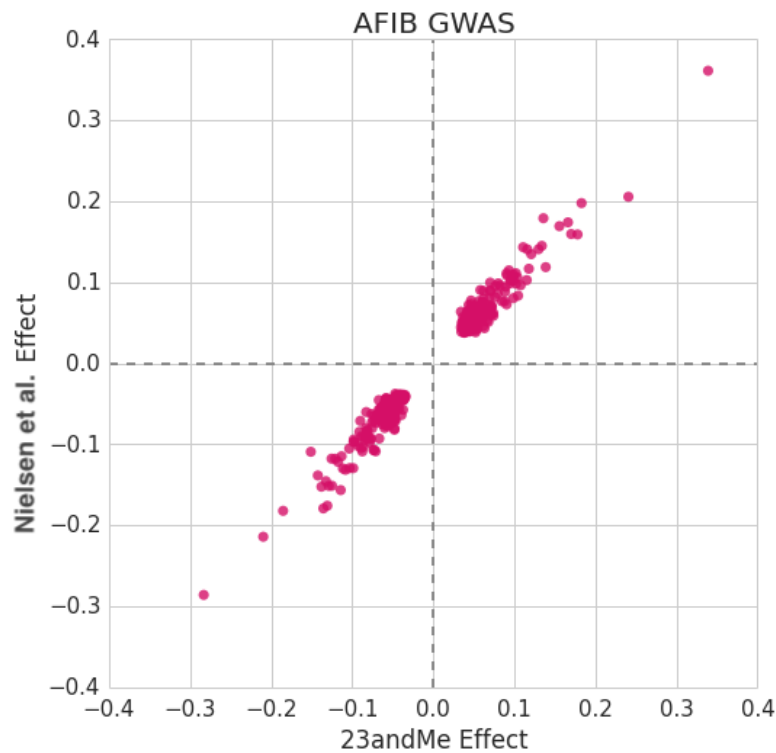


Figure 2-3: Scatter plot showing the estimated effect sizes (change in log-odds per unit predictor change) for 23andMe and Nielsen et al. (2018) genome-wide significant hits shared between the two GWAS for atrial fibrillation.

Table 2-2: Afib model performance characteristics

Ancestry Group	Full Model AUROC	Genetics Only AUROC	Odds Ratio top 5% versus average (95% CIs)	Odds Ratio top 5% versus bottom 5% (95% CIs)
European (Test Set)	0.7854	0.6198	2.69 (2.49 to 2.90)	5.07 (4.42 to 5.82)
Sub-Saharan African/African American	0.7118	0.5704	1.77 (1.39 to 2.27)	2.16 (1.51 to 3.10)
East/Southeast Asian	0.7990	0.6182	2.91 (2.04 to 4.17)	4.78 (2.49 to 9.17)
Hispanic/Latino	0.7577	0.5984	2.56 (2.25 to 2.92)	4.39 (3.48 to 5.54)
South Asian	0.7996	0.6128	2.41 (1.05 to 5.53)	9.09 (1.15 to 71.79)
Northern African/Western Asian	0.8352	0.6443	2.24 (1.19 to 4.24)	15.25 (2.01 to 115.71)

Table 2-3: Afib qualitative result characteristics

Ancestry Group (test sets)	Odds Ratio for Result Threshold	Percent Above Threshold	Likelihood Ratio of "Increased" Result (95% CIs)
European	1.5	17.68%	1.81 (1.76 to 1.86)
Sub-Saharan African/African American	1.5	5.39%	1.64 (1.36 to 1.98)
East/Southeast Asian	1.5	16.01%	1.77 (1.53 to 2.04)
Hispanic/Latino	1.5	13.40%	1.75 (1.65 to 1.86)
South Asian	1.5	16.01%	1.83 (1.27 to 2.62)
Northern African/Western Asian	1.5	20.40%	1.77 (1.44 to 2.18)

Baseline prevalences used for likelihood estimate

Ancestry- and sex-specific baseline prevalences of ever having had atrial fibrillation were derived from the 23andMe database because no large-scale nationally representative datasets were available (Table 2-4). Estimates were derived from consented 23andMe customers between the ages of 70 to 79 in each sex and ancestry group.

Table 2-4: Afib baseline prevalences in 23andMe data between ages 70 to 79

23andMe Ancestry Group	Sex	N	Afib Prevalence	95% CI
Sub-Saharan African/African American	Male	5,228	5.26%	4.65% to 5.87%
	Female	6,254	3.53%	3.08% to 3.99%
East/Southeast Asian	Male	2,329	5.75%	4.81% to 6.70%
	Female	2,763	3.04%	2.40% to 3.68%
European	Male	231,510	9.41%	9.30% to 9.53%
	Female	269,409	5.61%	5.53% to 5.70%
Hispanic/Latino	Male	13,434	5.95%	5.55% to 6.35%
	Female	16,513	3.72%	3.43% to 4.01%
Northern African/Western Asian	Male	1,358	5.96%	4.71% to 7.22%
	Female	895	2.91%	1.80% to 4.01%
South Asian	Male	837	2.87%	1.74% to 4.00%
	Female	437	1.83%	0.57% to 3.09%

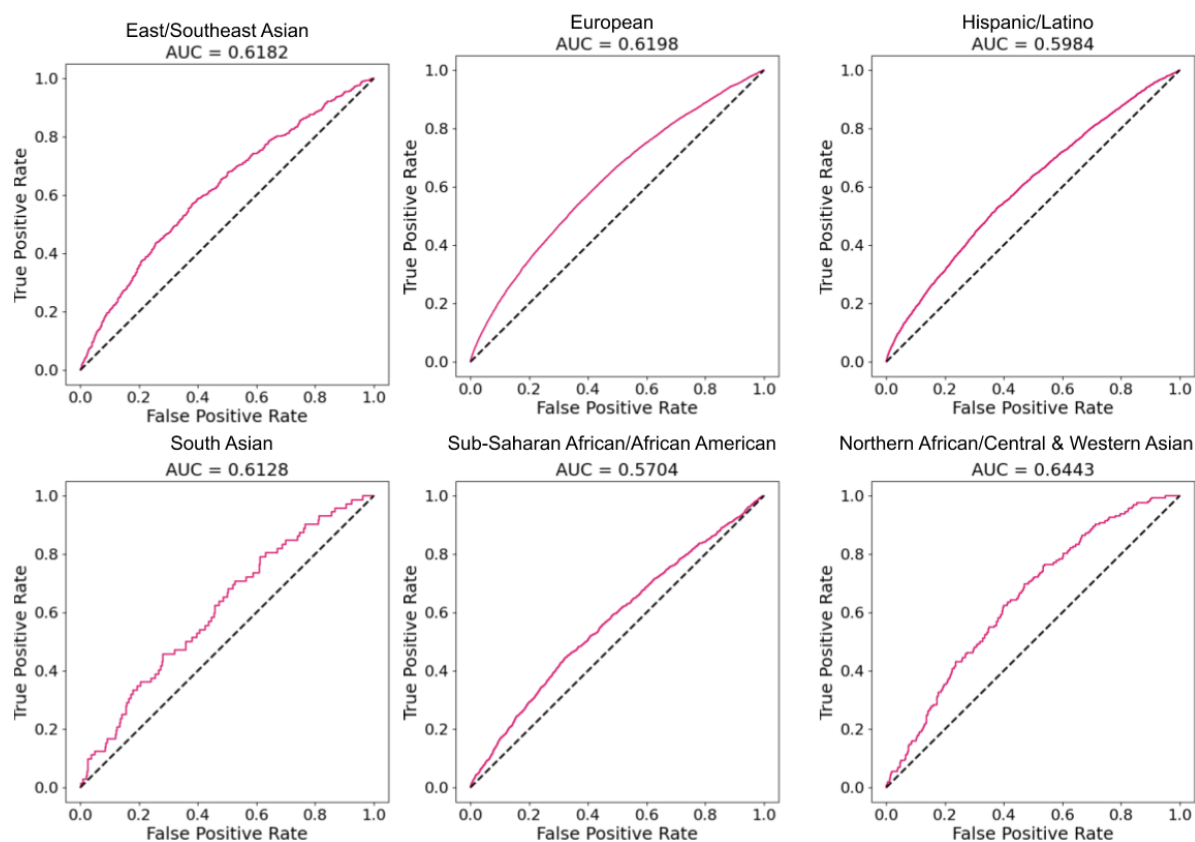
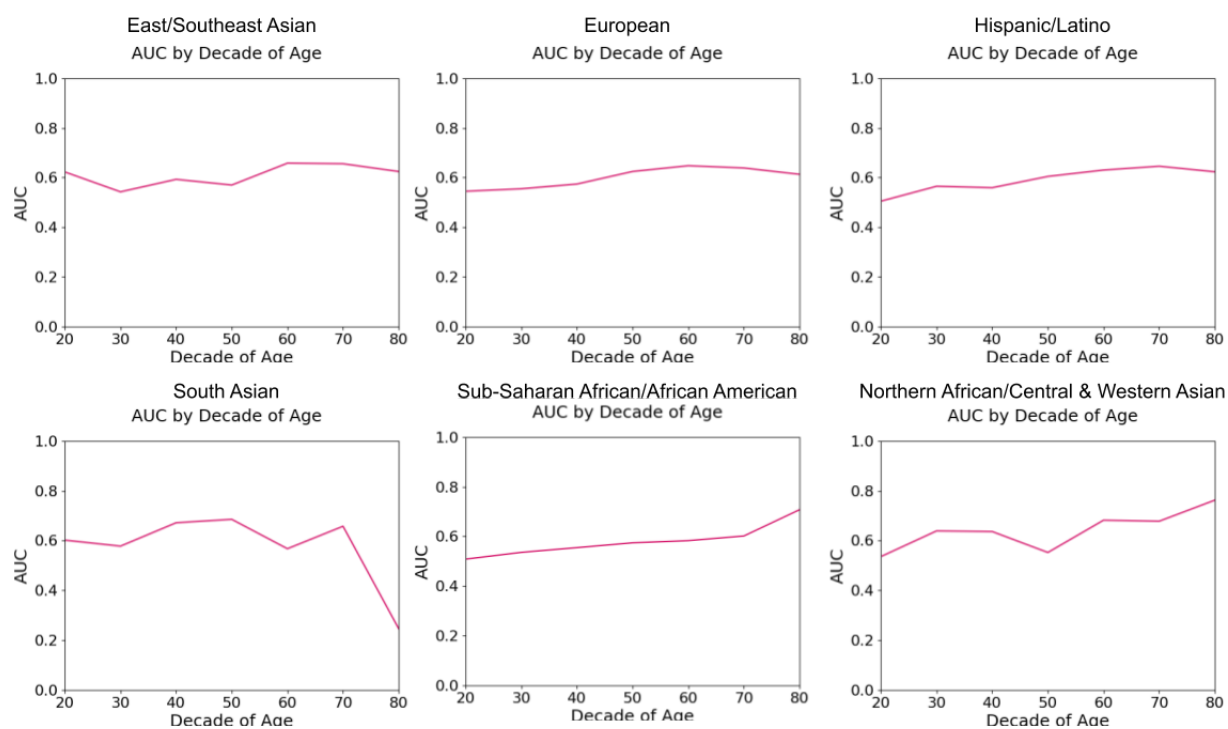
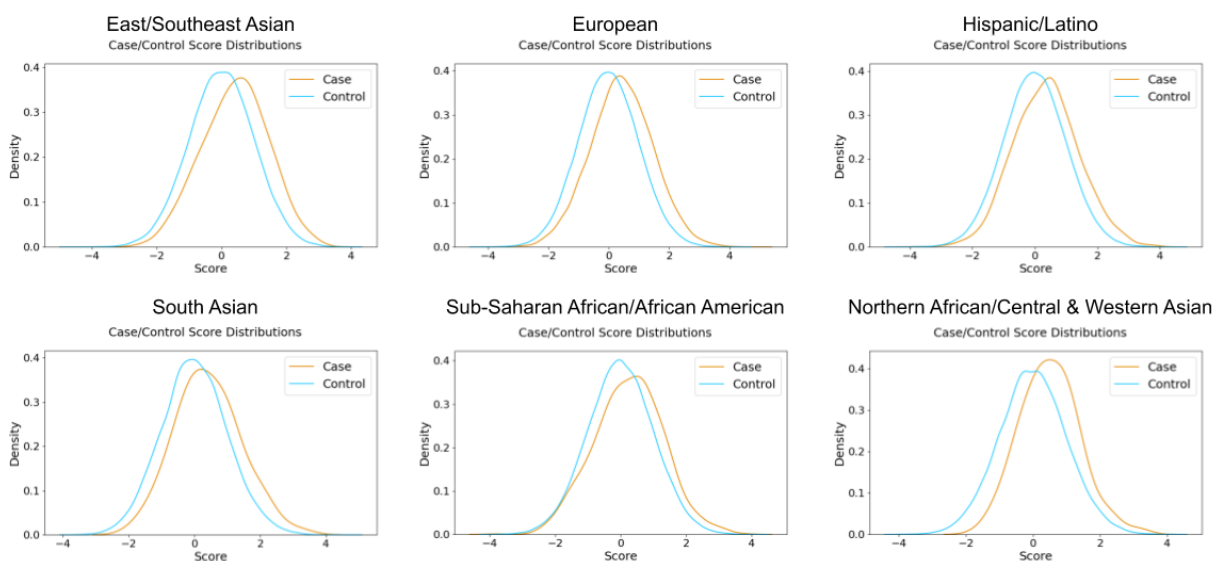
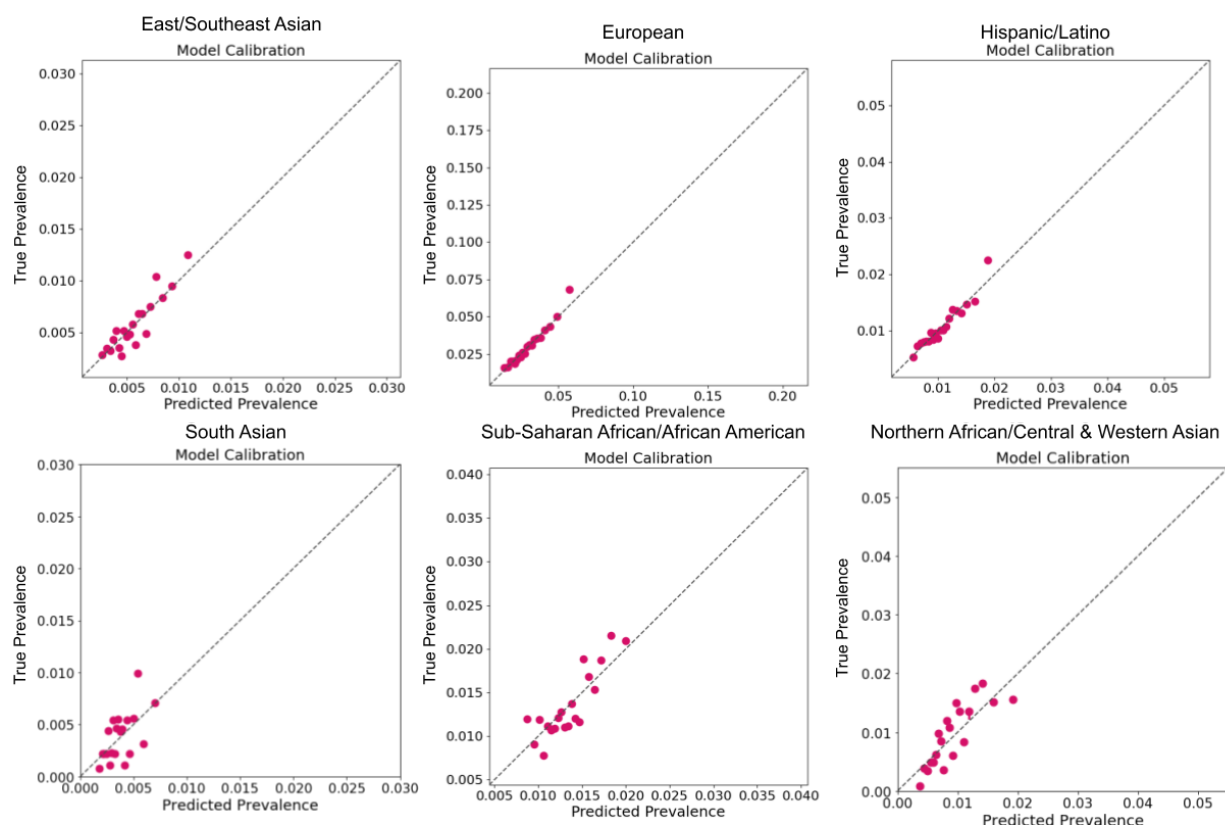
Figure 2-4: Afib AUROC across ancestry-specific test sets**Figure 2-5:** Afib AUROC within each decade of age across ancestry-specific test sets

Figure 2-6: Afib case/control standardized PGS distributions across ancestry-specific test sets**Figure 2-7:** Afib Platt-scaled calibration plots across ancestry-specific test sets

3. Coronary Artery Disease (CAD)

James Ashenhurst, PhD & Jianan Zhan, PhD, lead analysts

Report added May, 2020

Phenotype definition

Cases and controls were defined based on self-reported responses to a set of questions that indicate a positive history of coronary artery disease (CAD). This included: a) current or past diagnosis or treatment for CAD or coronary heart disease, b) history of heart attack, c) history of coronary bypass surgery, or d) history of coronary angioplasty. While some phrasings had minor differences, the general phrasing was: “Have you ever been diagnosed with or treated for coronary artery disease (CAD) or coronary heart disease (CHD)?” Participants who provided affirmative responses to any of these questions were counted as cases, whereas those who stated no history of any of these were counted as controls. If a participant answered multiple questions with some as a case and some as a control, they were counted as cases.

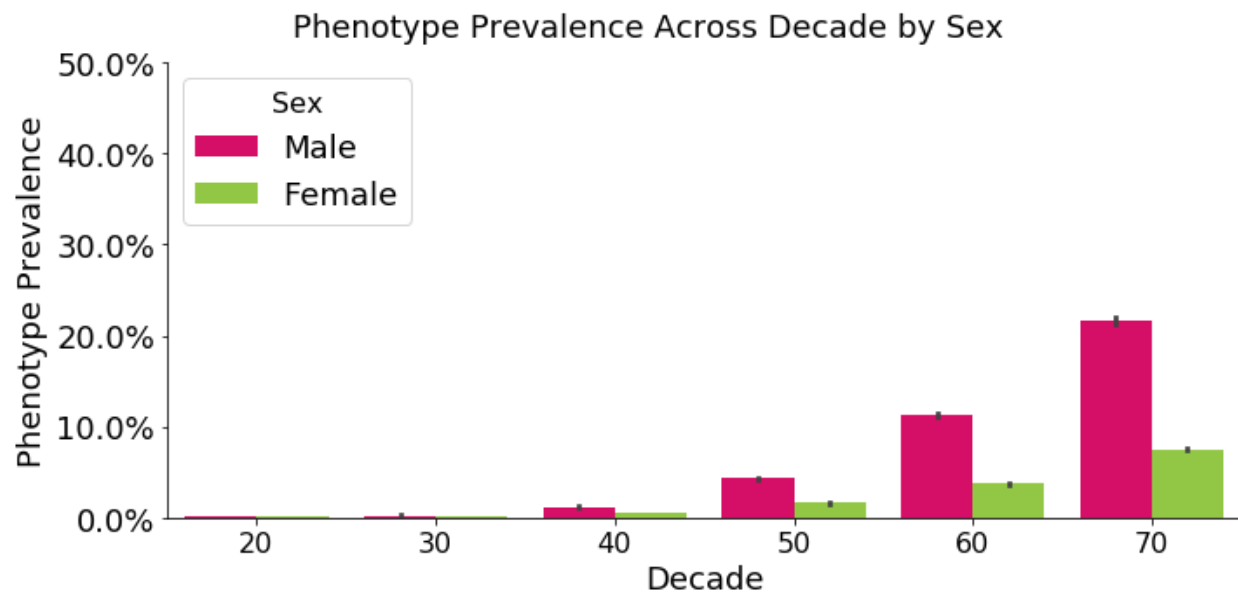
Model description

For the CAD report there are two models. In order to create a better-performing CAD model for the Hispanic/Latino ancestry group, this cohort was divided into one validation and one testing cohort. Many different variant sets (based on variant selection hyper-parameter permutations) were used to train candidate models in the European ancestry cohort, but performance was optimized in the Hispanic/Latino validation cohort. The final Hispanic/Latino model was trained on the European training+testing+validation cohorts, with out-of-sample final performance metrics evaluated on the Hispanic/Latino testing cohort. The data were insufficient to create custom models for the other groups.

Thus, the final two models selected for CAD were optimized for A) the European ancestry group (2,917 genetic variants) and B) the Hispanic/Latino ancestry group (2,456 genetic variants). Both models had age, sex, age*sex, age², and age²*sex interaction terms, as well as the first five principal components (PCs) derived from the European ancestry population. Genomic PCs are not used when determining an individual customer’s report result. The Hispanic/Latino model is provided to customers who self-identify as Hispanic/Latino, whereas the European-trained model is provided to all other customers, with personalized binary qualitative results and estimated likelihoods based on genetic ancestry and birth sex.

Table 3-1: CAD participant cohort descriptives

Platform	Ancestry group	Sample Use	N	Age Mean (SD)	Sex (% female)	CAD prevalence (%)
V1 to V5	European	GWAS	1,763,982	50.3 (15.9)	55.13%	4.09%
V5	European	Training the European Model	1,371,404	48.3 (15.9)	56.54%	3.58%
		Testing	342,775	48.3 (15.8)	56.45%	3.56%
	European	Training the Hispanic/Latino Model	1,714,179	48.3 (15.9)	56.52%	3.58%
	Sub-Saharan African/African American	Testing	135,080	41.5 (14.6)	58.91%	1.29%
	East/Southeast Asian	Testing	79,268	37.1 (13.1)	60.58%	0.78%
	Hispanic/Latino	Testing	159,781	39.6 (13.9)	57.77%	1.44%
	South Asian	Testing	19,426	38.7 (12.6)	42.74%	1.77%
	Northern African/Western Asian	Testing	17,624	42.3 (14.9)	44.46%	2.03%

Figure 3-1: CAD phenotype distributions in the European ancestry training set

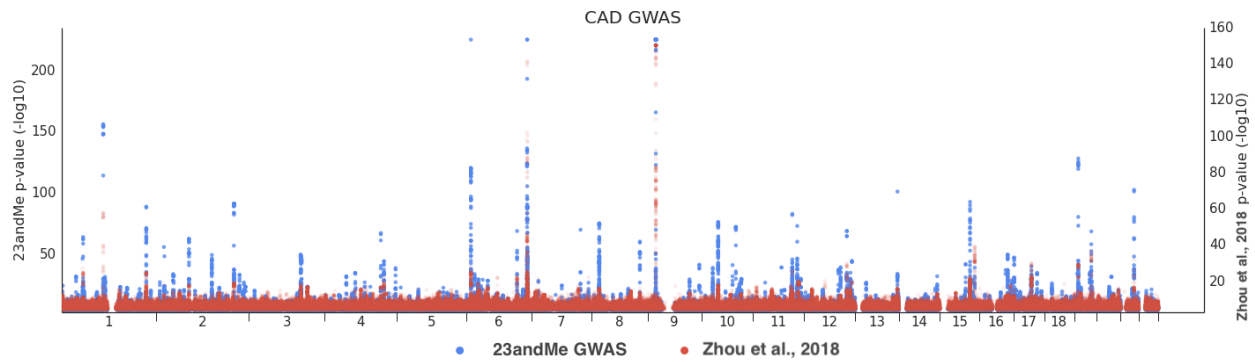


Figure 3-2: Manhattan plot of 23andMe and Zhou et al. (2018) GWAS summary statistics for CAD. The p -values for each variant shared between the 23andMe (blue) and Zhou et al. (red) genome-wide association studies (GWASs) are depicted. Chromosomal location is represented on the horizontal axis, and the negative log of the p -value is represented on the vertical axis (scaled separately for each analysis).

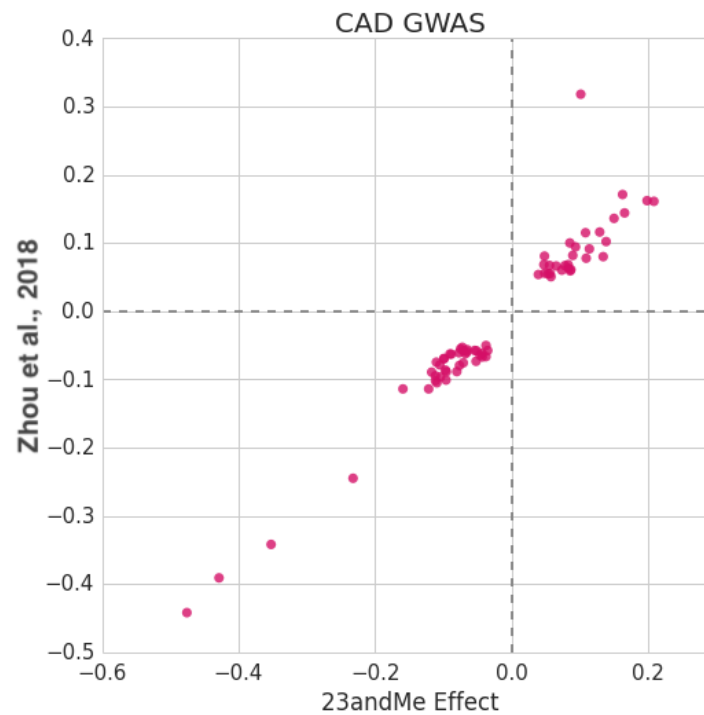


Figure 3-3: Scatter plot showing the estimated effect sizes (change in log-odds per unit predictor change) for 23andMe and Zhou et al. (2018) genome-wide significant hits shared between the two GWAS for CAD.

Table 3-2: CAD model performance characteristics

Ancestry Group (Test Sets)	Full Model AUROC	Genetics Only AUROC	Odds Ratio top 5% versus average (95% CIs)	Odds Ratio top 5% versus bottom 5% (95% CIs)
European	0.8702	0.6286	2.60 (2.43 to 2.78)	7.00 (6.06 to 8.10)
Sub-Saharan African/African American	0.8336	0.5638	1.10 n.s. (0.89 to 1.37)	2.47 (1.73 to 3.55)
East/Southeast Asian	0.8942	0.5846	1.98 (1.46 to 2.70)	3.08 (1.86 to 5.12)
Hispanic/Latino	0.8790	0.6595	2.14 (1.82 to 2.53)	6.41 (4.48 to 9.16)
South Asian	0.9258	0.5773	1.76 (1.16 to 2.67)	4.84 (2.13 to 11.01)
Northern African/Western Asian	0.8998	0.6016	2.10 (1.36 to 3.24)	4.71 (2.07 to 10.72)

Table 3-3: CAD qualitative result characteristics

Ancestry Group (test sets)	Odds Ratio for Result Threshold	Percent above threshold	Likelihood ratio of "increased" result (95% CIs)
European	1.5	19.26%	1.80 (1.76 to 1.85)
Sub-Saharan African/African American	1.5	2.22%	1.35 (1.03 to 1.77)
East/Southeast Asian	1.5	9.49%	1.86 (1.56 to 2.21)
Hispanic/Latino	1.5	14.97%	1.70 (1.58 to 1.83)
South Asian	1.5	7.13%	1.61 (1.19 to 2.17)
Northern African/Western Asian	1.5	12.05%	1.82 (1.48 to 2.22)

Baseline prevalences used for likelihood estimate

Ancestry- and sex-specific baseline prevalences of ever having had CAD were derived from the 2017 data release of the Behavioral Risk Factor Surveillance System (BRFSS; Centers for Disease Control and Prevention [CDC], 2017). The specific variable (coded cvdcrhd4) represents the question: “Ever told you had angina or coronary heart disease?” The ancestry variable used (coded _RACE) included the categories White non-Hispanic, Black non-Hispanic, Asian non-Hispanic, and Hispanic. Analysis was restricted only to those between the ages of 70 and 79, to capture this decade of age (coded _AGEG5YR). The baselines used for each sex and ancestry combination and how they map to each 23andMe ancestry group are shown in the table below.

Table 3-4: CAD baseline prevalences in BRFSS data between ages 70 to 79

Group	Matched 23andMe Population(s)	Sex	N	Prevalence	95% CI
White Non-Hispanic	European, Northern African/Western Asian, Other	Male	23,878	16.01%	15.55% to 16.48%
		Female	34,515	8.09%	7.80% to 8.38%
Black Non-Hispanic	Sub-Saharan African/African American	Male	1,378	11.10%	9.44% to 12.76%
		Female	2,871	10.28%	9.16% to 11.39%
Asian Non-Hispanic	East/Southeast Asian, South Asian	Male	332	10.54%	7.24% to 13.85%
		Female	398	2.01%	0.63% to 3.39%
Hispanic	Hispanic/Latino	Male	998	11.12%	9.17% to 13.07%
		Female	1,691	7.51%	6.25% to 8.77%

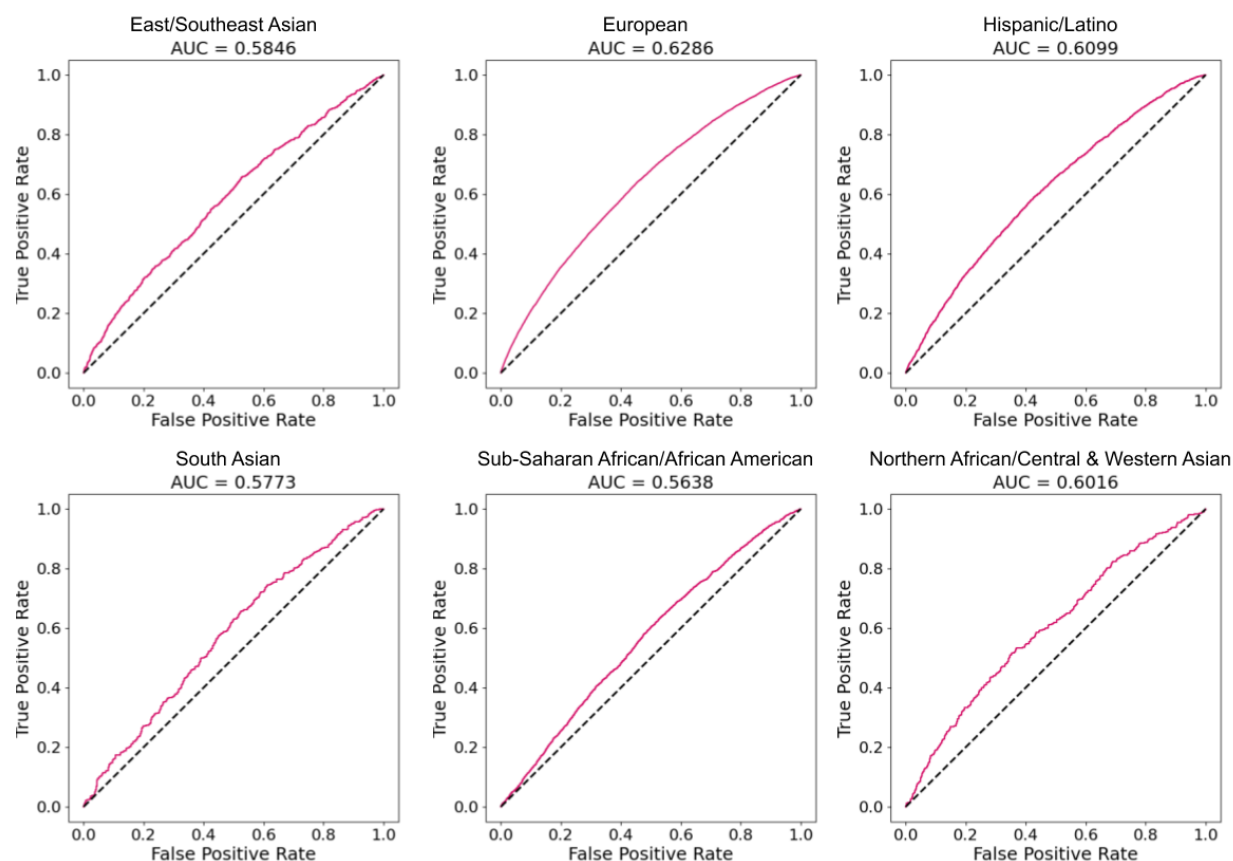
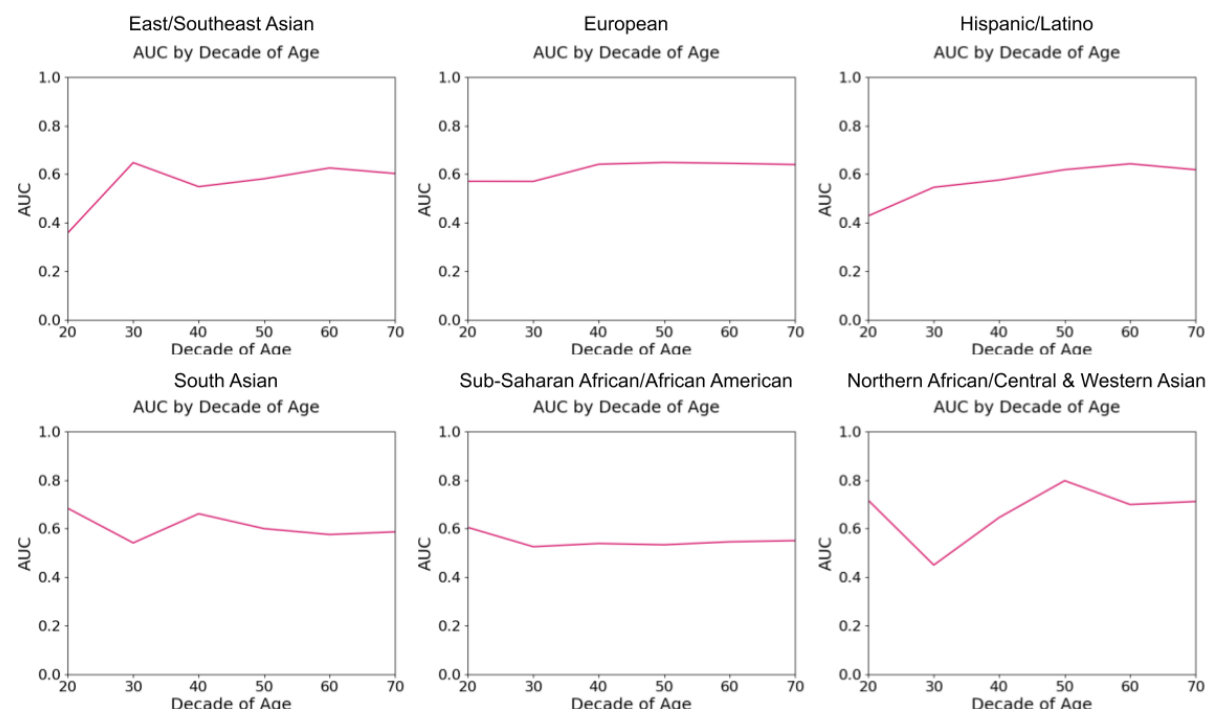
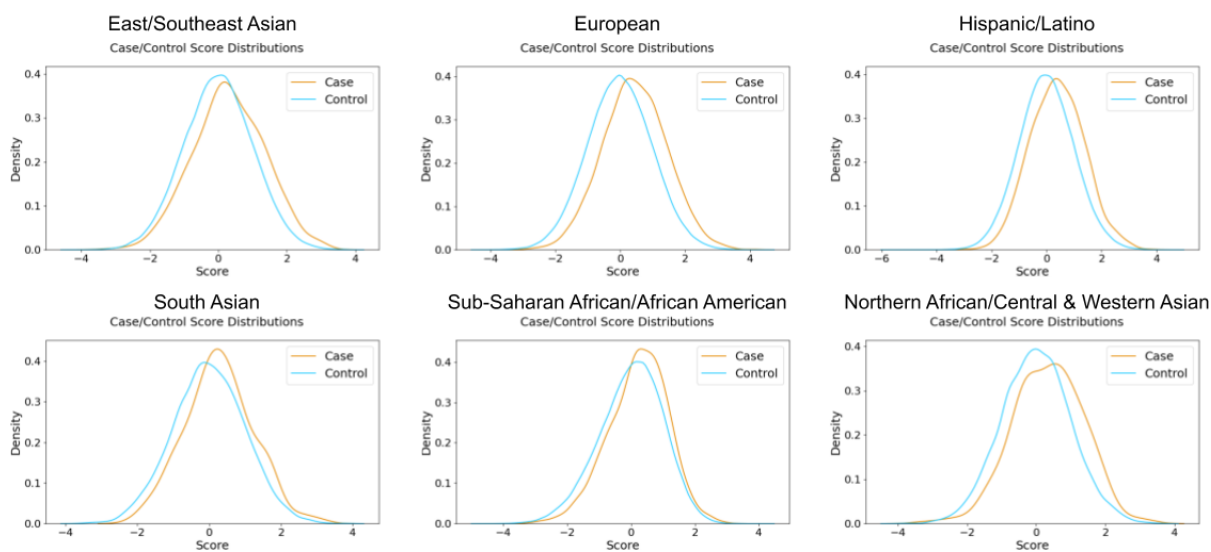
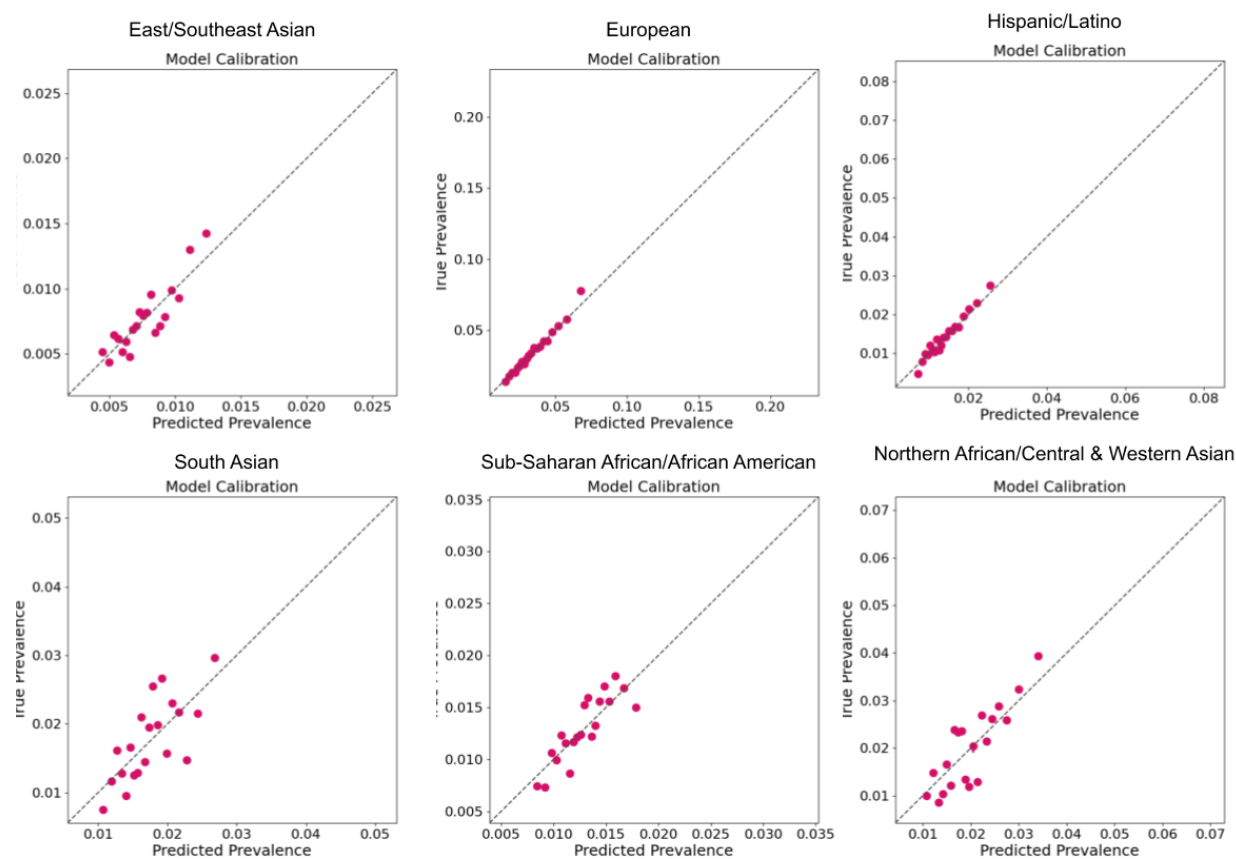
Figure 3-4: CAD AUROC across ancestry-specific test sets**Figure 3-5:** CAD AUROC within each decade of age across ancestry-specific test sets

Figure 3-6: CAD case/control standardized PGS distributions across ancestry-specific test sets**Figure 3-7:** CAD Platt-scaled calibration plots across ancestry-specific test sets

External Validation

As described in the generalized method White Paper, we conducted an external analysis on our CAD PGS model in the Atherosclerosis Risk in Communities (ARIC) study cohort. To determine CAD status, we combined [prvchd05](#) and [In_04sp](#), phenotypes for coronary heart disease at baseline and follow up, respectively, that combine both self reported history of procedures and/or conditions as well as clinical diagnosis, including myocardial infarction, coronary bypass, and balloon angioplasty. Any subject who was a case for either phenotype was coded as a case, and any subject who was not a case for either phenotype was coded as a control.

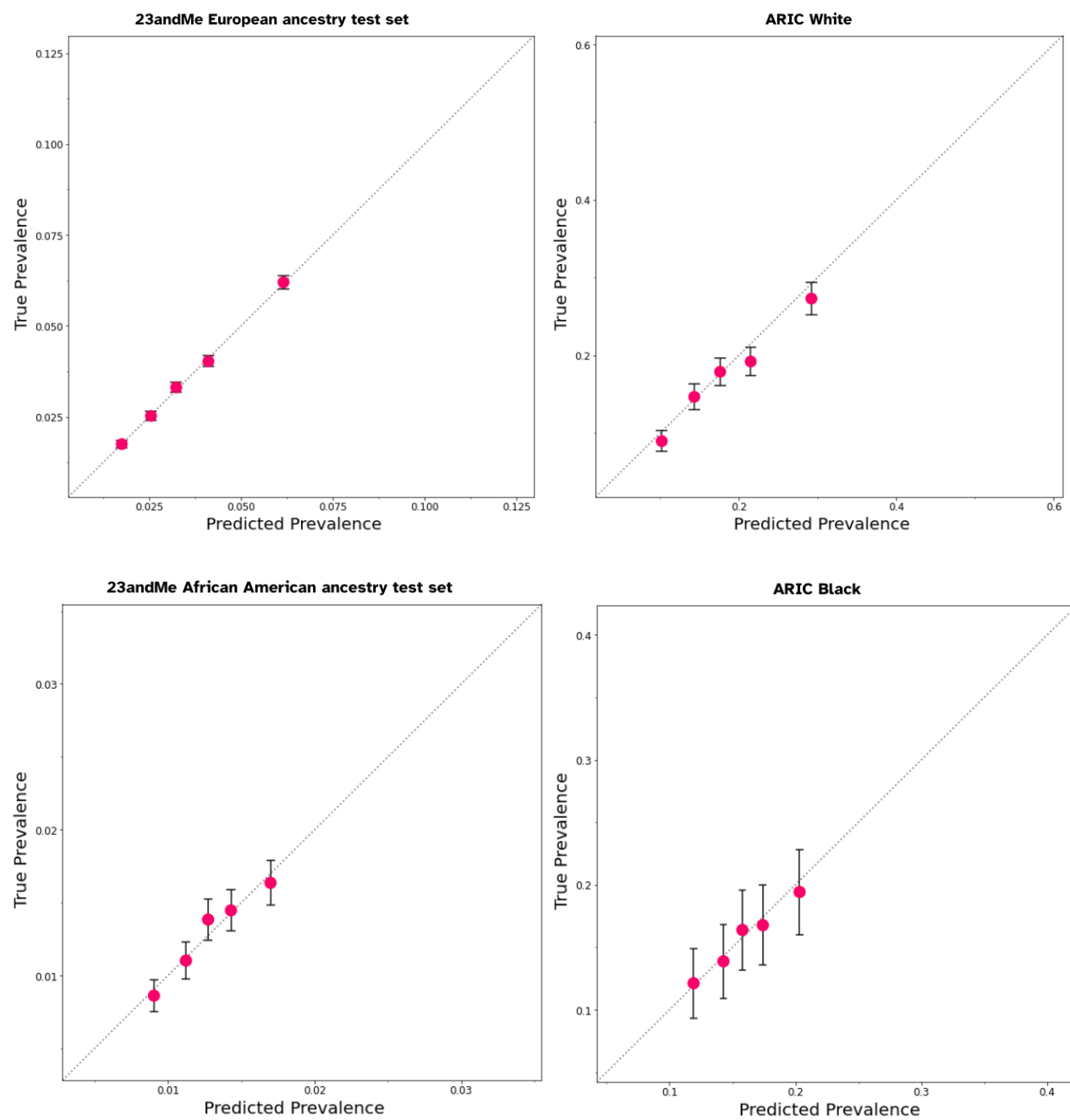
Table 3-5: ARIC cohort statistics for CHD

Cohort	N	Average earliest age	Average latest age	Average years in study	Sex (% female)	CHD Prevalence (%)
ARIC White	8,697	54.6	68.9	14.6	52.96%	17.60%
ARIC Black	2,597	53.7	67.7	14.4	62.11%	15.71%

Table 3-6: CAD PGS performance in 23andMe and ARIC cohorts: PGS stratifies risk similarly in 23andMe and external cohorts.

Cohort	Genetics Only AUC (95%CI)	Odds Ratio top 20% versus average (95%CI)	Odds Ratio top 20% versus bottom 20% (95%CI)
23andMe European ancestry test set	0.6286 (0.6232 to 0.6339)	1.93 (1.83 to 2.04)	3.70 (3.47 to 3.95)
ARIC White	0.6192 (0.603 to 0.6353)	1.72 (1.47 to 2.03)	3.81 (3.13 to 4.63)
23andMe African American ancestry test set	0.5638 (0.5497 to 0.5779)	1.18 (1.03 to 1.36)	1.91 (1.63 to 2.24)
ARIC Black	0.5548 (0.5237 to 0.5858)	1.23 (0.90 to 1.69)	1.75 (1.24 to 2.46)

Figure 3-8: Calibration of CAD PGS in ARIC cohorts: PGS scores remain well calibrated in both ARIC cohorts.



4. High Blood Pressure (HBP)

Ryo Kita, MD, PhD, lead analyst

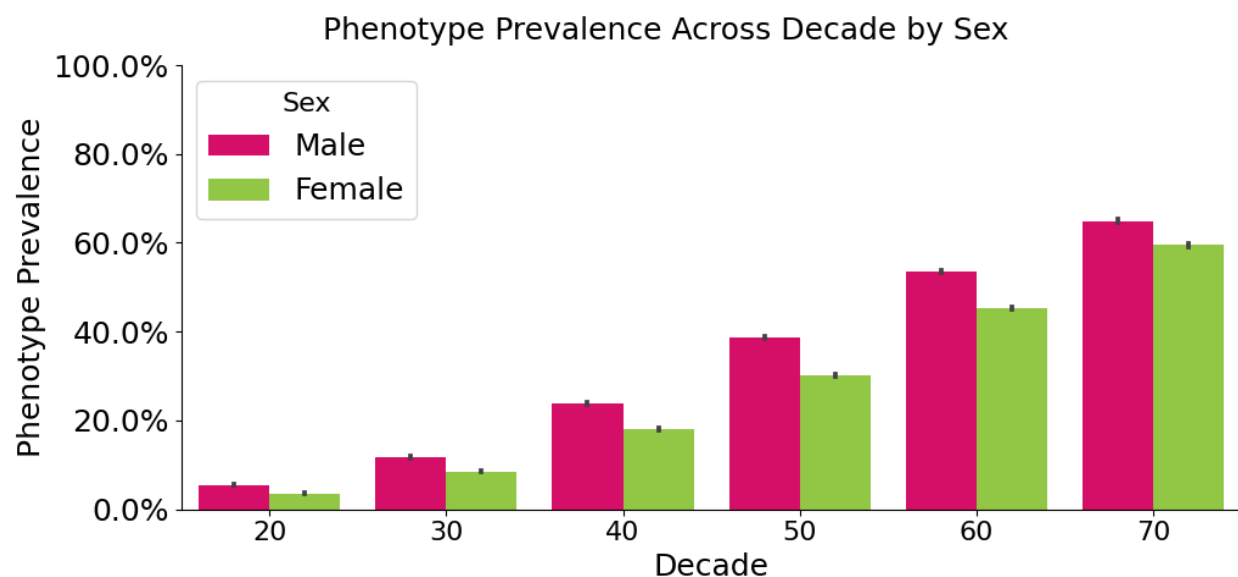
Report added May, 2020

Phenotype definition

Cases and controls were defined based on self-reported responses to questions about current or past diagnosis or treatment for high blood pressure. All question phrasings focused on receiving a diagnosis of high blood pressure, and some included the concept of treatment for high blood pressure. Although some phrasings had minor differences, the general phrasing was: “Have you ever been diagnosed with or treated for high blood pressure?” Participants answering affirmatively were treated as cases, while those who report no history of diagnosis or treatment for high blood pressure were counted as controls.

Model description

For the development of this report, the data were sufficient to attempt ancestry-specific GWAS. Four separate GWAS were conducted in East/Southeast Asian, European, Hispanic/Latino, and Sub-Saharan African/African American populations. Pruning and thresholding (P+T) using different hyperparameter combinations was performed on each of the four ancestry-specific GWAS. For the European model, the best SNP set was selected from the P+T candidate sets from the European GWAS. For the remaining three non-European groups, each of the P+T SNP sets for that ancestry-specific GWAS was combined with each of the P+T SNP sets from the European GWAS to generate a large list of candidate multi-ancestral unioned SNP sets. Each of these candidate feature sets were tested on the ancestry-specific validation sets to select the optimal feature set for that group. The final four variant sets were used to train models in a multi-ancestral training cohort, and evaluated in the ancestry-specific test sets (see Tables 4-0 and 4-1). All models had age, sex, age*sex, age², and age²*sex interaction terms, as well as the first five global principal components (PCs) to account for population stratification. Genomic PCs are not used when determining an individual customer’s report result. Personalized binary qualitative results and estimated likelihoods based on genetic ancestry and birth sex.

Figure 4-1: HBP phenotype distribution in the combined multi-ancestry training set**Table 4-0:** HBP model and variant set descriptions

Optimized Ancestry Group	Applicable Ancestry Group(s)	No. of variants
European	European, South Asian, Northern African/Western Asian	2,957
Hispanic/Latino	Hispanic/Latino	2,989
Sub-Saharan African/African American	Sub-Saharan African/African American	2,966
East/Southeast Asian	East/Southeast Asian	2,912

Note: Four models were generated, each optimized for performance in different ancestry groups. This table specifies the number of variants selected for each model, the mapping between groups the models were optimized for, and the groups that will receive these models based on genetic ancestry.

Table 4-1: HBP participant cohort descriptives

Platform	Ancestry Group	Sample Use	N	Age mean (SD)	Sex (% female)	HBP prevalence (%)
V1 to V5	East/Southeast Asian	GWAS	80,261	37.4 (13.3)	58.91%	13.47%
V1 to V5	European	GWAS	1,662,414	49.0 (15.9)	53.54%	29.63%
V1 to V5	Hispanic/Latino	GWAS	287,578	40.3 (14.2)	55.51%	19.71%
V1 to V5	Sub-Saharan African/African American	GWAS	93,054	42.5 (14.8)	56.58%	30.82%
V5	All	Training four ancestry-specific models	1,390,809	46.7 (16.0)	54.79%	26.39%
	European	Testing	131,350	48.7 (15.9)	54.45%	28.44%
	Sub-Saharan African/African American	Testing	7,273	42.7 (14.8)	57.36%	29.22%
	East/Southeast Asian	Testing	6,678	37.4 (13.0)	60.50%	12.86%
	Hispanic/Latino	Testing	25,241	40.0 (14.1)	56.76%	18.05%
	South Asian	Testing	1,604	39.0 (12.7)	41.83%	11.35%
	Northern African/Western Asian	Testing	1,501	42.2 (14.7)	43.97%	14.59%

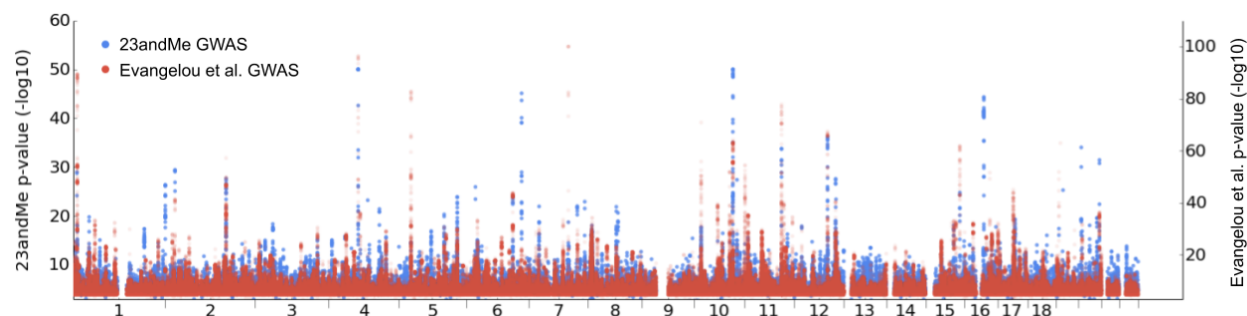


Figure 4-2: Manhattan plot of 23andMe (blue; ever had high blood pressure) and Evangelou et al. (2018; red; systolic blood pressure lab value) GWAS summary statistics for blood pressure measures. The p -values for each variant shared between the two GWAS are depicted. Chromosomal location is represented on the horizontal axis, and the negative log of the p -value is represented on the vertical axis (scaled separately for each analysis).

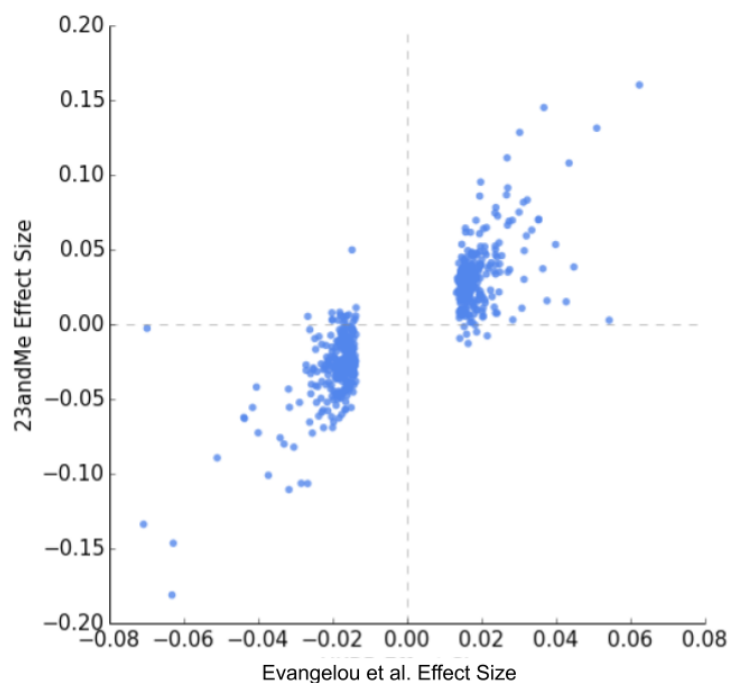


Figure 4-3: Scatter plot showing the estimated effect sizes (change in log-odds per unit predictor change) for 23andMe and Evangelou et al. (2018; linear betas) genome-wide significant hits shared between the two GWAS for blood pressure phenotypes.

Table 4-2: HBP model performance characteristics

Ancestry Group (test sets)	Full Model AUROC	Genetics Only AUROC	Odds Ratio top 5% versus average (95% CIs)	Odds Ratio top 5% versus bottom 5% (95% CIs)
European	0.8038	0.6218	2.47 (2.34 to 2.61)	6.35 (5.82 to 6.92)
Sub-Saharan African/African American	0.8114	0.6034	1.85 (1.46 to 2.35)	6.58 (4.43 to 9.76)
East/Southeast Asian	0.8202	0.6016	2.73 (2.01 to 3.72)	5.14 (3.03 to 8.70)
Hispanic/Latino	0.8195	0.6122	2.21 (1.92 to 2.54)	5.72 (4.50 to 7.27)
South Asian	0.8003	0.5973	1.61 (0.81 to 3.21)	4.97 (1.36 to 18.18)
Northern African/Western Asian	0.8024	0.6047	1.40 (0.73 to 2.67)	4.43 (1.40 to 14.04)

Table 4-3: HBP qualitative result characteristics

Ancestry group (test sets)	Odds Ratio for Result Threshold	Percent Above Threshold	Likelihood Ratio of "Increased" Result (95% CIs)
European	1.5	18.18%	1.82 (1.78 to 1.86)
Sub-Saharan African/African American	1.5	13.24%	1.65 (1.46 to 1.86)
East/Southeast Asian	1.5	14.42%	1.68 (1.46 to 1.93)
Hispanic/Latino	1.5	15.70%	1.67 (1.57 to 1.78)
South Asian	1.5	11.97%	1.74 (1.25 to 2.43)
Northern African/Western Asian	1.5	13.32%	1.75 (1.30 to 2.35)

Baseline prevalences used for likelihood estimate

Ancestry- and sex-specific baseline prevalences of ever having had HBP were derived from the 2017 data release of the Behavioral Risk Factor Surveillance System (BRFSS; Centers for Disease Control and Prevention [CDC], 2017). The specific calculated variable (coded _RFHYPE5) represents the concept: adults who have been told they have high blood pressure by a doctor, nurse, or other health professional. The ancestry variable used (coded _RACE) included the categories White non-Hispanic, Black non-Hispanic, Asian non-Hispanic, and Hispanic. Analysis was restricted only to those between the ages of 70 and 79, to capture this decade of age (coded _AGEG5YR). The baselines used for each sex and ancestry combination and how they map to each 23andMe ancestry group are shown in the table below.

Table 4-4: HBP baseline prevalences in BRFSS data between ages 70 to 79

Group	Matched 23andMe Population(s)	Sex	N	Prevalence	95% CI
White Non-Hispanic	European, Northern African/Western Asian, Other	Male	24,178	61.89%	61.28% to 62.50%
		Female	34,823	59.27%	58.75% to 59.78%
Black Non-Hispanic	Sub-Saharan African/African American	Male	1,410	74.61%	72.34% to 76.88%
		Female	2,919	82.22%	80.83% to 83.61%
Asian Non-Hispanic	East/Southeast Asian, South Asian	Male	335	56.42%	51.11% to 61.73%
		Female	402	60.95%	56.18% to 65.71%
Hispanic	Hispanic/Latino	Male	1,003	63.91%	60.94% to 66.88%
		Female	1,712	66.41%	64.18% to 68.65%

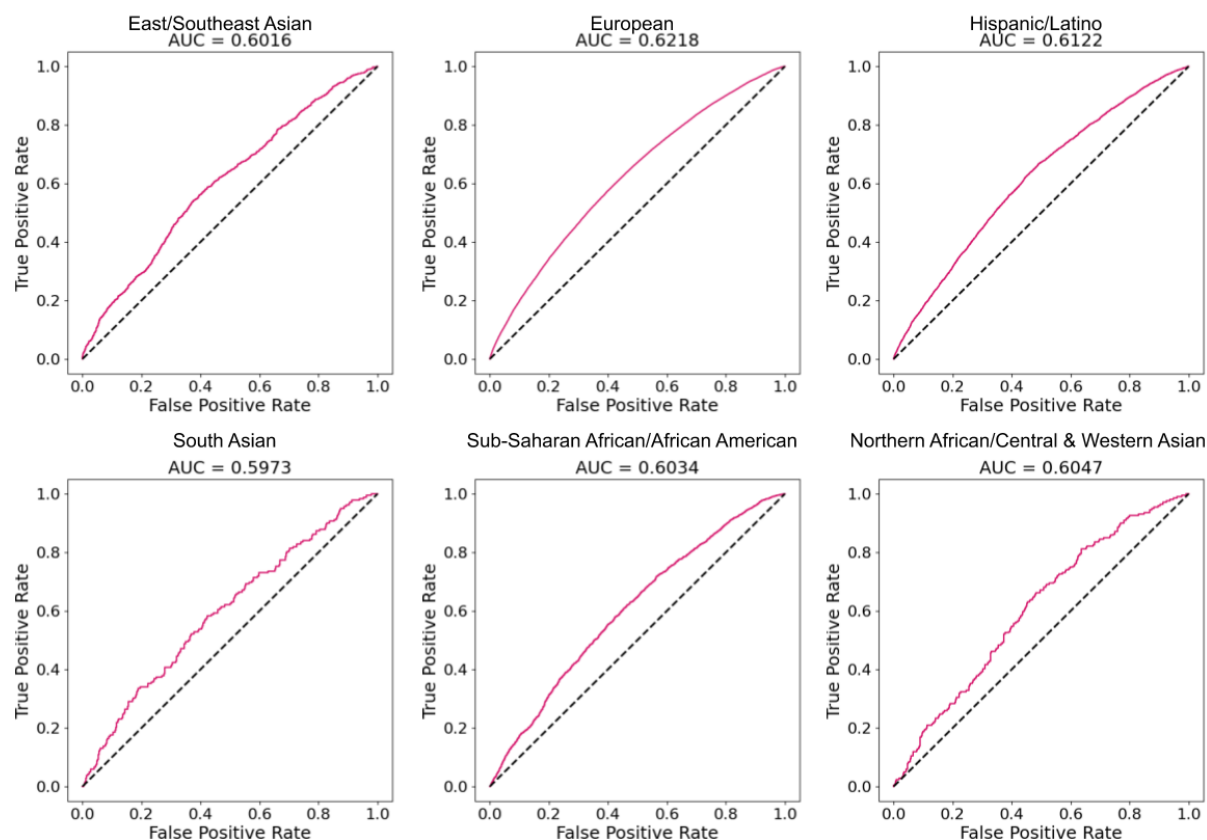
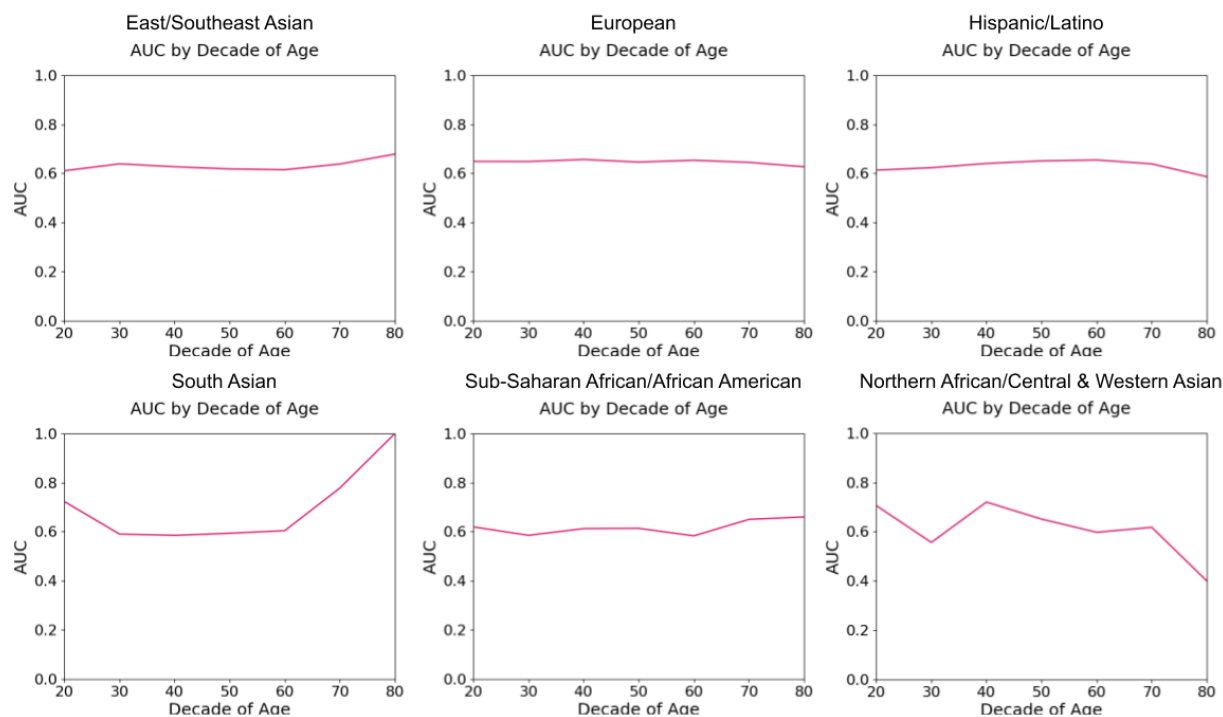
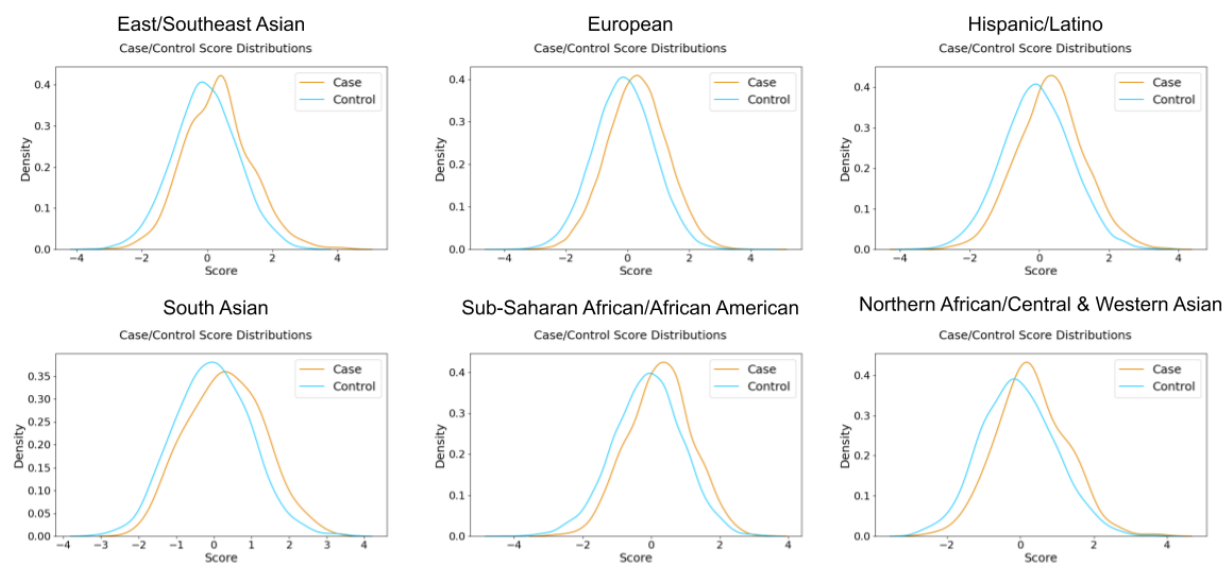
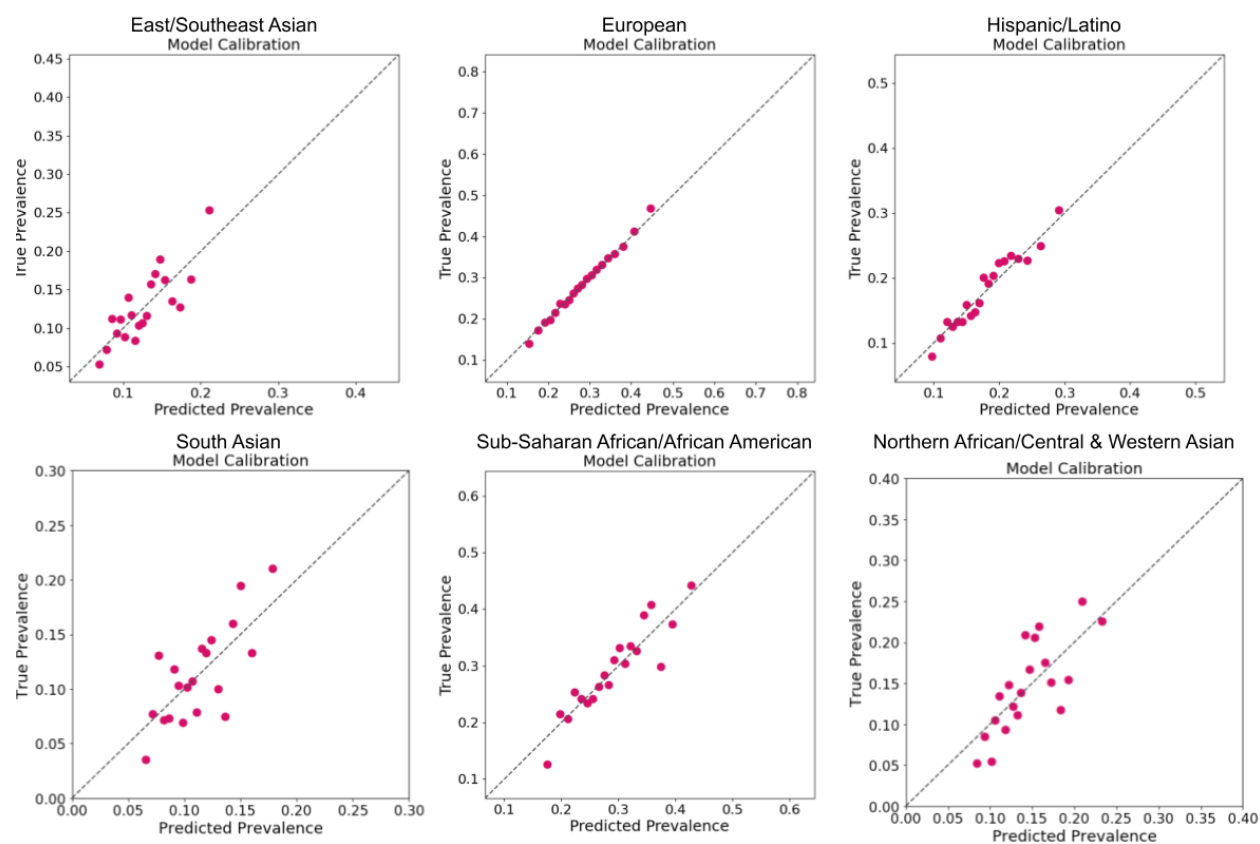
Figure 4-4: HBP AUROC across ancestry-specific test sets**Figure 4-5:** HBP AUROC within each decade of age across ancestry-specific test sets

Figure 4-6: HBP case/control standardized PGS distributions across ancestry-specific test sets**Figure 4-7:** HBP Platt-scaled calibration plots across ancestry-specific test sets

External Validation

As described in the generalized method White Paper, we conducted an external validation of our HBP PGS model in the Atherosclerosis Risk in Communities (ARIC) study cohort. To determine HBP status, we combined phenotypes [sbpa21](#), [sbpa22](#), [sbpd19](#), [sbpd20](#), [hypert05](#), [hyptmdcode01](#), and [Hyptmdcode41](#). These phenotypes report clinically measured systolic/diastolic blood pressure and self-reported use of hypertension lowering medications in the past two weeks collected at two different time points. Any subject who reported using hypertension lowering medication, had a systolic blood pressure ≥ 130 , or had a diastolic blood pressure ≥ 80 was determined to be a case. Any participant who did not meet these requirements was determined to be a control. If a participant's value for any phenotype was null, that phenotype did not contribute to the classification. If all phenotypes were null, the participant was removed from the analysis.

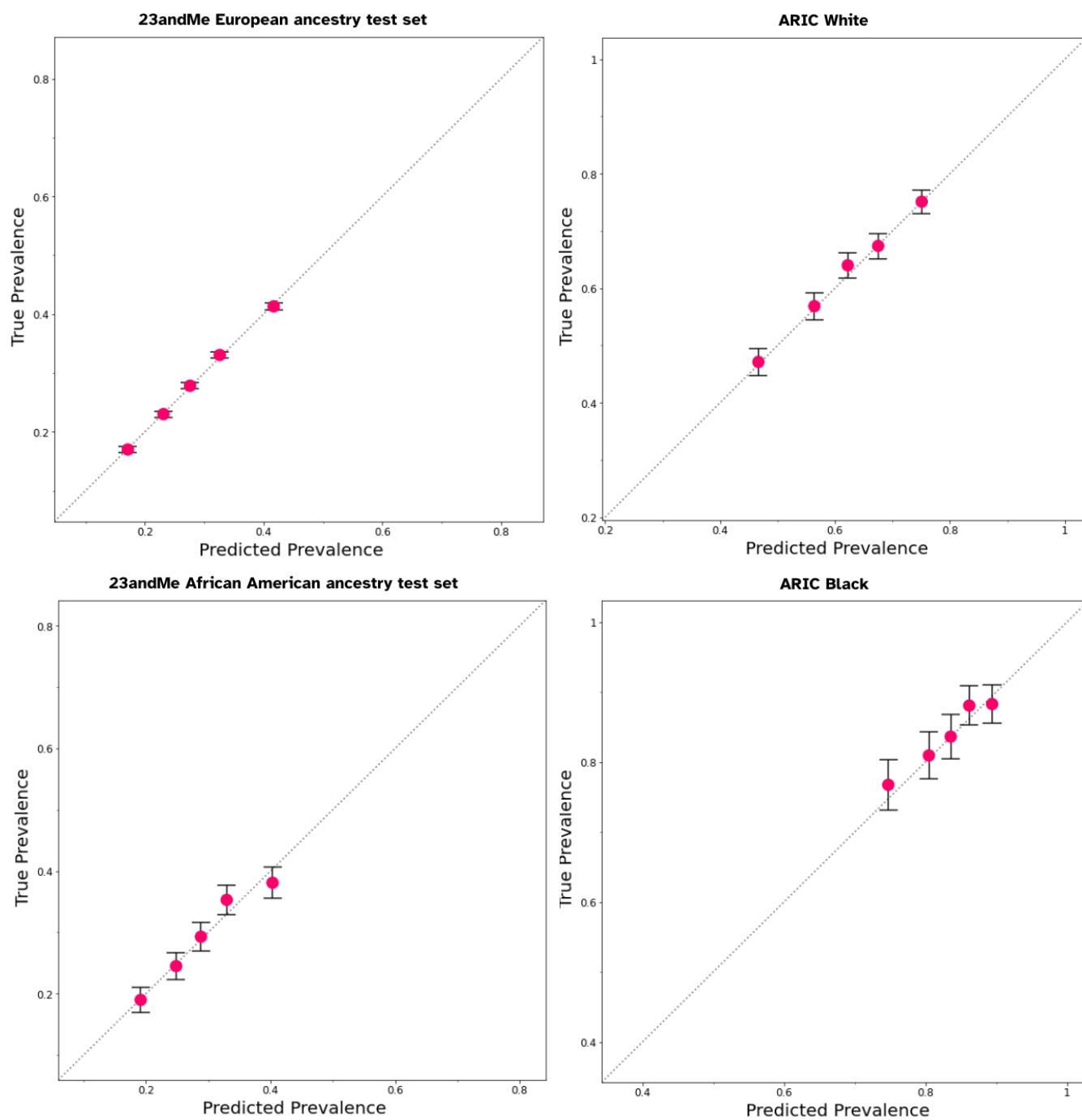
Table 4-5: ARIC cohort statistics for HBP

Cohort	N	Average earliest age	Average latest age	Average years in study	Sex (% female)	Hypertension Prevalence (%)
ARIC White	8,697	54.3	61.5	7.2	52.96%	62.10%
ARIC Black	2,597	53.4	59.5	6.2	62.11%	83.52%

Table 4-6: HBP PGS performance in 23andMe and ARIC cohorts: PGS stratifies risk similarly in 23andMe and external cohorts.

Cohort	Genetics Only AUC (95%CI)	Odds Ratio top 20% versus average (95%CI)	Odds Ratio top 20% versus bottom 20% (95%CI)
23andMe European ancestry test set	0.6218 (0.6184 to 0.6253)	1.83 (1.76 to 1.89)	3.44 (3.30 to 3.58)
ARIC White	0.6211 (0.6092 to 0.6329)	1.70 (1.47 to 1.97)	3.39 (2.93 to 3.91)
23andMe African American ancestry test set	0.6034 (0.5888 to 0.618)	1.49 (1.27 to 1.73)	2.63 (2.22 to 3.11)
ARIC Black	0.5947 (0.5667 to 0.6227)	1.47 (1.03 to 2.10)	2.28 (1.63 to 3.19)

Figure 4-8: Calibration of HBP PGS in ARIC cohorts: PGS scores remain well calibrated in both ARIC cohorts.



5. Migraine

Rafaela Bagur Quetglas, PhD, lead analyst

Report added July, 2020

Phenotype definition

Cases and controls were defined based on self-reported responses to questions about current or past diagnosis or treatment for migraines. All question phrasings focused on receiving a diagnosis of migraines, and some included the concept of treatment for migraines. Although some phrasings had minor differences, the general phrasing was: “Have you ever been diagnosed with or treated for migraines?” Participants answering affirmatively were treated as cases, while those who report no history of diagnosis or treatment for migraines were counted as controls.

Model description

For the development of this report, the data were sufficient to attempt ancestry-specific GWAS. Four separate GWAS were conducted in East/Southeast Asian, European, Hispanic/Latino, and Sub-Saharan African/African American populations. A fixed effect meta-analysis was performed to combine the four ancestry-specific GWASs, followed by pruning and thresholding (P+T) using different hyperparameter combinations. Each of these candidate feature sets were tested on the ancestry-specific validation sets to evaluate performance in all groups. We selected a single SNP set including 20,432 variants. This SNP set was used to train one model in a multi-ancestral training cohort, which was evaluated in the ancestry-specific test sets (see Table 5-1). This model had age, sex, an age*sex interaction term, as well as the first ten global principal components (PCs) to account for population stratification. Genomic PCs are not used when determining an individual customer’s report result. Personalized binary qualitative results and estimated likelihoods are based on genetic ancestry and birth sex.

Figure 5-1: Migraine phenotype distribution in the training set

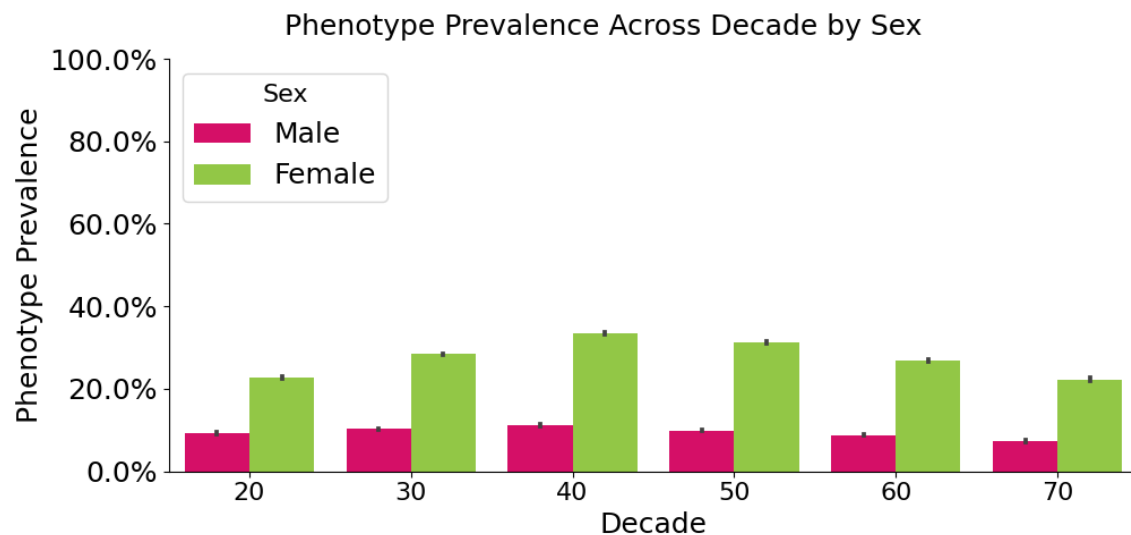


Table 5-1: Migraine participant cohort descriptives

Platform	Ancestry Group	Sample Use	N	Age mean (SD)	Sex (% female)	migraine prevalence (%)
V1 to V5	European	GWAS	1,954,622	50 (15.8)	56.68	22.08
V1 to V5	Sub-Saharan African/African American	GWAS	111,923	44 (14.9)	58.64	16.83
V1 to V5	East/Southeast Asian	GWAS	90,551	39 (13.3)	60.31	9.1
V1 to V5	Hispanic/Latino	GWAS	339,762	41 (14.3)	57.86	18.38
V5	European	Training multi-ancestral Model	1473761	46 (15.74)	58.32%	20.34%
	Sub-Saharan African/African American					
	East/Southeast Asian					
	Hispanic/Latino					
	European	Testing	272,167	48 (15.77)	58.24%	21.59%
	Hispanic/Latino	Testing	42,911	40 (13.71)	58.61%	17.41%
	Sub-Saharan African/African American	Testing	14,982	41 (14.44)	58.38%	16.17%
	East/Southeast Asian	Testing	12,909	37 (12.94)	60.92%	8.35%
	South Asian	Testing	11,729	39 (12.43)	42.77%	10.07%
	Northern African/Western Asian	Testing	10,497	42 (14.86)	44.99%	10.44%

Table 5-2: Migraine model performance characteristics

Ancestry Group (test sets)	Full Model AUROC (95%CIs)	Genetics Only AUROC (95%CIs)	Odds Ratio top 5% versus average (95%CIs)	Odds Ratio top 5% versus bottom 5% (95%CIs)
European	0.7009 (0.6987 to 0.7030)	0.6233 (0.6209 to 0.6258)	2.47 (2.37 to 2.57)	6.15 (5.75 to 6.57)
Hispanic/Latino	0.6831 (0.6771 to 0.6891)	0.6168 (0.6102 to 0.6234)	2.37 (2.13 to 2.64)	5.47 (4.56 to 6.56)
Sub-Saharan African/African American	0.6557 (0.6448 to 0.6666)	0.5736 (0.5616 to 0.5856)	1.81 (1.50 to 2.20)	3.35 (2.49 to 4.51)
East/Southeast Asian	0.6504 (0.6350 to 0.6659)	0.5973 (0.5808 to 0.6139)	2.25 (1.75 to 2.90)	5.57 (3.47 to 8.94)
South Asian	0.6882 (0.6741 to 0.7022)	0.5986 (0.5826 to 0.6147)	1.82 (1.41 to 2.37)	4.90 (3.03 to 7.91)
Northern African/Western Asian	0.6871 (0.6725 to 0.7018)	0.6161 (0.5998 to 0.6325)	2.35 (1.82 to 3.04)	6.07 (3.74 to 9.87)

Table 5-3: Migraine qualitative result characteristics

Ancestry group (test sets)	Odds Ratio for Result Threshold	Percent Above Threshold	Likelihood Ratio of "Increased" Result (95% CIs)
European	1.5	18.40%	1.81 (1.78 to 1.84)
Hispanic/Latino	1.5	17.20%	1.83 (1.75 to 1.91)
Sub-Saharan African/African American	1.5	6.37%	1.68 (1.46 to 1.94)
East/Southeast Asian	1.5	12.85%	1.71 (1.51 to 1.94)
South Asian	1.5	12.71%	1.63 (1.47 to 1.85)
Northern African/Western Asian	1.5	16.94%	1.78 (1.60 to 1.98)

Baseline prevalences used for likelihood estimate

Ancestry- and sex-specific baseline prevalences of ever been diagnosed with or treated for migraine were derived from the 23andMe database because no large-scale nationally representative datasets were available (Table 5-2). Estimates were derived from consented 23andMe customers between the ages of 40 to 49. We selected this specific age range because it is around the peak prevalence of the condition, in each sex and ancestry group. This age range of peak prevalence is similar to a previously published cross-sectional study of migraine prevalence (Lipton et al., 2007).

Table 5-2: Migraine baseline prevalences in 23andMe data between ages 40 to 49.

23andMe Population	Sex	N	Prevalence	95% CI
European	Male	204960	12.26%	12.12% to 12.40%
	Female	262148	35.48%	35.30% to 35.66%
Hispanic/Latino	Male	34965	9.45%	9.14% to 9.76%
	Female	43964	29.27%	28.84% to 29.70%
Sub-Saharan African/African American	Male	12551	9.09%	8.50% to 9.59%
	Female	16678	28.55%	27.86% to 29.24%
East/Southeast Asian	Male	9323	5.76%	5.29% to 6.23%
	Female	13556	14.98%	14.38% to 15.58%
South Asian	Male	2926	5.74%	4.9% to 6.58%
	Female	1886	17.44%	15.73% to 19.15%
Northern African/Western Asian	Male	2102	5.47%	4.5% to 6.44%
	Female	1665	19.76 %	17.85% to 21.67%

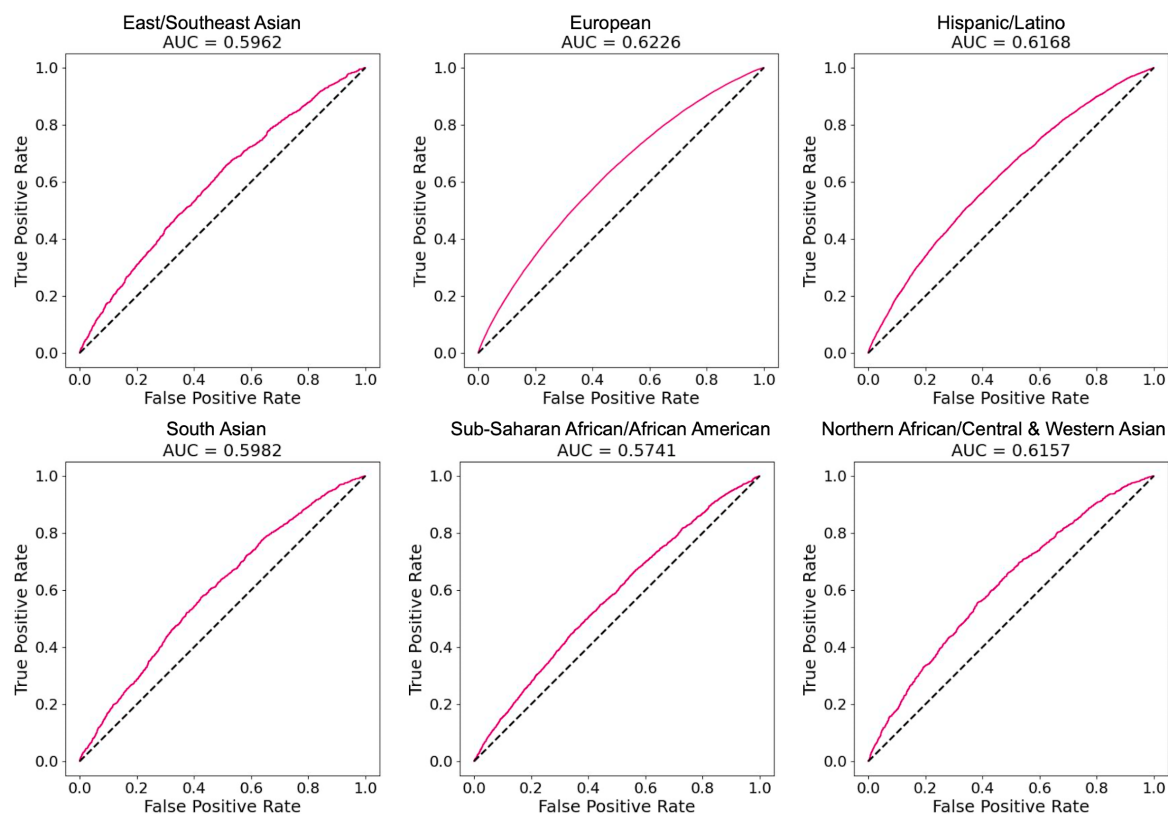
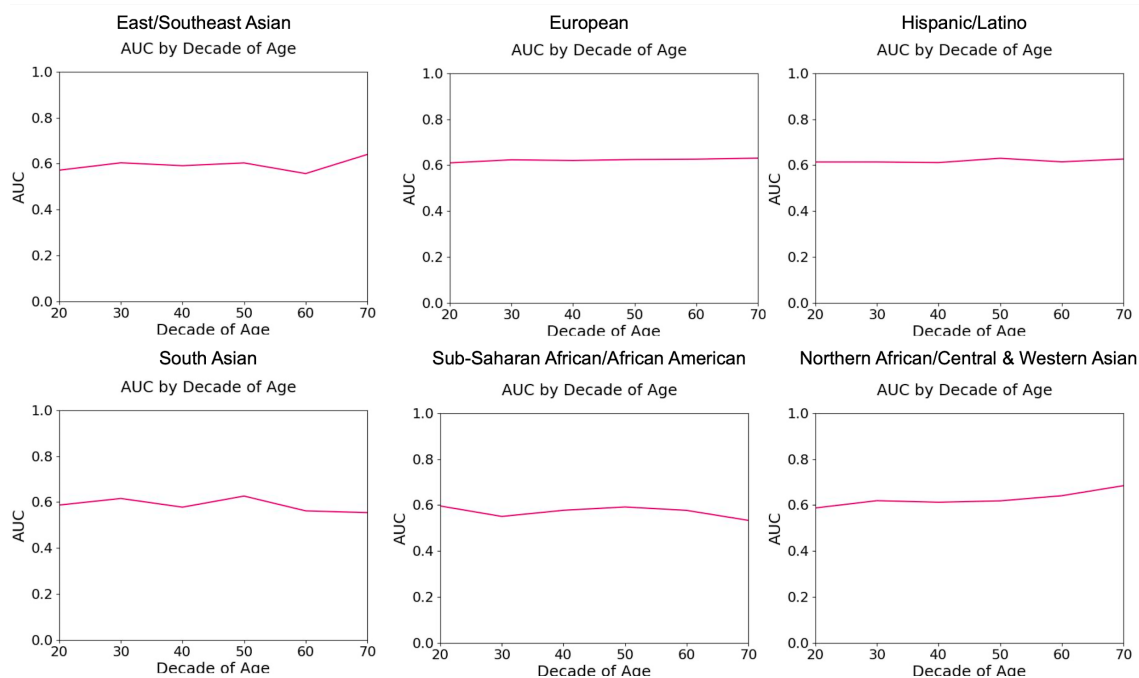
Figure 5-3: Migraine AUROC across ancestry-specific test sets**Figure 5-4: Migraine AUROC within each decade of age across ancestry-specific test sets**

Figure 5-5: Migraine case/control standardized PGS distributions across ancestry-specific test sets

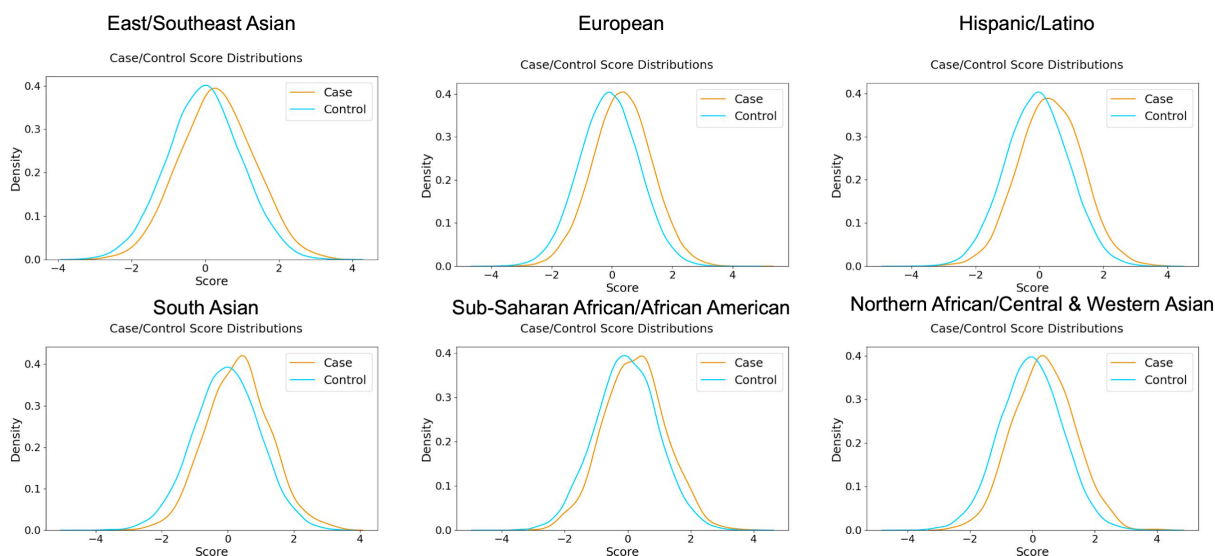
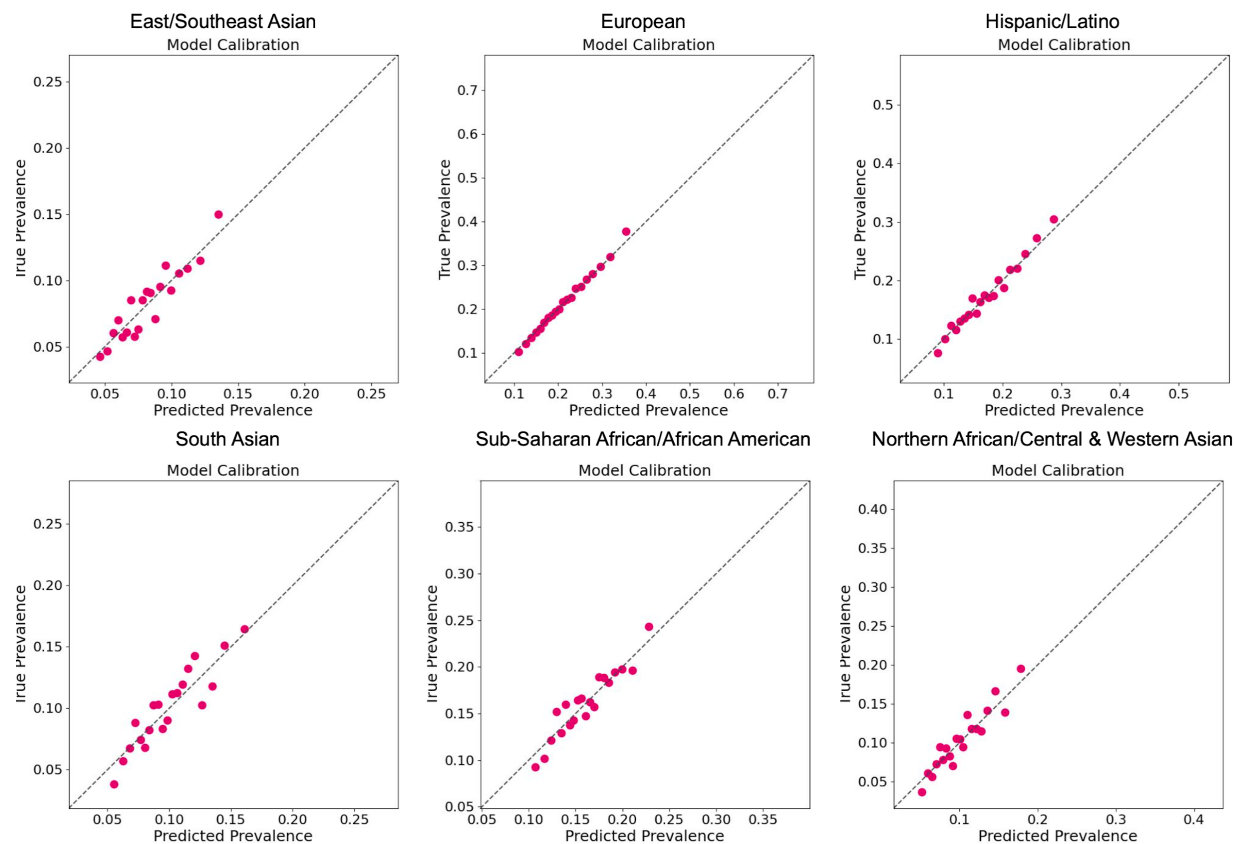


Figure 5-6: Migraine Platt-scaled calibration plots across ancestry-specific test sets



6. Uterine Fibroids

Olga Sazonova, PhD & Jianan Zhan, PhD, lead analysts. Report added August, 2020

Phenotype definition

Cases and controls were defined based on self-reported responses to questions about current or past diagnosis or treatment for uterine fibroids (UF) among female research participants. All question phrasings focused on receiving a diagnosis of UF, and some included the concept of treatment for UF. Although some phrasings had minor differences, the general phrasing was: “Have you ever been diagnosed with, or treated for, uterine fibroids or leiomyomas?” Participants answering affirmatively were treated as cases, while those who report no history of diagnosis or treatment for UF were counted as controls.

Model description

For the UF report there are two models. The first was trained on a European ancestry cohort based on results of a European ancestry GWAS, with genetic features optimized against a European ancestry validation set. This model is provided to all customers except those identifying as Sub-Saharan African/African American. In order to create a better-performing UF model for this group, we performed a three way fixed effect meta-analysis of European, Sub-Saharan African/African American, and Hispanic/Latino GWAS. Individuals from these three cohorts were combined into a large multi-ancestral training set. Multiple variant sets (based on variant selection hyper-parameter permutations) were used to train candidate models in this mixed cohort, with performance optimized in a Sub-Saharan African/African American validation cohort. The training and validation cohorts were then combined to train the final model, for which model performance metrics were measured in an out-of-sample Sub-Saharan African/African American test cohort.

Thus, the final two models selected for UF were optimized for A) the European ancestry group (2,898 genetic variants) and B) the Sub-Saharan African/African American ancestry group (1,834 genetic variants). Both models had age and age² terms, but only the European model included the first five principal components (PCs) derived from the European ancestry population. Genomic PCs are not used when determining an individual customer’s report result.

Table 6-0: UF model and variant set descriptions

Optimized Ancestry Group	Applicable Ancestry Group(s)	No. of variants
European	European, South Asian, Northern African/Western Asian, Hispanic/Latino, East/Southeast Asian	2,898
Sub-Saharan African/African American	Sub-Saharan African/African American	1,834

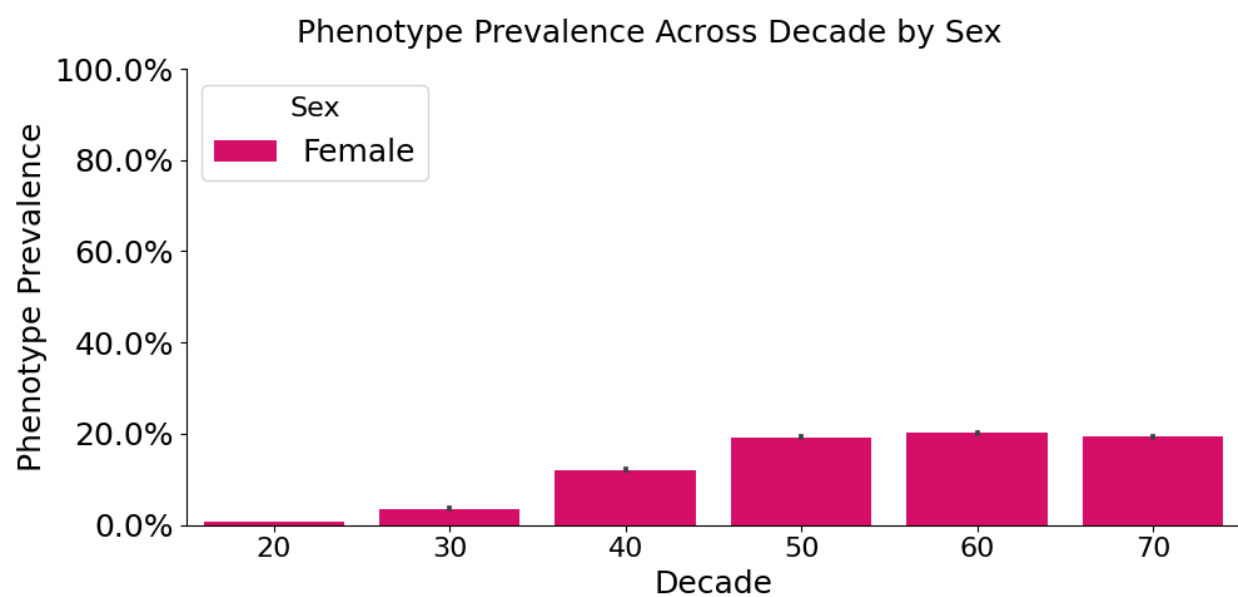
Figure 6-1: UF phenotype distribution in the European ancestry training set

Table 6-1: UF participant cohort descriptives

Platform	Ancestry Group(s)	Sample Use	N	Age mean (SD)	Sex (% female)	UF prevalence (%)
V1 to V5	Sub-Saharan African/African American	GWAS	57,524	42.6 (14.9)	100.00%	20.56%
V1 to V5	Hispanic/Latino	GWAS	192,241	40.6 (14.4)	100.00%	9.40%
V1 to V5	European	GWAS	1,004,539	49.7 (15.8)	100.00%	14.46%
V5	European	Training the European Model	828,393	48.4 (15.7)	100.00%	12.32%
		Testing	206,951	48.4 (15.7)	100.00%	12.30%
	Sub-Saharan African/African American & European & Hispanic/Latino	Training the African American Model	1,270,675	46.9 (15.7)	100.00%	11.96%
	Sub-Saharan African/African American	Testing	5,091	42.2 (14.7)	100.00%	18.07%
	East/Southeast Asian	Testing	46,945	37.5 (13.0)	100.00%	4.89%
	Hispanic/Latino	Testing	18,947	40.1 (13.9)	100.00%	8.16%
	South Asian	Testing	8,208	38.2 (12.3)	100.00%	5.93%
	Northern African/Western Asian	Testing	11,450	41.6 (14.8)	100.00%	6.04%

Table 6-2: UF model performance characteristics

Ancestry Group (test sets)	Full Model AUROC (95%CIs)	Genetics Only AUROC (95%CIs)	Odds Ratio top 5% versus average (95%CIs)	Odds Ratio top 5% versus bottom 5% (95%CIs)
European	0.7356 (0.7327 to 0.7384)	0.5933 (0.5898 to 0.5968)	2.12 (2.01 to 2.24)	4.12 (3.76 to 4.52)
Sub-Saharan African/African American	0.7778 (0.7636 to 0.792)	0.5854 (0.5659 to 0.6049)	1.45 (1.05 to 2.01)	4.02 (2.35 to 6.87)
East/Southeast Asian	0.8268 (0.8204 to 0.8332)	0.5916 (0.5804 to 0.6028)	1.7 (1.43 to 2.02)	3.66 (2.71 to 4.95)
Hispanic/Latino	0.7887 (0.7793 to 0.7981)	0.59 (0.5761 to 0.604)	1.85 (1.48 to 2.31)	2.53 (1.82 to 3.53)
South Asian	0.7929 (0.7768 to 0.809)	0.5933 (0.5689 to 0.6177)	1.96 (1.34 to 2.88)	3.78 (1.96 to 7.3)
Northern African/Western Asian	0.7678 (0.7532 to 0.7825)	0.5836 (0.5629 to 0.6043)	1.46 (1.05 to 2.02)	3.71 (2.06 to 6.68)

Table 6-3: UF qualitative result characteristics

Ancestry Group (test sets)	Odds Ratio for Result Threshold	Percent Above Threshold	Likelihood Ratio of "Increased" Result (95% CIs)
European	1.5	11.56%	1.75 (1.70 to 1.80)
Sub-Saharan African/African American	1.5	8.80%	1.37 (1.11 to 1.69)
East/Southeast Asian	1.5	10.37%	1.65 (1.50 to 1.82)
Hispanic/Latino	1.5	9.12%	1.61 (1.41 to 1.84)
South Asian	1.5	11.23%	1.57 (1.28 to 1.93)
Northern African/Western Asian	1.5	8.75%	1.55 (1.27 to 1.90)

Baseline prevalences used for likelihood estimate

Ancestry- and sex-specific baseline prevalences of ever been diagnosed with or treated for UF were derived from the 23andMe database because no large-scale nationally representative datasets were available (Table 6-2). Estimates were derived from consented 23andMe customers between the ages of 55 to 64. We selected this specific age range because it is around the observed peak prevalence of the condition in each sex and ancestry group and represents the end of childbearing years.

Table 6-2: UF baseline prevalences in 23andMe data between ages 55 to 64.

UF				
23andMe Population	Sex	N	Prevalence	95% CI
Sub-Saharan African/African American	Female	14,502	34.77%	0.78%
East/Southeast Asian	Female	6,717	15.24%	0.86%
European	Female	424,498	17.94%	0.12%
Hispanic/Latino	Female	43,869	17.08%	0.35%
Northern African/Western Asian	Female	1,983	13.72%	1.51%
South Asian	Female	1,185	16.37%	2.11%

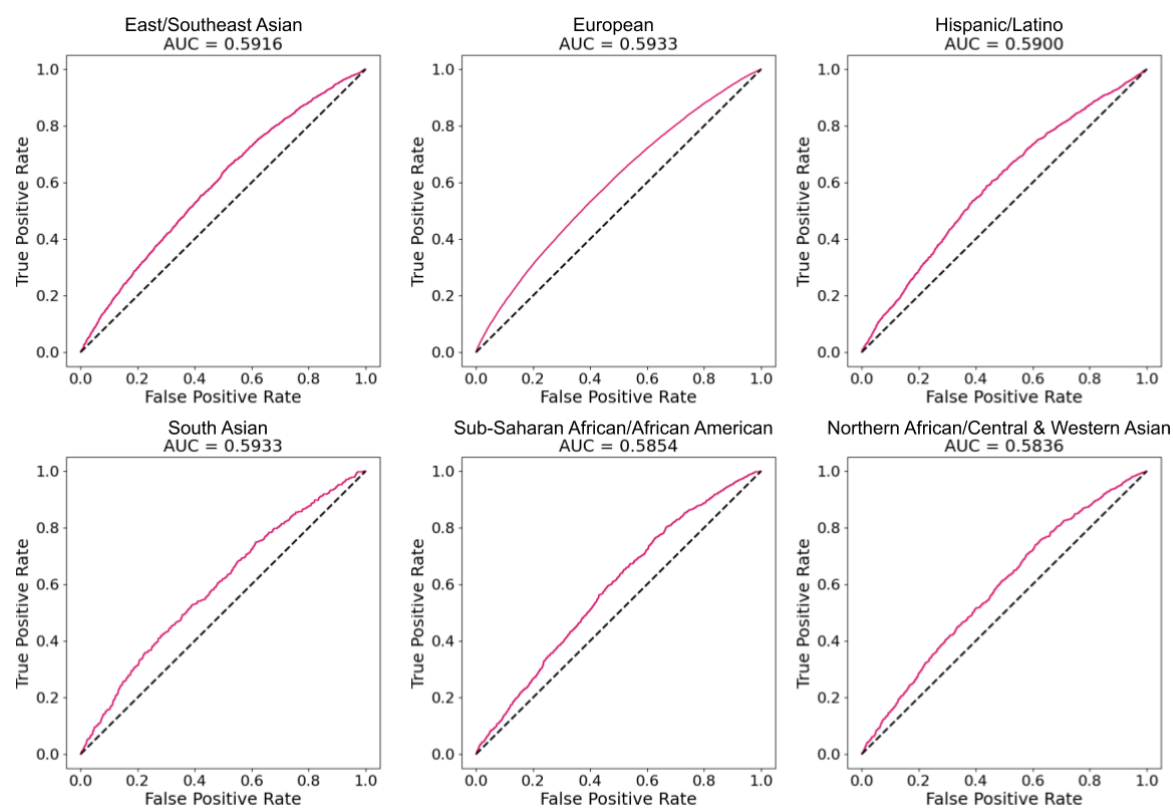
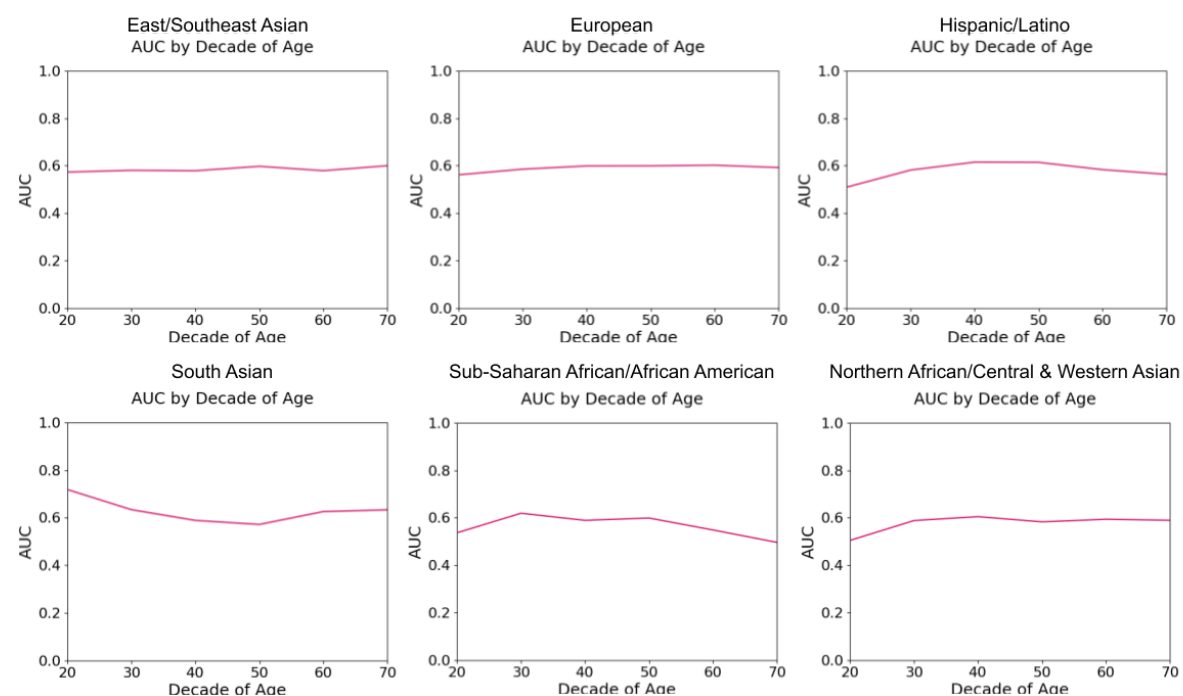
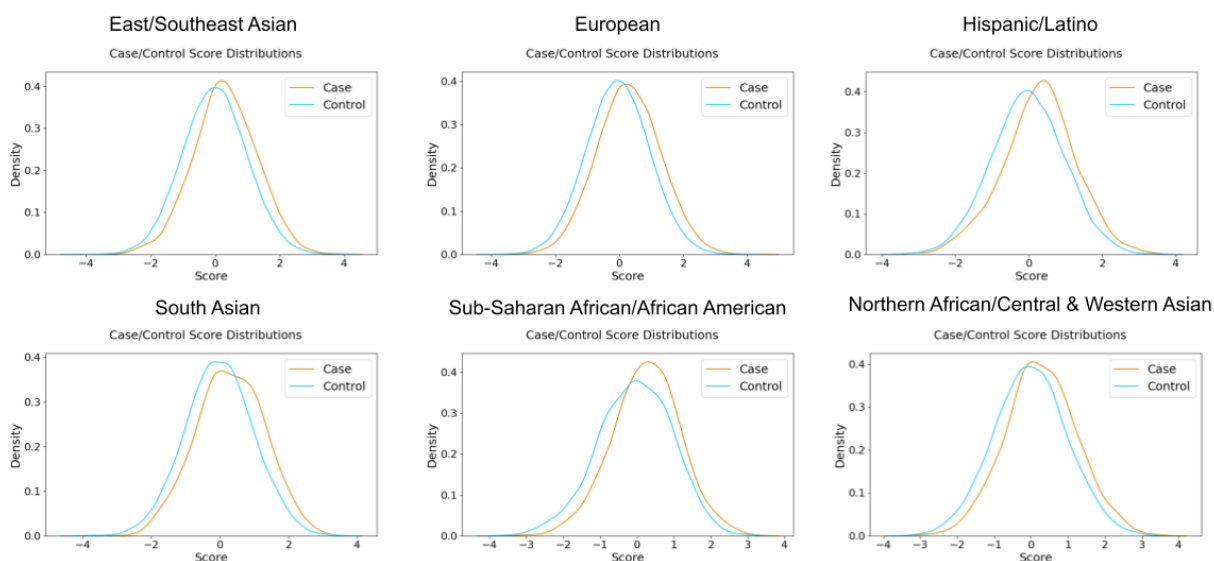
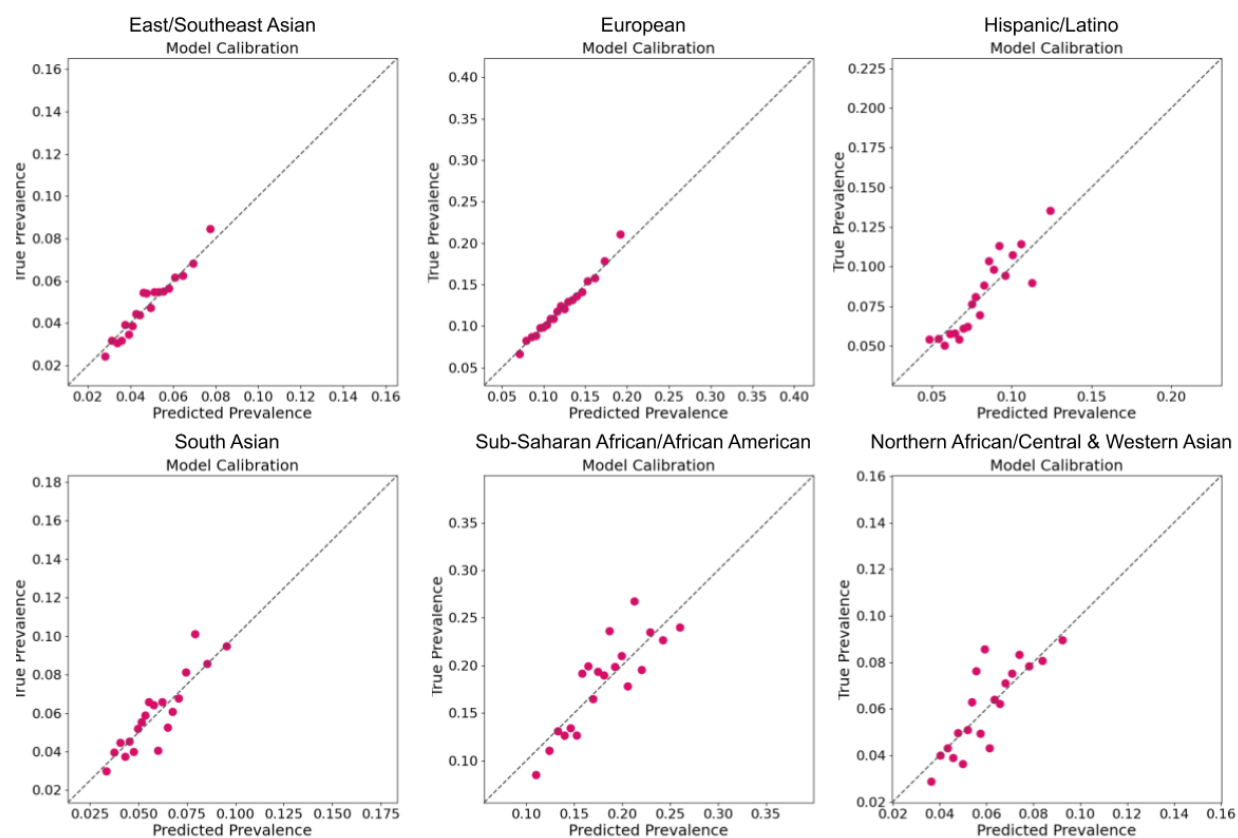
Figure 6-3: UF AUROC across ancestry-specific test sets**Figure 6-4: UF AUROC within each decade of age across ancestry-specific test sets**

Figure 6-5: UF case/control standardized PGS distributions across ancestry-specific test sets**Figure 6-6:** UF Platt-scaled calibration plots across ancestry-specific test sets

7. Obstructive Sleep Apnea

Peter Chisnell, PhD, lead analyst. Report added September, 2020.

Phenotype definition

Cases and controls were defined based on self-reported responses to questions about current or past diagnosis of sleep apnea. All question phrasings focused on receiving a diagnosis of or treatment for sleep apnea. Although some phrasings had minor differences, the general phrasing was: “Have you ever been diagnosed with or treated for sleep apnea?” Participants answering affirmatively were treated as cases, while those who report no history of sleep apnea diagnosis were counted as controls.

Model description

For the development of this report, the data were sufficient to attempt ancestry-specific GWAS. Four separate GWAS were conducted in East/Southeast Asian, European, Hispanic/Latino, and Sub-Saharan African/African American populations. A fixed effect meta-analysis was performed to combine the four ancestry-specific GWASs, followed by pruning and thresholding (P+T) using different hyperparameter combinations. Each of these candidate feature sets were tested on the ancestry-specific validation sets to evaluate performance in all groups. We selected a single SNP set including 23,021 variants. This SNP set was used to train one model in a multi-ancestral training cohort, which was evaluated in the ancestry-specific test sets (see Table 7-1). This model had age, sex, an age*sex interaction term, as well as the first ten global principal components (PCs) to account for population stratification. Genomic PCs are not used when determining an individual customer’s report result. Personalized binary qualitative results and estimated likelihoods are based on genetic ancestry and birth sex.

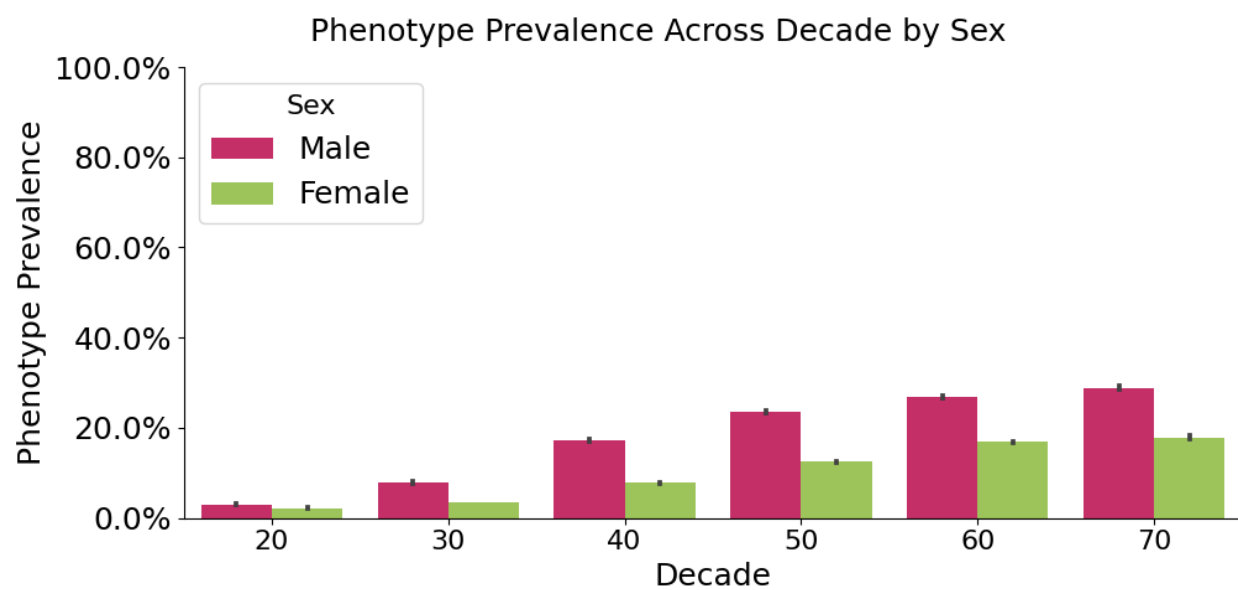
Figure 7-1: Sleep apnea phenotype distribution in the multi-ancestry training set

Table 7-1: Sleep apnea participant cohort descriptives

Platform	Ancestry Group	Sample Use	N	Age mean (SD)	Sex (% female)	Sleep apnea prevalence (%)
V1 to V5	European	GWAS	1,903,422	51 (15.89)*	55.4%	13.6%
V1 to V5	Sub-Saharan African/African American	GWAS	107,539	44 (15.00)*	58.4%	11.7%
V1 to V5	Hispanic/Latino	GWAS	331,370	41 (14.39)*	57.1%	9.3%
V1 to V5	East/Southeast Asian	GWAS	81,244	39 (13.34)*	60.7%	4.9%
V5	European	Training the mixed-ancestry Model	1,480,954	47 (15.85)	56.8%	12.33%
	Sub-Saharan African/African American					
	East/Southeast Asian					
	Hispanic/Latino					
	European	Testing	259,358	48 (15.87)	56.3%	13.24%
	Hispanic/Latino	Testing	40,002	40 (13.85)	57.3%	8.94%
	Sub-Saharan African/African American	Testing	13,892	42 (14.44)	58.3%	11.81%
	East/Southeast Asian	Testing	19,576	37 (12.96)	60.5%	4.97%
	South Asian	Testing	11,513	39 (12.70)	42.5%	5.39%
	Northern African/Western Asian	Testing	10,282	42 (14.75)	44.8%	5.79%

Table 7-2: Sleep apnea model performance characteristics

Ancestry Group (test sets)	Full Model AUROC (95%CIs)	Genetics Only AUROC (95%CIs)	Odds Ratio top 5% versus average (95%CIs)	Odds Ratio top 5% versus bottom 5% (95%CIs)
European	0.7395 (0.7370, 0.7419)	0.6116 (0.6086, 0.6146)	2.38 (2.27, 2.50)	5.36 (4.94, 5.83)
Hispanic/Latino	0.7723 (0.7657, 0.7788)	0.6116 (0.6026, 0.6206)	2.58 (2.25, 2.96)	4.66 (3.68, 5.90)
Sub-Saharan African/African American	0.7422 (0.7313, 0.7530)	0.5880 (0.5741, 0.6019)	1.97 (1.58, 2.46)	4.41 (3.00, 6.48)
East/Southeast Asian	0.7674 (0.7552, 0.7796)	0.5893 (0.5720, 0.6065)	1.87 (1.43, 2.44)	3.03 (1.97, 4.68)
South Asian	0.7493 (0.7332, 0.7655)	0.5764 (0.5545, 0.5984)	1.96 (1.38, 2.76)	3.19 (1.82, 5.60)
Northern African/Western Asian	0.7589 (0.7427, 0.7750)	0.5808 (0.5584, 0.6031)	2.01 (1.43, 2.83)	3.82 (2.13, 6.88)

Table 7-3: Sleep apnea qualitative result characteristics

Ancestry Group (test sets)	Odds Ratio for Result Threshold	Percent Above Threshold	Likelihood Ratio of "Increased" Result (95% CIs)
European	1.5	15.77%	1.76 (1.73, 1.80)
Hispanic/Latino	1.5	16.27%	1.81 (1.71, 1.92)
Sub-Saharan African/African American	1.5	9.73%	1.69 (1.48, 1.92)
East/Southeast Asian	1.5	9.79%	1.64 (1.40, 1.91)
South Asian	1.5	7.24%	1.76 (1.41, 2.20)
Northern African/Western Asian	1.5	7.24%	1.63 (1.32, 2.01)

Baseline prevalences used for likelihood estimate

Ancestry- and sex-specific baseline prevalences of ever been diagnosed with or treated for insomnia were derived from the 23andMe database because no large-scale nationally representative datasets were available (Table 7-4). Estimates were derived from consented 23andMe customers between the ages of 70 to 79.

Table 7-4: Sleep apnea baseline prevalences in 23andMe data between ages 70 to 79.

23andMe Population	Sex	N	Prevalence (%)	95% CI (%)
European	Male	1,651,845	14.98%	0.05%
	Female	2,276,620	7.83%	0.03%
Hispanic/Latino	Male	265,079	11.60%	0.12%
	Female	378,026	5.36%	0.07%
Sub-Saharan African/African American	Male	77,675	14.57%	0.25%
	Female	114,422	8.45%	0.16%
East/Southeast Asian	Male	54,984	8.17%	0.23%
	Female	84,863	2.56%	0.11%
South Asian	Male	21,023	7.38%	0.35%
	Female	14,991	2.6%	0.25%
Northern African/Western Asian	Male	18,706	8.4%	0.40%
	Female	15,216	2.9%	0.27%

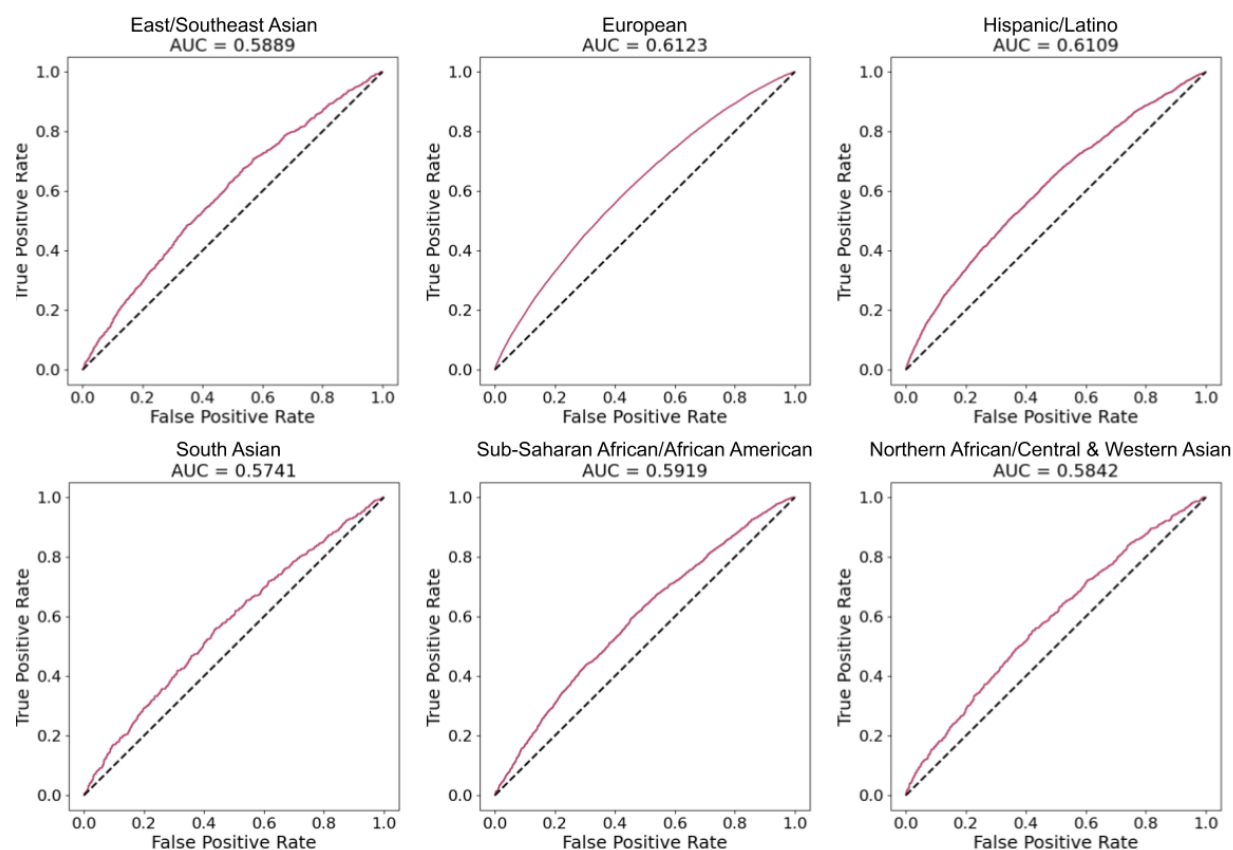
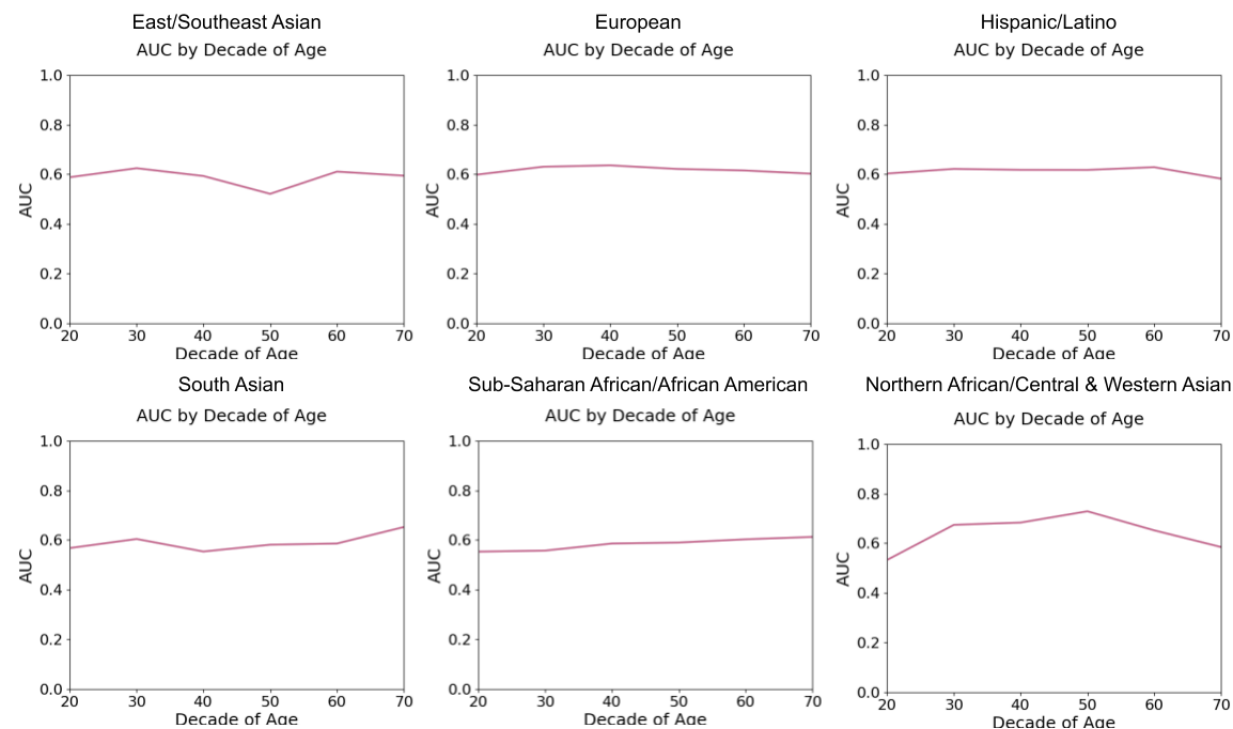
Figure 7-2: Sleep apnea AUROC across ancestry-specific test sets**Figure 7-3:** Sleep apnea AUROC within each decade of age across ancestry-specific test sets

Figure 7-4: Sleep apnea case/control standardized PGS distributions across ancestry-specific test sets

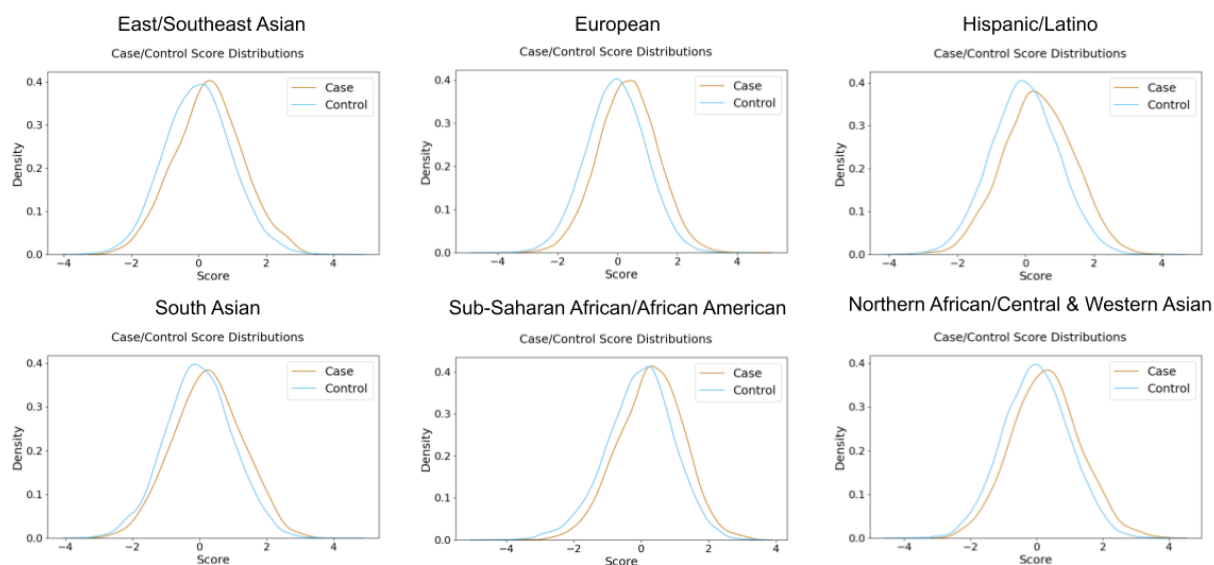
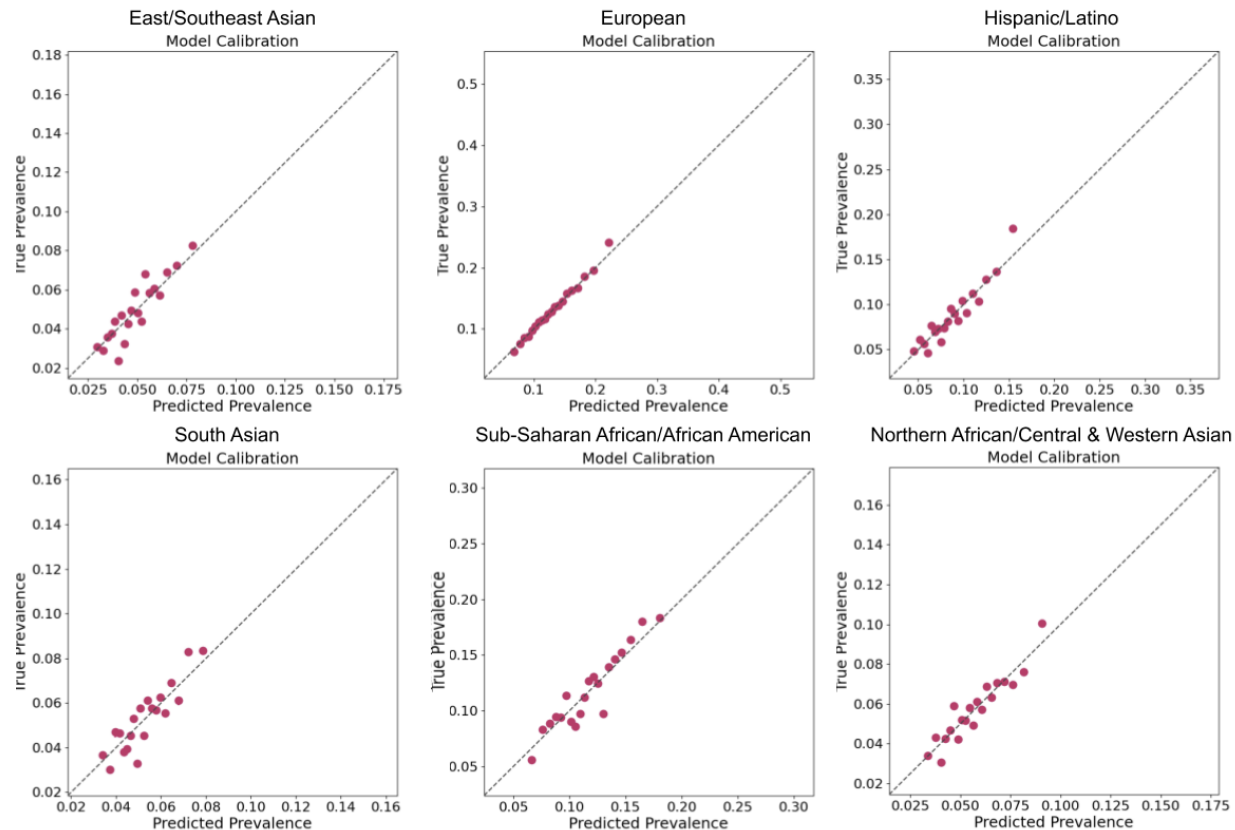


Figure 7-5: Sleep apnea Platt-scaled calibration plots across ancestry-specific test sets



8. Restless Legs Syndrome

Peter Chisnell, PhD, lead analyst. Report added October, 2020.

Phenotype definition

Cases and controls were defined based on self-reported responses to questions about current or past diagnosis of restless legs syndrome. All question phrasings focused on receiving a diagnosis of or treatment for restless legs syndrome. Although some phrasings had minor differences, the general phrasing was: “Have you ever been diagnosed with or treated for restless legs syndrome?” Participants answering affirmatively were treated as cases, while those who report no history of restless legs syndrome diagnosis were counted as controls.

Model description

For the development of this report, the data were sufficient to attempt ancestry-specific GWAS. Three separate GWAS were conducted in European, Hispanic/Latino, and Sub-Saharan African/African American populations. A fixed effect meta-analysis was performed to combine the three ancestry-specific GWASs, followed by pruning and thresholding (P+T) using different hyperparameter combinations. Each of these candidate feature sets were tested on the ancestry-specific validation sets to evaluate performance in all groups. We selected a single SNP set including 23,099 variants. This SNP set was used to train one model in a multi-ancestral training cohort, which was evaluated in the ancestry-specific test sets (see Table 7-1). This model had age, sex, an age*sex interaction term, as well as the first ten global principal components (PCs) to account for population stratification. Genomic PCs are not used when determining an individual customer’s report result. Personalized binary qualitative results and estimated likelihoods are based on genetic ancestry and birth sex.

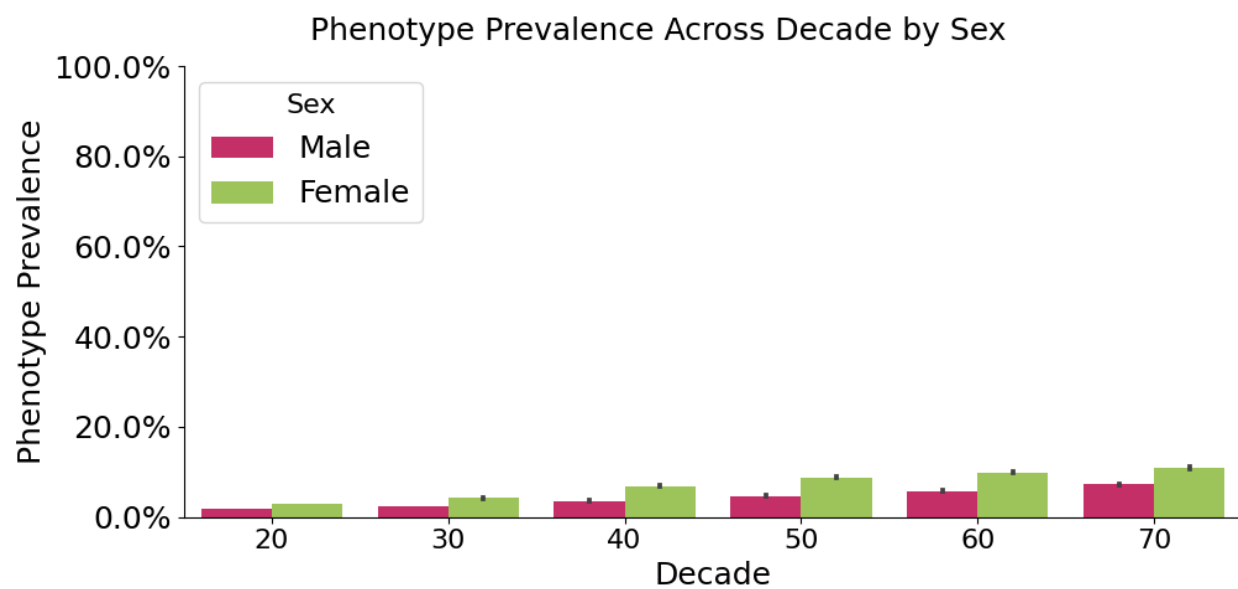
Figure 8-1: Restless legs syndrome phenotype distribution in the multi-ancestry training set

Table 8-1: Restless legs syndrome participant cohort descriptives

Platform	Ancestry Group	Sample Use	N	Age mean (SD)	Sex (% female)	RLS prevalence (%)
V1 to V5	European	GWAS	1,916,314	51 (15.89)	55.4	5.28
V1 to V5	Sub-Saharan African/African American	GWAS	107,949	44 (15.00)	58.4	2.37
V1 to V5	Hispanic/Latino	GWAS	332,489	41 (14.39)	57.1	3.07
V5	European	Training the mixed-ancestry Model	1,379,449	47 (15.83)	57.1	5.56
	Sub-Saharan African/African American					
	East/Southeast Asian					
	Hispanic/Latino					
	European	Testing	262,532	48 (15.84)	57.1	6.04
	Hispanic/Latino	Testing	40,520	40 (13.80)	57.9	3.21
	Sub-Saharan African/African American	Testing	37,212	42 (14.47)	58.0	2.38
	East/Southeast Asian	Testing	46,436	37 (13.02)	60.7	1.03
	South Asian	Testing	2,2986	39 (12.62)	42.8	1.64
	Northern African/Western Asian	Testing	20,516	42 (14.74)	44.6	2.04

Table 8-2: Restless legs syndrome model performance characteristics

Ancestry Group (test sets)	Full Model AUROC (95%CIs)	Genetics Only AUROC (95%CIs)	Odds Ratio top 5% versus average (95%CIs)	Odds Ratio top 5% versus bottom 5% (95%CIs)
European	0.7050 (0.7014, 0.7085)	0.6609 (0.6570, 0.6648)	3.44 (3.24, 3.65)	11.31 (9.82, 13.04)
Hispanic/Latino	0.7068 (0.6947, 0.7189)	0.6620 (0.6488, 0.6752)	3.71 (3.05, 4.53)	14.35 (8.46, 24.35)
Sub-Saharan African/African American	0.6517 (0.6355, 0.6679)	0.5801 (0.5622, 0.5980)	2.18 (1.67, 2.84)	2.99 (1.97, 4.55)
East/Southeast Asian	0.6184 (0.5955, 0.6413)	0.5938 (0.5701, 0.6175)	2.33 (1.61, 3.35)	4.25 (2.19, 8.22)
South Asian	0.6426 (0.6176, 0.6677)	0.6136 (0.5875, 0.6398)	1.97 (1.29, 3.00)	3.06 (1.54, 6.08)
Northern African/Western Asian	0.6610 (0.6378, 0.6841)	0.6212 (0.5965, 0.6459)	2.31 (1.59, 3.35)	5.84 (2.74, 12.44)

Table 8-3: Restless legs syndrome qualitative result characteristics

Ancestry Group (test sets)	Odds Ratio for Result Threshold	Percent Above Threshold	Likelihood Ratio of "Increased" Result (95% CIs)
European	1.5	24.49%	1.90 (1.86, 1.93)
Hispanic/Latino	1.5	25.27%	1.83 (1.72, 1.95)
Sub-Saharan African/African American	1.5	8.64%	1.75 (1.49, 2.06)
East/Southeast Asian	1.5	11.56%	1.93 (1.63, 2.29)
South Asian	1.5	14.69%	1.80 (1.52, 2.14)
Northern African/Western Asian	1.5	17.53%	1.88 (1.63, 2.17)

Baseline prevalences used for likelihood estimate

Ancestry- and sex-specific baseline prevalences of ever been diagnosed with or treated for restless legs syndrome were derived from the 23andMe database because no large-scale nationally representative datasets were available (Table 8-4). Estimates were derived from consented 23andMe customers between the ages of 70 to 79.

Table 8-4: Restless legs syndrome baseline prevalences in 23andMe data between ages 70 to 79.

23andMe Population	Sex	N	Prevalence (%)	95% CI (%)
European	Male	237,626	6.01	0.10
	Female	277,045	8.68	0.10
Hispanic/Latino	Male	13,891	4.89	0.36
	Female	17,051	6.42	0.37
Sub-Saharan African/African American	Male	5,266	3.30	0.48
	Female	6,463	4.70	0.52
East/Southeast Asian	Male	2,272	2.11	0.59
	Female	2,652	2.34	0.58
South Asian	Male	824	3.52	1.26
	Female	398	2.01	1.38
Northern African/Western Asian	Male	1,361	3.82	1.02
	Female	887	4.74	1.40

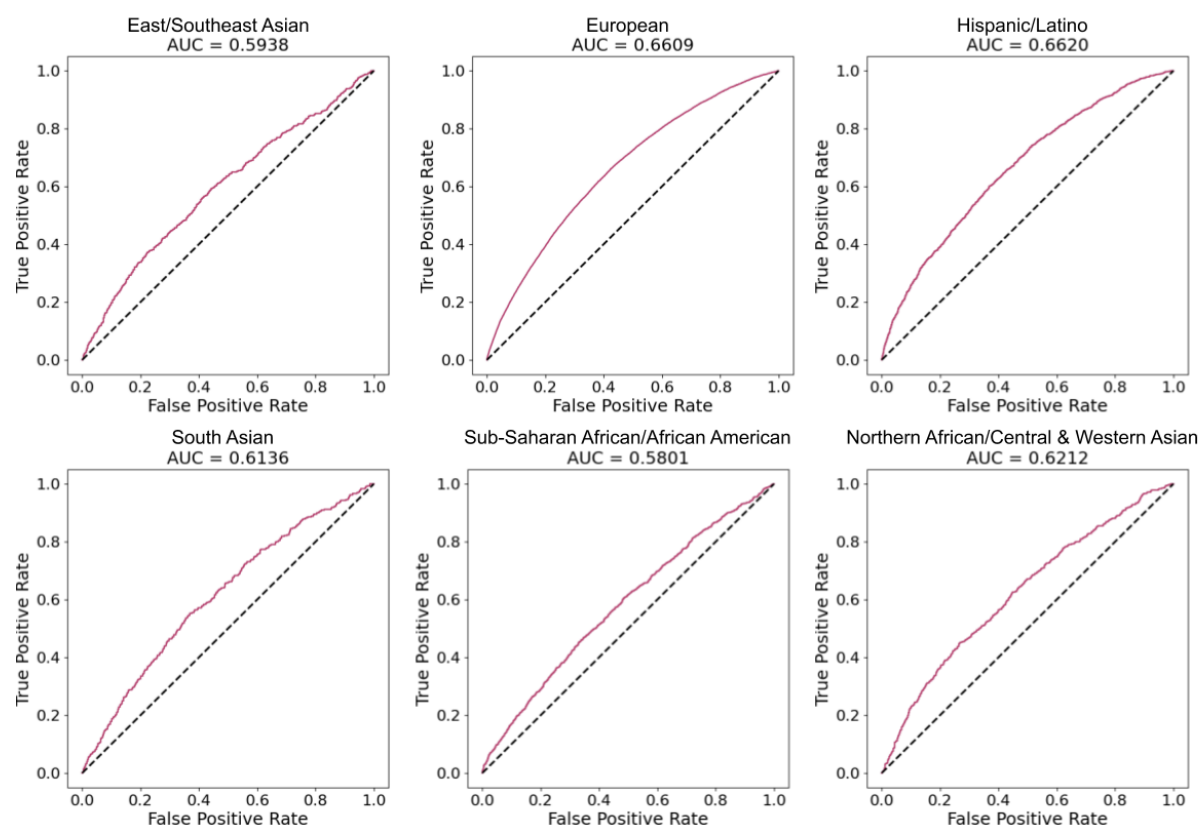
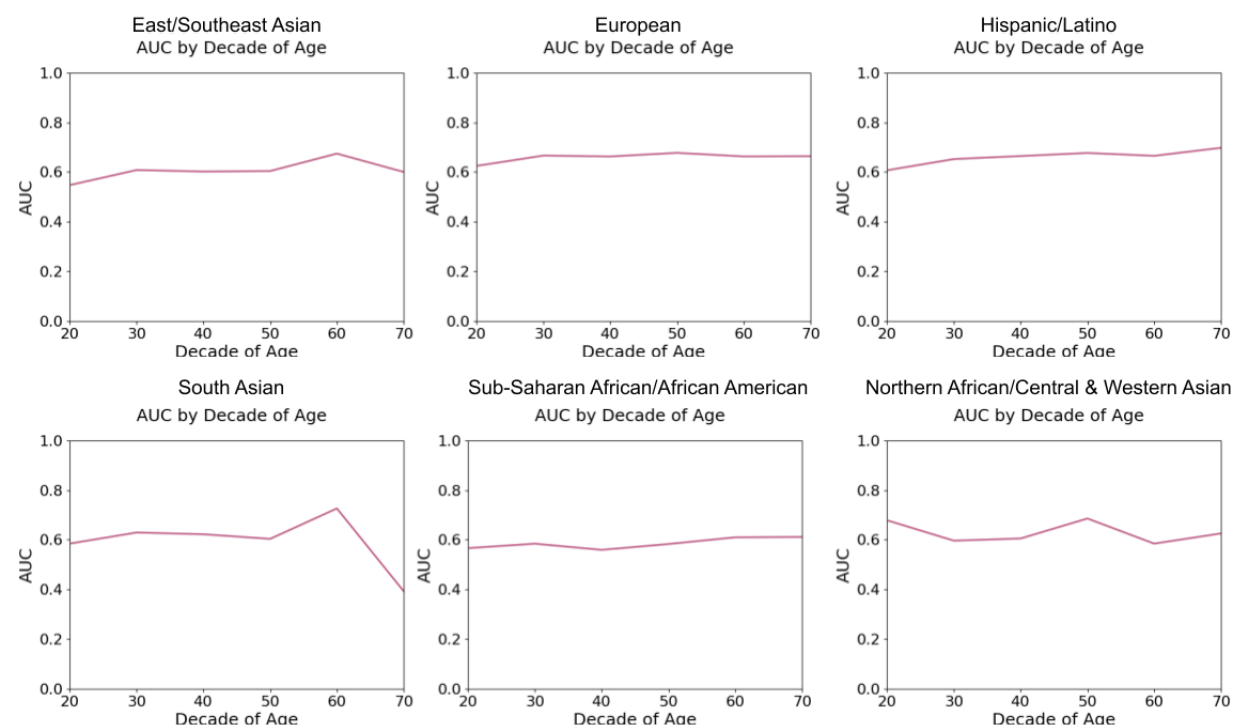
Figure 8-2: Restless legs syndrome AUROC across ancestry-specific test sets**Figure 8-3:** Restless legs syndrome AUROC within each decade of age across ancestry-specific test sets

Figure 8-4: Restless legs syndrome case/control standardized PGS distributions across ancestry-specific test sets

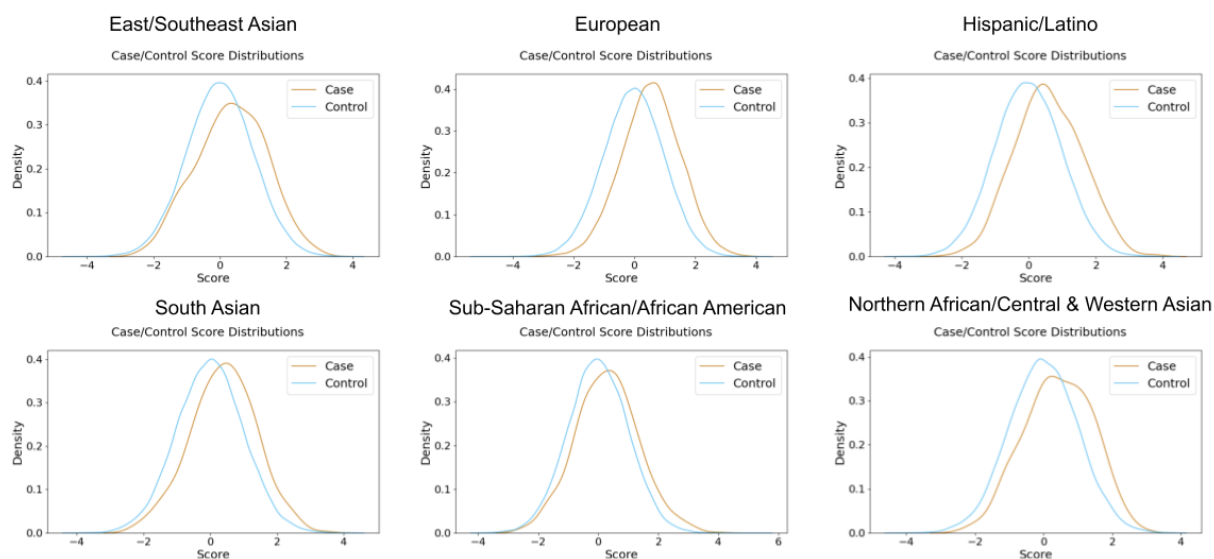
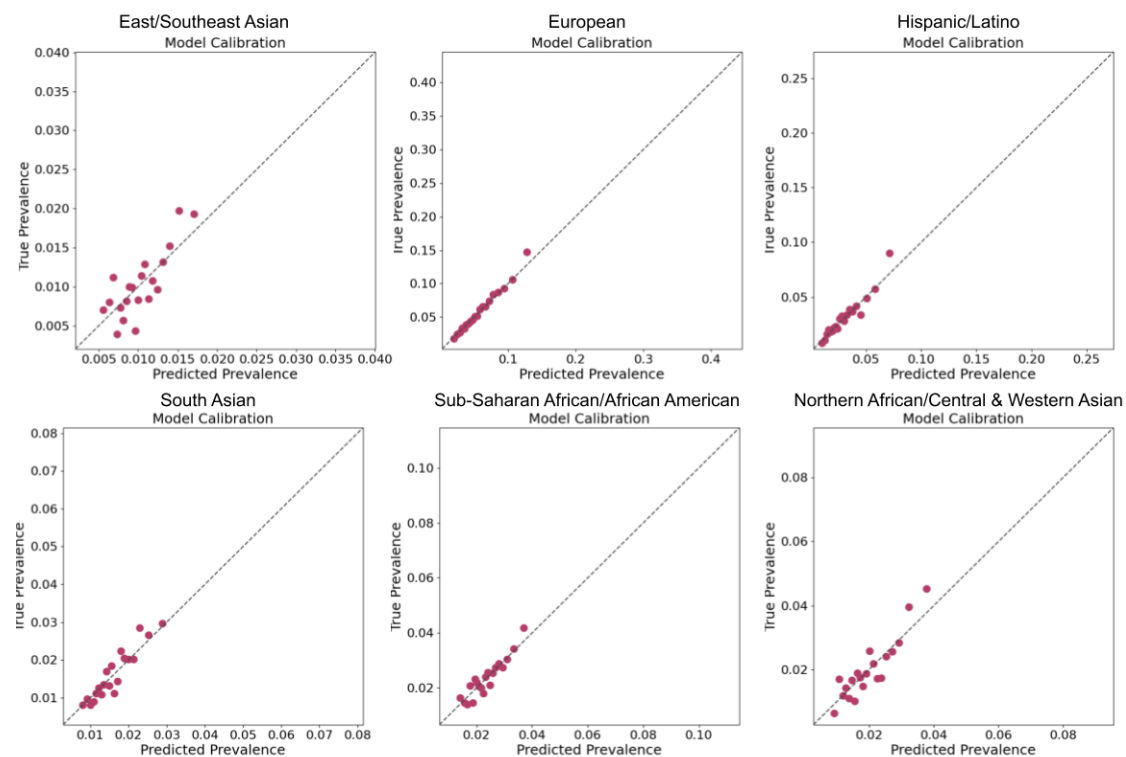


Figure 8-5: Restless leg syndrome Platt-scaled calibration plots across ancestry-specific test sets



9. Gout

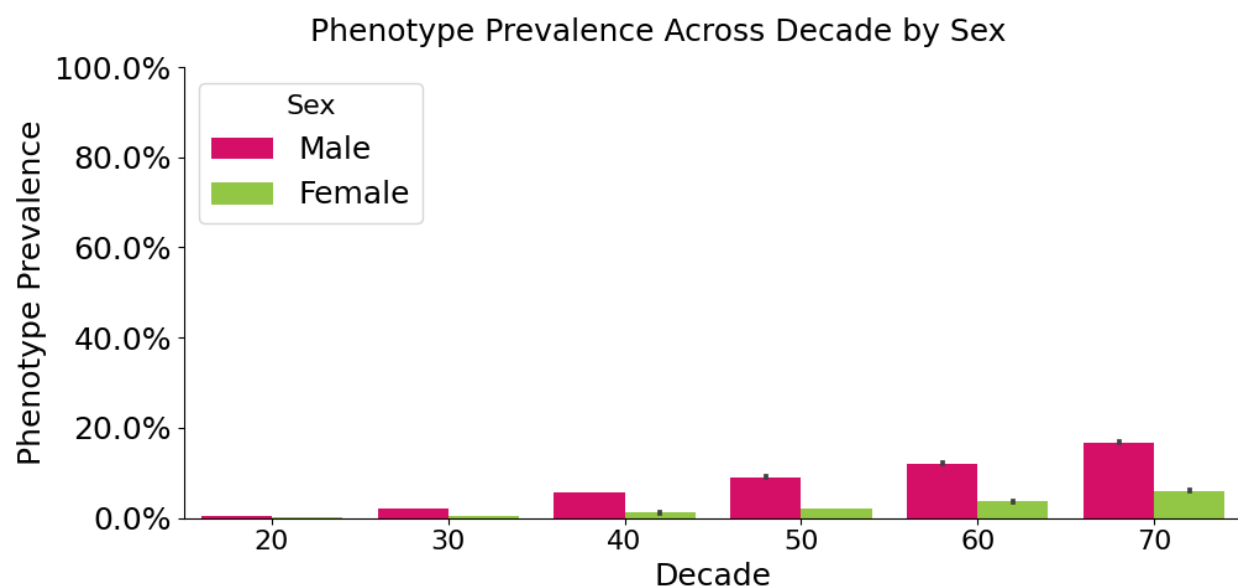
Rafaela Bagur Quetglas, PhD, lead analyst. Report added November, 2020.

Phenotype definition

Cases and controls were defined based on self-reported responses to questions about current or past diagnosis of gout. All question phrasings focused on receiving a diagnosis of or treatment for gout. Although some phrasings had minor differences, the general phrasing was: “Have you ever been diagnosed with or treated for gout?” Participants answering affirmatively were treated as cases, while those who report no history of gout diagnosis were counted as controls.

Model description

For the development of this report, the data were insufficient to attempt ancestry-specific GWAS. Therefore, one unique GWAS was conducted in European ancestry cohort. The model was trained on a European ancestry cohort based on results of a European ancestry GWAS, with genetic features optimized against a European ancestry validation set. This model is provided to all customers and was evaluated in all the ancestry-specific test sets (see Table 9-1). The gout model includes 21,381 genetic variants, age, sex, an age*sex interaction term, as well as the first ten global principal components (PCs) to account for population stratification. Genomic PCs are not used when determining an individual customer’s report result. Personalized binary qualitative results and estimated likelihoods are based on genetic ancestry and birth sex.

Figure 9-1: Gout phenotype distribution in the european training set**Table 9-1:** Gout participant cohort descriptives

Platform	Ancestry Group	Sample Use	N	Age mean (SD)	Sex (% female)	Gout prevalence (%)
V1 to V5	European	GWAS	2,615,838	49 (16.0)	55.81%	4.70%
V5	European	Training the European Model	2,137,145	47 (15.9)	56.79%	4.18%
	European	Testing	107,723	48 (15.9)	56.75%	4.31%
	Hispanic/Latino	Testing	184,039	40 (13.9)	57.61%	2.15%
	Sub-Saharan African/African American	Testing	115,108	41 (14.5)	58.09%	3.19%
	East/Southeast Asian	Testing	96,483	37 (12.9)	60.62%	2.96%
	South Asian	Testing	23,949	39 (12.6)	42.56%	1.89%
	Northern African/Western Asian	Testing	21,217	42 (14.6)	44.37%	2.33%

Table 9-2: Gout model performance characteristics

Ancestry Group (test sets)	Full Model AUROC	Genetics Only AUROC	Odds Ratio top 5% versus average (95% CIs)	Odds Ratio top 5% versus bottom 5% (95% CIs)
European	0.8168	0.6749	3.98 (3.57, 4.43)	12.37 (9.51, 16.09)
Hispanic/Latino	0.8402	0.6762	3.71 (3.32, 4.15)	13.28 (9.88, 17.84)
Sub-Saharan African/African American	0.8099	0.5967	2.07 (1.81, 2.36)	4.55 (3.56, 5.83)
East/Southeast Asian	0.8636	0.6745	3.57 (3.12, 4.08)	18.92 (12.52, 28.61)
South Asian	0.8341	0.6833	4.11 (2.90, 5.82)	7.58 (3.76, 15.29)
Northern African/Western Asian	0.8498	0.6790	4.34 (3.11, 6.06)	15.60 (6.28, 38.78)

Table 9-3: Gout qualitative result characteristics

Ancestry Group (test sets)	Odds Ratio for Result Threshold	Percent Above Threshold	Likelihood Ratio of "Increased" Result (95% CIs)
European	1.5	26.49%	1.97 (1.91, 2.03)
Hispanic/Latino	1.5	26.21%	1.92 (1.86, 1.99)
Sub-Saharan African/African American	1.5	12.09%	1.66 (1.55, 1.78)
East/Southeast Asian	1.5	26.25%	1.92 (1.85, 2.00)
South Asian	1.5	27.11%	2.06 (1.89, 2.25)
Northern African/Western Asian	1.5	26.87%	2.02 (1.85, 2.20)

Baseline prevalences used for likelihood estimate

Ancestry- and sex-specific baseline prevalences of ever been diagnosed with or treated for gout were derived from the 23andMe database because no large-scale nationally representative datasets were available (Table 9-4). Estimates were derived from consented 23andMe customers between the ages of 70 to 79. We selected this specific age range because it is around the peak prevalence of the condition, in each sex and ancestry group.

Table 9-4: Gout baseline prevalences in 23andMe data between ages 70 to 79.

23andMe Population	Sex	N	Prevalence (%)	95% CI (%)
European	Male	218,962	13.1%	0.14
	Female	297,110	4.3%	0.08
Hispanic/Latino	Male	24,521	11.3%	0.55
	Female	33,702	3.8%	0.30
Sub-Saharan African/African American	Male	10,401	17.3%	1.08
	Female	13,955	7.5%	0.68
East/Southeast Asian	Male	5,172	19.4%	1.70
	Female	7,654	4.7%	0.84
South Asian	Male	1,593	10.5%	2.22
	Female	1,056	3.7%	1.99
Northern African/Western Asian	Male	1,567	10.8%	1.71
	Female	1,260	3.3%	1.21

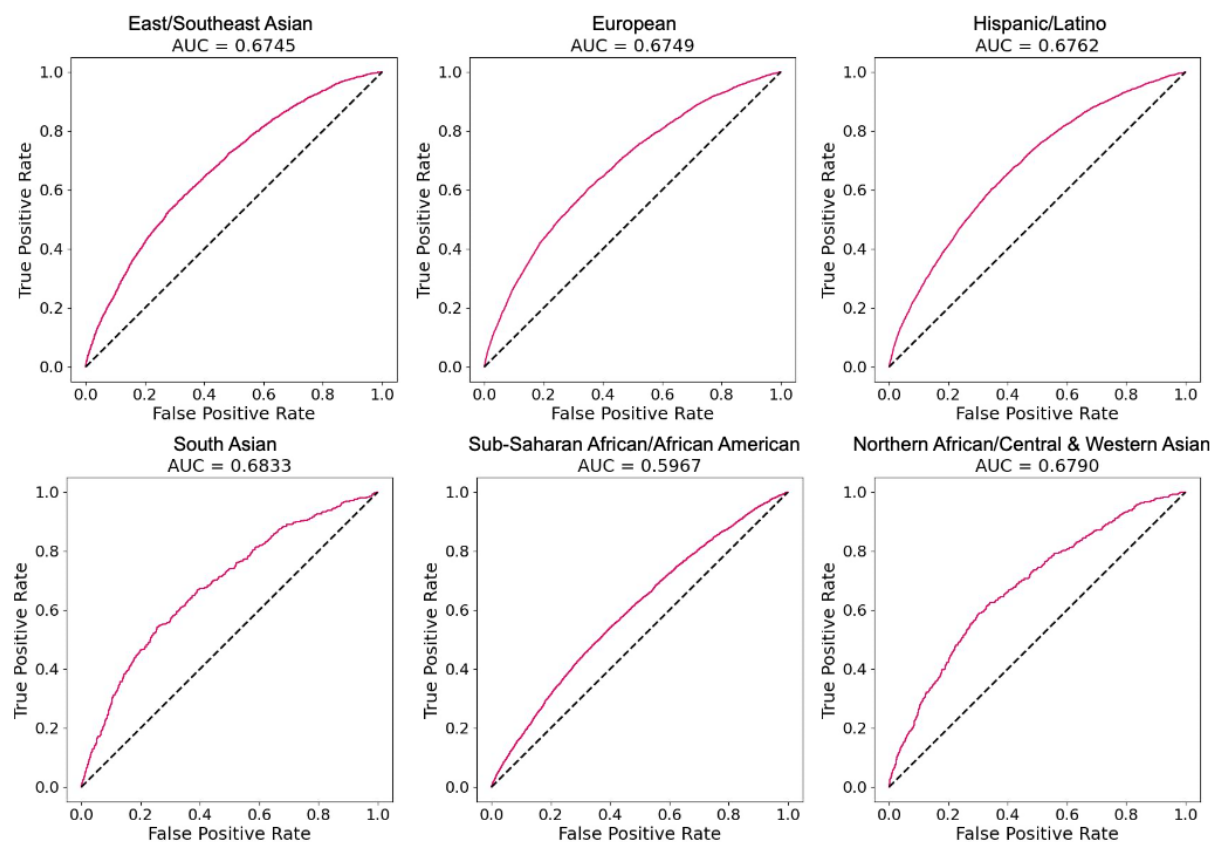
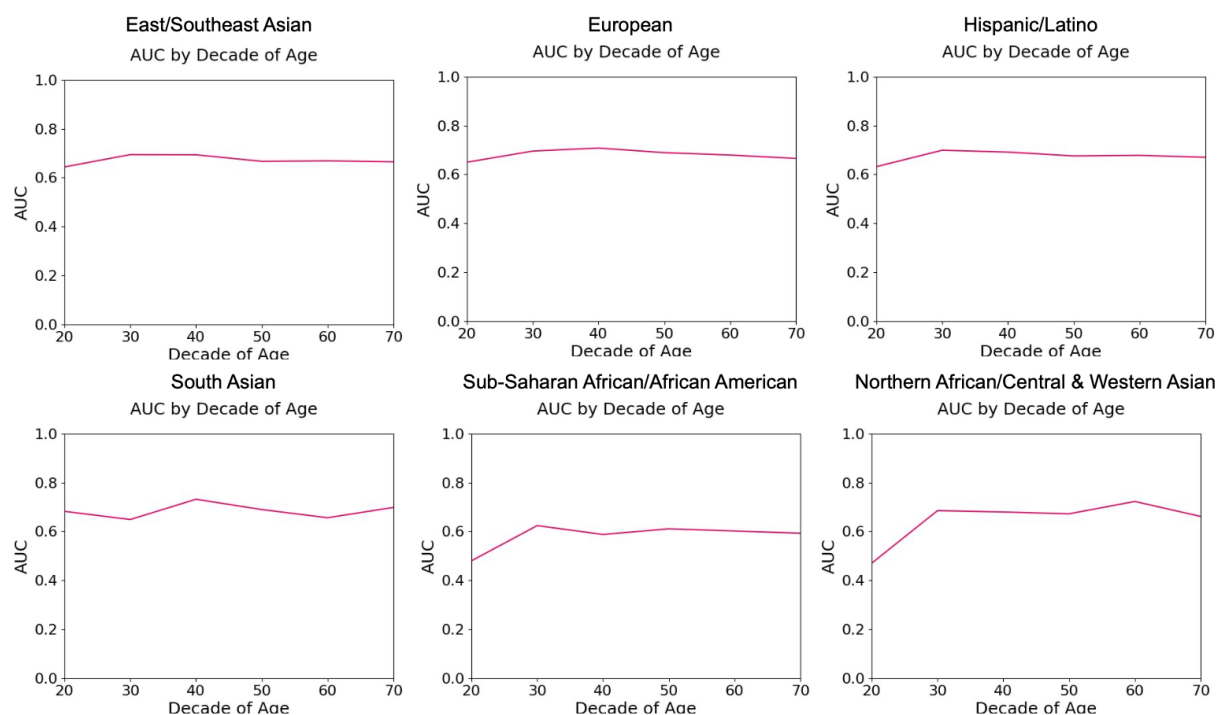
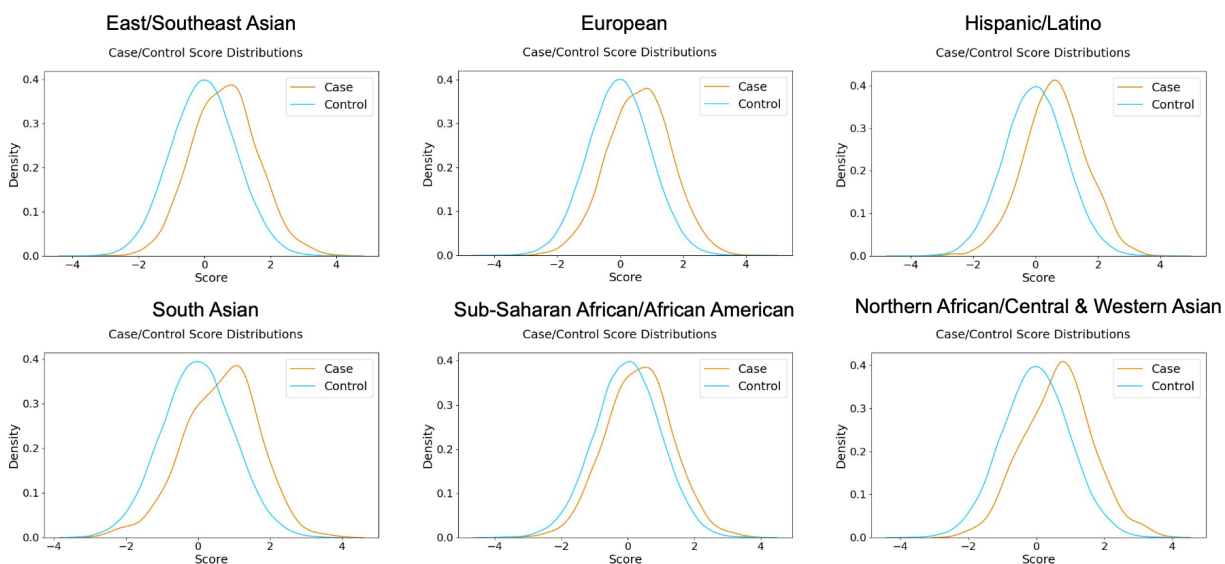
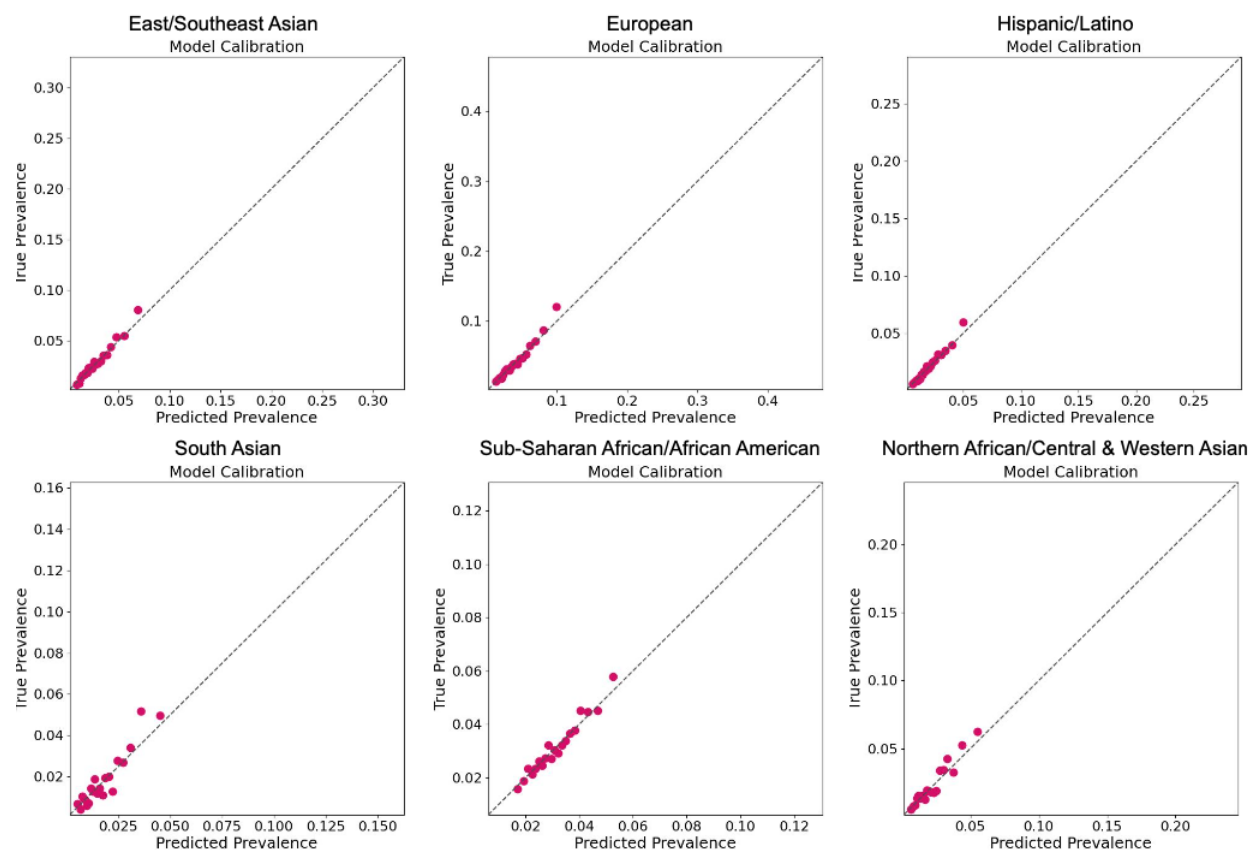
Figure 9-2: Gout AUROC across ancestry-specific test sets**Figure 9-3:** Gout AUROC within each decade of age across ancestry-specific test sets

Figure 9-4: Gout case/control standardized PGS distributions across ancestry-specific test sets**Figure 9-5:** Gout Platt-scaled calibration plots across ancestry-specific test sets

10. Nonalcoholic Fatty Liver Disease (NAFLD)

James R. Ashenhurst, PhD, lead analyst. Report added December, 2020.

Phenotype definition

Cases and controls were defined based on self-reported responses to questions about current or past diagnosis of nonalcoholic fatty liver disease (NAFLD) or nonalcoholic steatohepatitis (NASH). All question phrasings focused on receiving a diagnosis of or treatment for NAFLD or NASH. Although some phrasings had minor differences, the general phrasing was: “Have you ever been diagnosed with, or treated for, fatty liver disease, including NASH?” A follow-up question asked: “Did your doctor tell you that your fatty liver was mainly caused by any of the following? Alcohol, pregnancy, overweight, autoimmune or inherited, other, not told, or not sure.” Those who reported fatty liver caused by alcohol, pregnancy, or autoimmune or inherited conditions were excluded from analysis. Alcohol consumption reported in other surveys was also taken into account. Anyone reporting a personal history of alcoholism or alcohol use disorder was excluded from analysis, as was anyone who reported consuming 4 or more alcoholic drinks every day across a two week period. Participants answering affirmatively to current or past diagnosis of NAFLD or NASH who were not otherwise excluded were treated as cases, while those who reported no history of fatty liver diagnosis or treatment were included as controls.

Model description

For the development of this report, there were sufficient data to attempt ancestry-specific GWAS among those of European and Hispanic/Latino descent. Therefore, two unique GWAS were conducted and meta-analyzed. For model training, the European and Hispanic/Latino training cohorts were combined into a mega-cohort, with genetic features optimized against ancestry-specific validation sets comprised of individuals of European, Hispanic/Latino, East/Southeast Asian, South Asian, and Sub-Saharan African/African American descent. Three variant sets / models (Table 10-0) were selected based on performance in the validation sets, and were finally assessed in ancestry-specific test sets (see Table 10-1). Additional features in the model were age, age², sex, an age*sex interaction term, as well as the first five global principal components (PCs) to account for population stratification. Genomic

PCs are not used when determining an individual customer's report result. Personalized binary qualitative results and estimated likelihoods are based on genetic ancestry and birth sex.

Table 10-0: NAFLD model and variant set descriptions

Optimized Ancestry Group	Applicable Ancestry Group(s)	No. of variants
European, Hispanic/Latino	European, South Asian, Northern African/Western Asian, Hispanic/Latino	13,269
Sub-Saharan African/African American	Sub-Saharan African/African American	1,420
East/Southeast Asian	East/Southeast Asian	17,860

Figure 10-1: NAFLD phenotype distribution in the European + Hispanic/Latino training set

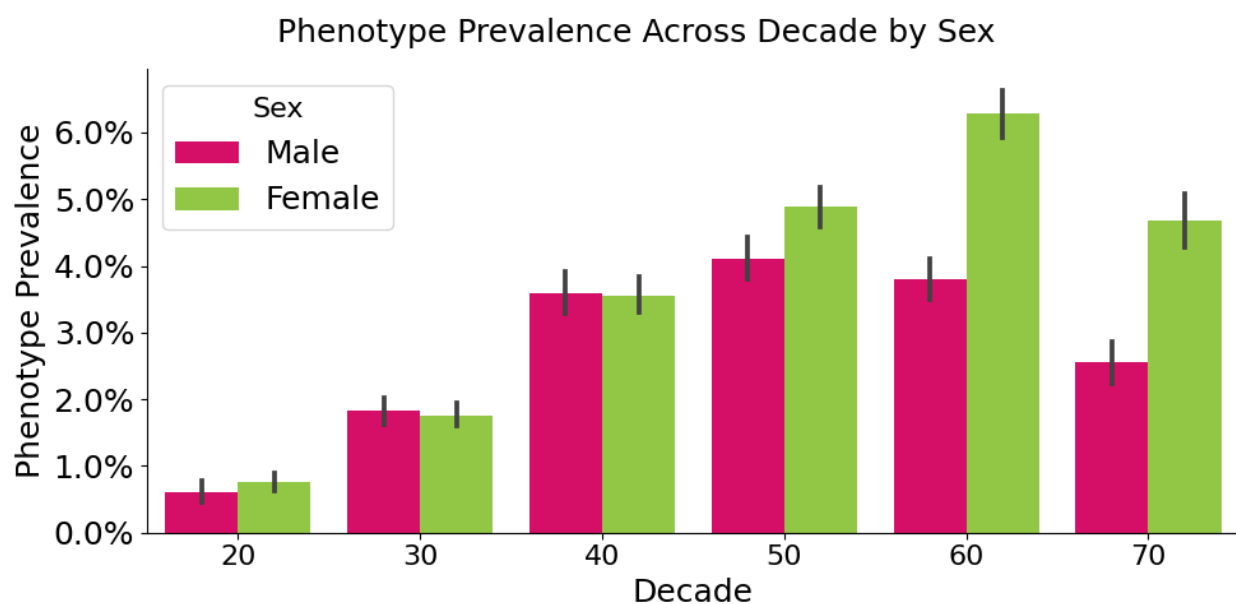


Table 10-1: NAFLD participant cohort descriptives

Platform	Ancestry Group	Sample Use	N	Age mean (SD)	Sex (% female)	NAFLD prevalence (%)
V1 to V5	Hispanic/Latino	GWAS	300,210	40.26 (15.71)	57.08%	3.67%
V1 to V5	European	GWAS	1,840,672	50.00 (17.41)	55.68%	3.40%
V5	European, Hispanic/Latino	Training three models	1,837,435	47.23 (15.83)	56.63%	3.30%
	European	Testing	173,677	48.46 (15.83)	56.64%	3.28%
	Sub-Saharan African/African American	Testing	47,539	42.00 (14.59)	58.12%	1.46%
	East/Southeast Asian	Testing	40,815	37.12 (12.87)	60.87%	2.42%
	Hispanic/Latino	Testing	30,572	40.09 (13.91)	57.40%	3.51%
	South Asian	Testing	19,961	39.00 (12.58)	42.66%	2.23%
	Northern African/Western Asian	Testing	17,751	42.08 (14.52)	44.52%	2.69%

Table 10-2: NAFLD model performance characteristics

Ancestry Group (test sets)	Full Model AUROC	Genetics Only AUROC	Odds Ratio top 5% versus average (95% CIs)	Odds Ratio top 5% versus bottom 5% (95% CIs)
European	0.7032	0.6423	3.16 (2.86, 3.49)	7.47 (6.07, 9.21)
Sub-Saharan African/African American	0.7238	0.5728	2.26* (1.66, 3.07)	2.00* (1.31, 3.05)
East/Southeast Asian	0.7026	0.5924	2.03 (1.57, 2.63)	3.88 (2.46, 6.11)
Hispanic/Latino	0.7266	0.6305	2.67 (2.11, 3.39)	7.08 (4.29, 11.69)
South Asian	0.6930	0.5936	2.73 (1.92, 3.87)	4.02 (2.22, 7.29)
Northern African/Western Asian	0.6935	0.6106	1.73 (1.20, 2.50)	7.30 (3.09, 17.26)

*The bottom 5% versus average comparison was not significant, likely due to sample size and power.

Table 10-3: NAFLD qualitative result characteristics

Ancestry group (test sets)	Odds Ratio for Result Threshold	Percent Above Threshold	Likelihood Ratio of "Increased" Result (95% CIs)
European	1.5	21.92%	1.86 (1.80, 1.93)
Sub-Saharan African/African American	1.5	6.80%	1.91 (1.57, 2.33)
East/Southeast Asian	1.5	11.08%	1.65 (1.44, 1.89)
Hispanic/Latino	1.5	19.12%	1.78 (1.63, 1.94)
South Asian	1.5	12.49%	1.73 (1.44, 2.08)
Northern African/Western Asian	1.5	14.99%	1.62 (1.38, 1.91)

Baseline prevalences used for likelihood estimate

Ancestry- and sex-specific baseline prevalences of ever been diagnosed with or treated for NAFLD were derived from external estimates as this condition is considerably under-recognized (Table 10-4). Ethnicity- and sex-specific prevalence estimates from individuals age 60 or older were chosen from two publications that analyzed data from several waves of the National Health and Nutrition Examination Survey (NHANES; Golabi et al., 2019; Lazo et al., 2013).

Table 10-4: NAFLD baseline prevalences from NHANES samples age 60 or older.

Source	Group	Matched 23andMe Population(s)	Sex	N	Prevalence
PMID: 23703888 (Table 1)	White Non-Hispanic	European, Northern African/Western Asian, Other	Male	*	28.1%
			Female	*	25.7%
PMID: 23703888 (Table 1)	Black Non-Hispanic	Sub-Saharan African/African American	Male	*	22.6%
			Female	*	23.9%
PMID: 23703888 (Table 1)	Mexican American	Hispanic/Latino	Male	*	33.4%
			Female	*	34.4%
PMID: 30597715 (Table S6)	Asian American	East/Southeast Asian, South Asian	Male & Female	**	24.5%

* Total N for PMID: 23703888 was 12,454, but specific sample sizes for ancestry- and sex-specific estimates were not provided.

** Total Asian American analysis sample size in PMID: 30597715 was 816, but the count of individuals age 60 or older was not provided.

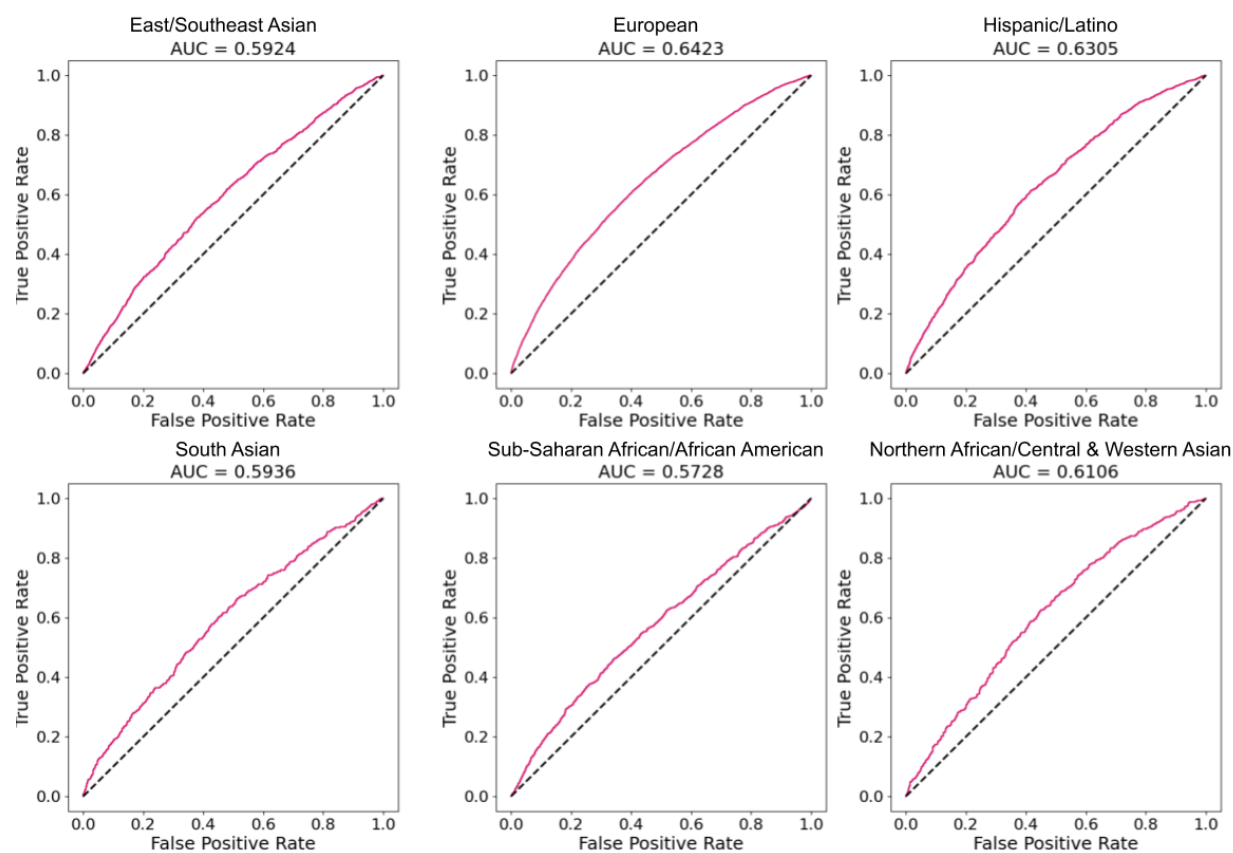
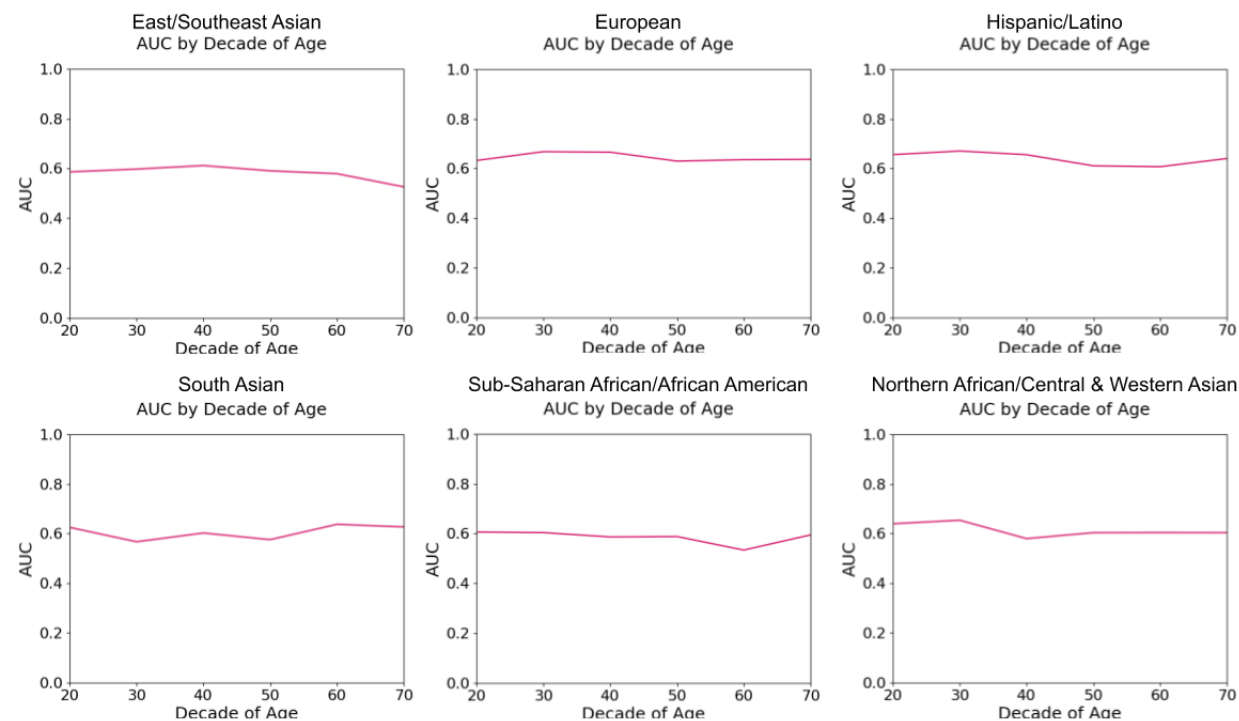
Figure 10-2: NAFLD AUROC across ancestry-specific test sets**Figure 10-3:** NAFLD AUROC within each decade of age across ancestry-specific test sets

Figure 10-4: NAFLD case/control standardized PGS distributions across ancestry-specific test sets

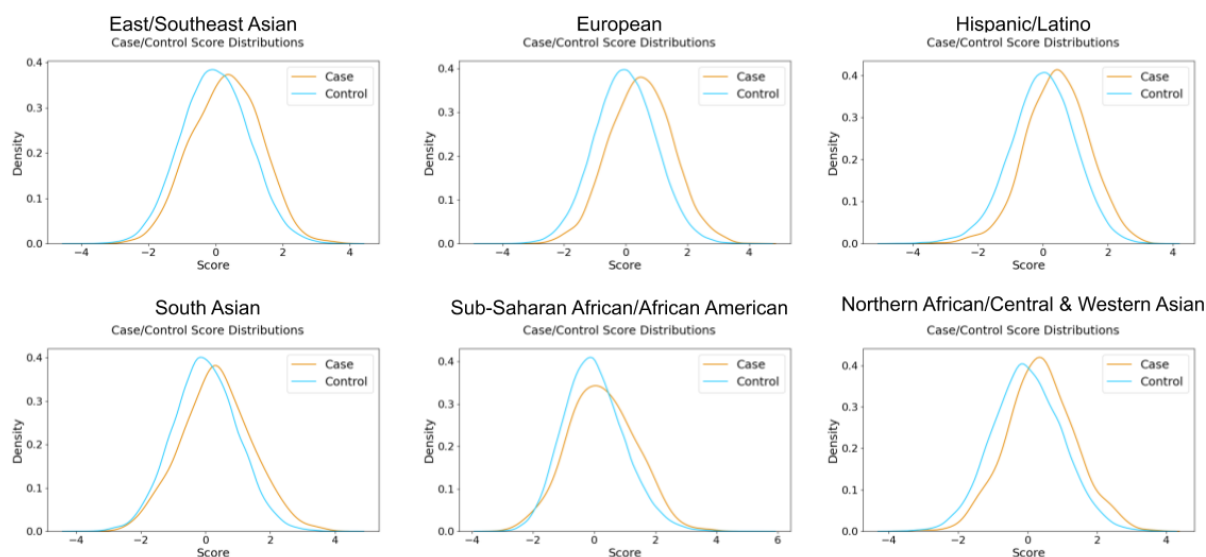
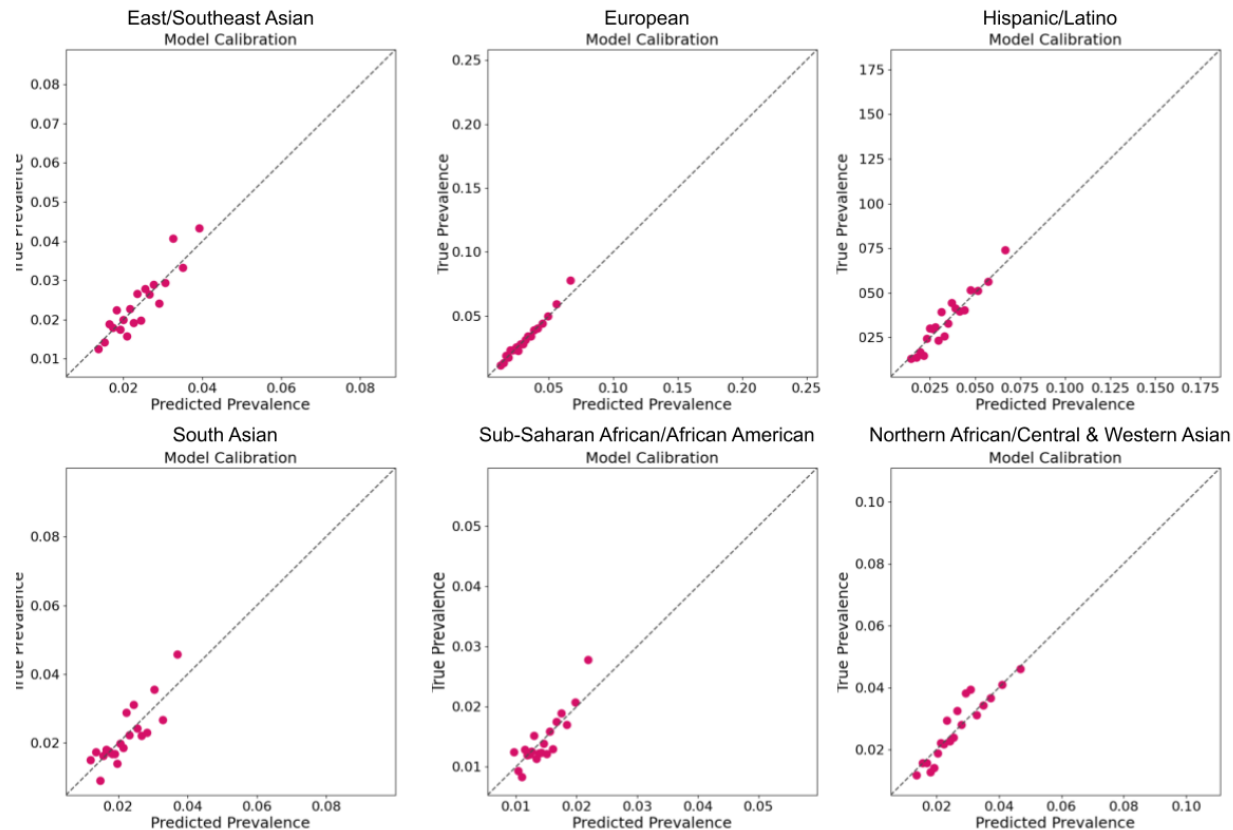


Figure 10-5: NAFLD Platt-scaled calibration plots across ancestry-specific test sets



References

- Centers for Disease Control and Prevention (CDC). (2017). *Behavioral Risk Factor Surveillance System Survey Data*. Centers for Disease Control and Prevention.
https://www.cdc.gov/brfss/annual_data/annual_2017.html
- Evangelou, E., Warren, H. R., Mosen-Ansorena, D., Mifsud, B., Pazoki, R., Gao, H., Ntritsos, G., Dimou, N., Cabrera, C. P., Karaman, I., Ng, F. L., Evangelou, M., Witkowska, K., Tzanis, E., Hellwege, J. N., Giri, A., Velez Edwards, D. R., Sun, Y. V., Cho, K., ... Caulfield, M. J. (2018). Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nature Genetics*, 50(12), 1755.
<https://doi.org/10.1038/s41588-018-0297-3>
- Golabi, P., Paik, J., Hwang, J. P., Wang, S., Lee, H. M., & Younossi, Z. M. (2019). Prevalence and outcomes of non-alcoholic fatty liver disease (NAFLD) among Asian American adults in the United States. *Liver International: Official Journal of the International Association for the Study of the Liver*, 39(4), 748–757. <https://doi.org/10.1111/liv.14038>
- Lazo, M., Hernaez, R., Eberhardt, M. S., Bonekamp, S., Kamel, I., Guallar, E., Koteish, A., Brancati, F. L., & Clark, J. M. (2013). Prevalence of nonalcoholic fatty liver disease in the United States: The Third National Health and Nutrition Examination Survey, 1988-1994. *American Journal of Epidemiology*, 178(1), 38–45. <https://doi.org/10.1093/aje/kws448>
- Lipton, R. B., Bigal, M. E., Diamond, M., Freitag, F., Reed, M. L., Stewart, W. F., & AMPP Advisory Group. (2007). Migraine prevalence, disease burden, and the need for preventive therapy. *Neurology*, 68(5), 343–349. <https://doi.org/10.1212/01.wnl.0000252808.97649.21>
- Nielsen, J. B., Thorolfsdottir, R. B., Fritsche, L. G., Zhou, W., Skov, M. W., Graham, S. E., Herron, T. J., McCarthy, S., Schmidt, E. M., Sveinbjornsson, G., Surakka, I., Mathis, M. R., Yamazaki, M., Crawford, R. D., Gabrielsen, M. E., Skogholt, A. H., Holmen, O. L., Lin, M., Wolford, B. N.,

... Willer, C. J. (2018). Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. *Nature Genetics*, 50(9), 1234–1239.

<https://doi.org/10.1038/s41588-018-0171-3>

Willer, C. J., Schmidt, E. M., Sengupta, S., Peloso, G. M., Gustafsson, S., Kanoni, S., Ganna, A., Chen, J., Buchkovich, M. L., Mora, S., Beckmann, J. S., Bragg-Gresham, J. L., Chang, H.-Y., Demirkan, A., Den Hertog, H. M., Do, R., Donnelly, L. A., Ehret, G. B., Esko, T., ... Global Lipids Genetics Consortium. (2013). Discovery and refinement of loci associated with lipid levels. *Nature Genetics*, 45(11), 1274–1283. <https://doi.org/10.1038/ng.2797>

Zhou, W., Nielsen, J. B., Fritsche, L. G., Dey, R., Gabrielsen, M. E., Wolford, B. N., LeFaive, J., VandeHaar, P., Gagliano, S. A., Gifford, A., Bastarache, L. A., Wei, W.-Q., Denny, J. C., Lin, M., Hveem, K., Kang, H. M., Abecasis, G. R., Willer, C. J., & Lee, S. (2018). Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nature Genetics*, 50(9), 1335–1341.

<https://doi.org/10.1038/s41588-018-0184-y>