SUPPLEMENTAL METHODS

Neighborhood-level Characteristics

Neighborhood variables at the census tract level included a range of sociodemographic characteristics and neighborhood physical characteristics. The following neighborhood sociodemographic characteristics were included: percent male, percent 65 years of age or older, percent owner-occupied housing units, racial/ethnic composition (percent Hispanic, percent non-Hispanic Black), education (percent bachelor's degree or higher), percent 21-64 years of age with a disability, percent 65 years of age or older with a disability, and economic indicators (percent unemployed, log median household income, percent households receiving public assistance, percent families living in poverty, percent 65+ population living in poverty).

In addition, we assessed physical features of the neighborhood related to population density, housing density, and street connectivity. Housing density was operationalized as number of housing units per square mile. Street connectivity was assessed with three measures: median block length (per 1,000 feet), alpha (the ratio of the actual number of complete loops to the maximum number of possible loops given the number of intersections), and gamma (the ratio of actual street segments to maximum possible given the number of intersections). Higher ratios for alpha and gamma, and lower values of block length, indicate greater connectivity.

During the 1990s, when many HRS participants were enrolled, annual measures at the tract level for the entire country are not available. For this reason, neighborhood characteristics were drawn from the 2000 Decennial Census and linked to HRS participants based on census tract of residence in 2000.

Evaluation of Model Performance

We conducted formal evaluations of model performance. To do so, we first split the sample into a random 70% subsample (the “training set”) and a 30% subsample (the “test set”). We sampled by household rather than respondent ID, since spouses of HRS respondents may also be included in the data set. We then fit regressions on the training set (logistic regressions for binary outcomes, linear regressions for continuous outcomes). We used coefficients from these models to predict outcomes among those in the test set. For binary outcomes, we used the actual and predicted outcomes to calculate sensitivity, specificity, accuracy, and Brier scores. For continuous outcomes, we obtained R-squared values by regressing the predicted values on the actual values in the test set. While no set of predictors demonstrated consistently improved performance across different outcomes (eTables 4-6), these evaluations were limited in part by the relatively small size of the sample (N = 7,522 for White participants and 1,198 for Black participants).

eTable 1. Percent of variance explained by demographic, socioeconomic, and genetic covariates among white participants

	R^2 (95% Confidence Interval)						
Model Covariates	*Diabetes*	*Heart disease*	*Depression*	*Smoking*	*BMI*	*Total cholesterol*	*HDL*
Demographic	1.2 (0.7, 1.7)	5.3 (4.3, 6.3)	2.5 (1.8, 3.2)	4.9 (3.9, 5.8)	5.6 (4.6, 6.6)	8.0 (6.8, 9.2)	7.5 (6.2, 8.8)
Demographic + SEP	2.2 (1.5, 2.9)	6.7 (5.6, 7.8)	6.9 (5.8, 8.0)	6.6 (5.5, 7.7)	8.1 (6.9, 9.3)	8.4 (7.2, 9.6)	9.8 (8.4, 11.2)
Demographic + PCs	1.5 (1.0, 2.0)	5.6 (4.6, 6.6)	3.0 (2.2, 3.8)	5.4 (4.4, 6.4)	5.9 (4.9, 6.9)	8.1 (6.9, 9.3)	7.8 (6.5, 9.1)
Demographic + PGS	2.2 (1.5, 2.9)	5.6 (4.6, 6.6)	3.8 (3.0, 4.6)	6.2 (5.1, 7.3)	11.9 (10.5, 13.3)	8.6 (7.4, 9.8)	9.3 (7.9, 10.7)
Demographic + PCs + PGS	2.8 (2.1, 3.5)	6.0 (5.0, 7.0)	4.0 (3.1, 4.9)	6.5 (5.4, 7.6)	12.5 (11.1, 13.9)	8.7 (7.5, 9.9)	9.6 (8.2, 11.0)
Demographic + SEP + PCs + PGS	3.7 (2.9, 4.5)	7.4 (6.3, 8.5)	8.0 (6.8, 9.0)	8.1 (6.9, 9.3)	14.3 (12.8, 15.8)	9.0 (7.7, 10.3)	11.6 (10.1, 13.1)

N = 7,522. Percent of variance explained is the adjusted R-squared value from a multivariate linear regression of the given outcome on the given combination of covariates. BMI: body mass index; HDL: high-density lipoprotein cholesterol; PCs: principal components for genetic ancestry; PGS: polygenic score specific to relevant health condition; SEP: socioeconomic position.

eTable 2. Percent of variance explained by demographic, socioeconomic, and genetic covariates among Black participants

	R^2 (95% Confidence Interval)						
Model Covariates	*Diabetes*	*Heart disease*	*Depression*	*Smoking*	*BMI*	*Total cholesterol*	*HDL*
Demographic	2.4 (0.7, 4.1)	1.8 (0.3, 3.3)	2.0 (0.4, 3.6)	6.9 (4.1, 9.7)	8.1 (5.2, 11.0)	3.1 (1.1, 5.1)	3.4 (1.1, 5.7)
Demographic + SEP	4.2 (1.8, 6.6)	3.8 (1.5, 6.1)	9.2 (5.9, 12.5)	7.6 (4.5, 10.7)	9.7 (6.3, 13.1)	4.3 (1.8, 6.8)	4.4 (1.7, 7.1)
Demographic + PCs	3.5 (1.5, 5.5)	2.5 (0.8, 4.2)	3.7 (1.6, 5.8)	7.5 (4.7, 10.3)	8.2 (5.3, 11.1)	4.2 (1.9, 6.5)	4.3 (1.8, 6.8)
Demographic + PGS	2.7 (0.9, 4.5)	1.8 (0.3, 3.3)	3.0 (1.1, 4.9)	7.0 (4.2, 9.8)	9.0 (5.9, 12.1)	3.6 (1.5, 5.7)	3.8 (1.4, 6.2)
Demographic + PCs + PGS	4.2 (2.0, 6.4)	2.5 (0.8, 4.2)	3.8 (1.7, 5.9)	7.7 (4.8, 10.6)	9.8 (6.7, 12.9)	4.7 (2.3, 7.1)	4.9 (2.2, 7.6)
Demographic + SEP + PCs + PGS	6.1 (3.3, 8.9)	4.7 (2.2, 7.2)	10.3 (6.9, 13.7)	8.4 (5.2, 11.6)	11.2 (7.7, 14.7)	6.2 (3.3, 8.6)	6.0 (2.9, 9.1)

N = 1,198. Percent of variance explained is the adjusted R-squared value from a multivariate linear regression of the given outcome on the given combination of covariates. BMI: body mass index; HDL: high-density lipoprotein cholesterol; PCs: principal components for genetic ancestry; PGS: polygenic score specific to relevant health condition; SEP: socioeconomic position.

eTable 3. P-values obtained from likelihood-ratio tests between each model and the base model that included only demographic predictors, by race

	Diabetes	*Heart Disease*	*Depression*	*Smoker*
White participants				
Demographic + SEP	< 0.001	< 0.001	< 0.001	< 0.001
Demographic + PCs	0.0049	0.042	< 0.001	< 0.001
Demographic + PGS	< 0.001	< 0.001	< 0.001	< 0.001
Demographic + PCs + PGS	< 0.001	< 0.001	< 0.001	< 0.001
Demographic + SEP + PCs + PGS	< 0.001	< 0.001	< 0.001	< 0.001
Black participants				
Demographic + SEP	< 0.001	< 0.001	< 0.001	< 0.001
Demographic + PCs	0.469	0.52	0.016	0.6355
Demographic + PGS	0.12	0.71	0.0014	0.0413
Demographic + PCs + PGS	0.26	0.60	0.022	0.352
Demographic + SEP + PCs + PGS	< 0.001	< 0.001	< 0.001	< 0.001

N = 7,522 white and 1,198 Black participants. Underlying models involved multivariate linear regression of the given outcome on the given combination of covariates. PCs: principal components for genetic ancestry; PGS: polygenic score specific to relevant health condition; SEP: socioeconomic position.

eTable 4. Measures of model performance for binary outcomes in the sample of white participants

Model covariates	Sensitivity	Specificity	Accuracy	Brier Score
Diabetes				
Demographic	50.5	58.5	56.8	0.17
Demographic + SEP	61.7	51.9	54.1	0.17
Demographic + PCs	51.9	55.2	54.4	0.17
Demographic + PGs	60.2	58.6	58.9	0.17
Demographic + PCs + PGs	58.6	60.4	60	0.17
Demographic + SEP + PCs + PGs	53.7	61.9	60.2	0.17
Heart Disease				
Demographic	47.9	61.4	54.5	0.25
Demographic + SEP	44.6	67.2	55.7	0.25
Demographic + PCs	52.1	59.4	55.6	0.25
Demographic + PGs	52.1	60.1	56.0	0.25
Demographic + PCs + PGs	55.1	56.1	55.6	0.25
Demographic + SEP + PCs + PGs	59.0	52.4	55.7	0.25
Depression				
Demographic	53.9	56.9	55.6	0.24
Demographic + SEP	55.0	60.0	58.0	0.24
Demographic + PCs	60.5	51.4	55.1	0.24
Demographic + PGs	57.9	56.2	56.9	0.24
Demographic + PCs + PGs	54.6	59.3	57.4	0.24
Demographic + SEP + PCs + PGs	58.5	60.5	59.7	0.24
Smoker				
Demographic	60.7	65.0	62.7	0.23
Demographic + SEP	65.0	58.7	62.0	0.23
Demographic + PCs	58.7	65.4	61.9	0.24
Demographic + PGs	61.4	62.9	62.1	0.23
Demographic + PCs + PGs	54.0	68.1	60.7	0.24
Demographic + SEP + PCs + PGs	59.1	63.7	61.3	0.23

N = 7,522. PCs: principal components for genetic ancestry; PGS: polygenic score specific to relevant health condition; SEP: socioeconomic position.

eTable 5. Measures of model performance for binary outcomes in the sample of Black participants

Model covariates	Sensitivity	Specificity	Accuracy	Brier Score
Diabetes				
Demographic	48.4	57.8	54.4	0.23
Demographic + SEP	53.7	58.9	57.0	0.23
Demographic + PCs	55.6	50.0	52.0	0.24
Demographic + PGs	46.8	61.5	56.1	0.23
Demographic + PCs + PGs	55.6	53.2	54.1	0.23
Demographic + SEP + PCs + PGs	54.6	60.4	58.3	0.23
Heart Disease				
Demographic	64.4	48.6	54.7	0.26
Demographic + SEP	56.2	53.6	54.7	0.25
Demographic + PCs	61.4	48.6	53.5	0.27
Demographic + PGs	63.6	49.5	55.0	0.26
Demographic + PCs + PGs	65.9	45.7	53.5	0.27
Demographic + SEP + PCs + PGs	57.9	54.7	56.0	0.26
Depression				
Demographic	61.6	46.5	54.7	0.25
Demographic + SEP	59.7	65.8	62.7	0.23
Demographic + PCs	49.2	61.1	54.7	0.25
Demographic + PGs	64.3	50.3	57.9	0.25
Demographic + PCs + PGs	50.8	60.5	55.3	0.23
Demographic + SEP + PCs + PGs	56.5	66.4	61.3	0.23
Smoker				
Demographic	68.9	66.2	67.7	0.22
Demographic + SEP	61.0	70.4	65.2	0.22
Demographic + PCs	72.6	60.3	67.2	0.22
Demographic + PGs	61.6	69.5	65.1	0.22
Demographic + PCs + PGs	61.6	70.9	65.7	0.22
Demographic + SEP + PCs + PGs	61.0	66.7	63.5	0.23

N = 1,198. PCs: principal components for genetic ancestry; PGS: polygenic score specific to relevant health condition; SEP: socioeconomic position.

eTable 6. Adjusted R-squared values obtained from regressing the predicted values from each of the models on the actual values in the random 30% subsample

Model covariates	BMI	Cholesterol	HDL
White participants			
Demographic	5.9	5.3	7.7
Demographic + SEP	7.1	5.6	7.1
Demographic + PCs	6.0	5.3	7.2
Demographic + PGs	10.3	6.0	8.2
Demographic + PCs + PGs	10.8	6.0	8.2
Demographic + SEP + PCs + PGs	12.2	6.3	7.9
Black participants			
Demographic	8.8	0.3	5.1
Demographic + SEP	9.2	0.02	3.9
Demographic + PCs	8.5	0.4	4.9
Demographic + PGs	9.1	0.04	5.4
Demographic + PCs + PGs	9.4	0.2	5.3
Demographic + SEP + PCs + PGs	9.1	0.01	4.4

N = 7,522 white and 1,198 Black participants. Percent of variance explained is the adjusted R-squared value from a multivariate linear regression of the given outcome on the given combination of covariates. BMI: body mass index; HDL: high-density lipoprotein cholesterol; PCs: principal components for genetic ancestry; PGS: polygenic score specific to relevant health condition; SEP: socioeconomic position.