The American Journal of Human Genetics, Volume *109*

Supplemental information

Polygenic transcriptome risk scores for COPD and lung function improve cross-ethnic portability of prediction in the NHLBI TOPMed program

Xiaowei Hu, Dandi Qiao, Wonji Kim, Matthew Moll, Pallavi P. Balte, Leslie A. Lange, Traci M. Bartz, Rajesh Kumar, Xingnan Li, Bing Yu, Brian E. Cade, Cecelia A. Laurie, Tamar Sofer, Ingo Ruczinski, Deborah A. Nickerson, Donna M. Muzny, Ginger A. Metcalf, Harshavardhan Doddapaneni, Stacy Gabriel, Namrata Gupta, Shannon Dugan-Perez, L. Adrienne Cupples, Laura R. Loehr, Deepti Jain, Jerome I. Rotter, James G. Wilson, Bruce M. Psaty, Myriam Fornage, Alanna C. Morrison, Ramachandran S. Vasan, George Washko, Stephen S. Rich, George T. O'Connor, Eugene Bleecker, Robert C. Kaplan, Ravi Kalhan, Susan Redline, Sina A. Gharib, Deborah Meyers, Victor Ortega, Josée Dupuis, Stephanie J. London, Tuuli Lappalainen, Elizabeth C. Oelsner, Edwin K. Silverman, R. Graham Barr, Timothy A. Thornton, Heather E. Wheeler, TOPMed Lung Working Group, Michael H. Cho, Hae Kyung Im, and Ani Manichaikul

Supplemental Figures

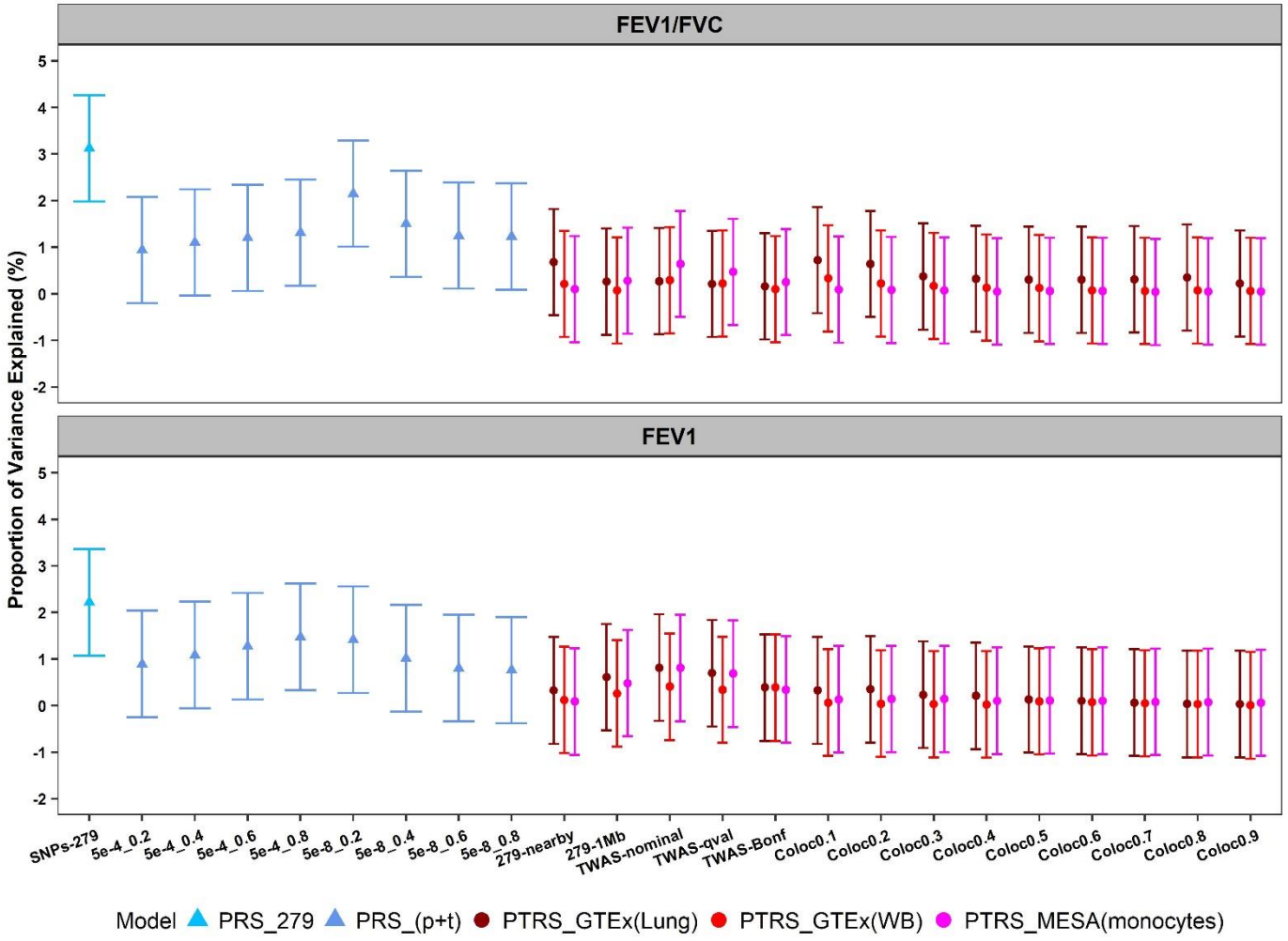

Figure S1: Prediction performance of all risk score candidates in multi-ethnic population/family-based cohorts for FEV1/FVC ratio and FEV1. Proportion of variance explained (%), estimated by 100*squared correlation between the observed phenotypes and the predicted phenotypes by risk score only; Data are shown as proportion of variance explained with 95% CI; PRS_279, PRS derived by previously published 279 variants for FEV1/FVC ratio or FEV1; PRS_(p+t), PRS derived by pruning and thresholding, a range of p-value and pairwise correlation thresholds were used to create eight candidates (5e-4_0.2 to 5e-8_0.8); 279-nearby and 279-1Mb, PTRS derived by genes nearby and within +/- 1Mb region of previously published 279 variants (for FEV1/FVC ratio or FEV1) respectively; TWAS-nominal, TWAS-qval, and TWAS-Bonf, PTRS derived by genes passing TWAS p-value threshold of 0.05, q-value, and Bonferroni respectively; Coloc0.1 to Coloc0.9, PTRS derived by genes with regional colocalization probability ranging from 0.1 to 0.9.

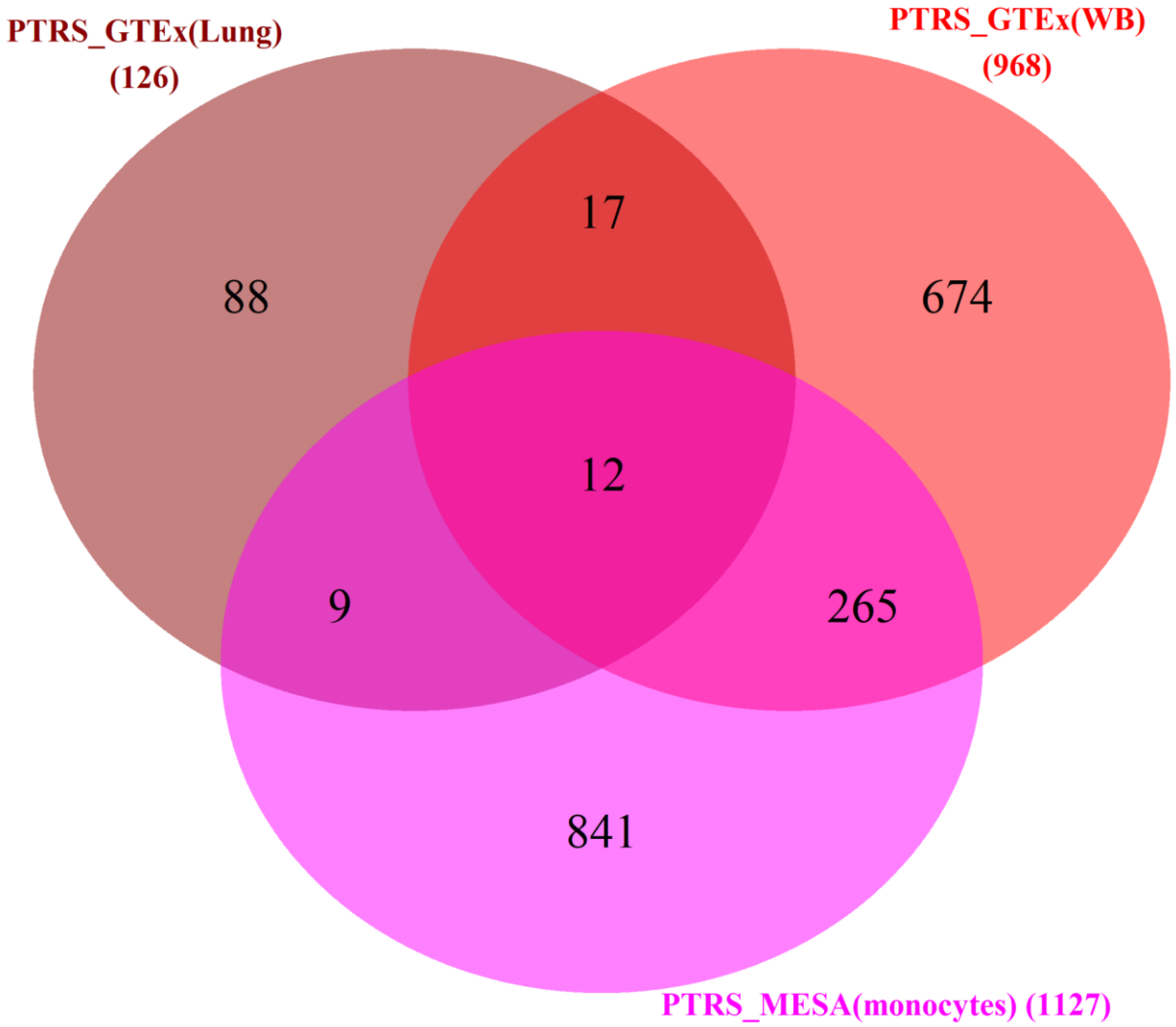

Figure S2: Venn diagram of the number of genes included in the best risk score candidate of each model for two COPD traits. The best candidates of three models are PTRS_GTEx(Lung): 279-nearby that was derived by genes nearby previously published 279 variants for FEV1/FVC ratio; PTRS_GTEx(WB): TWAS-qval that was derived by genes passing TWAS q-value threshold; PTRS_MESA(monocytes): TWAS-nominal that was derived by genes passing TWAS p-value threshold of 0.05. The number in the parenthesis shows the gene size of the best candidate.

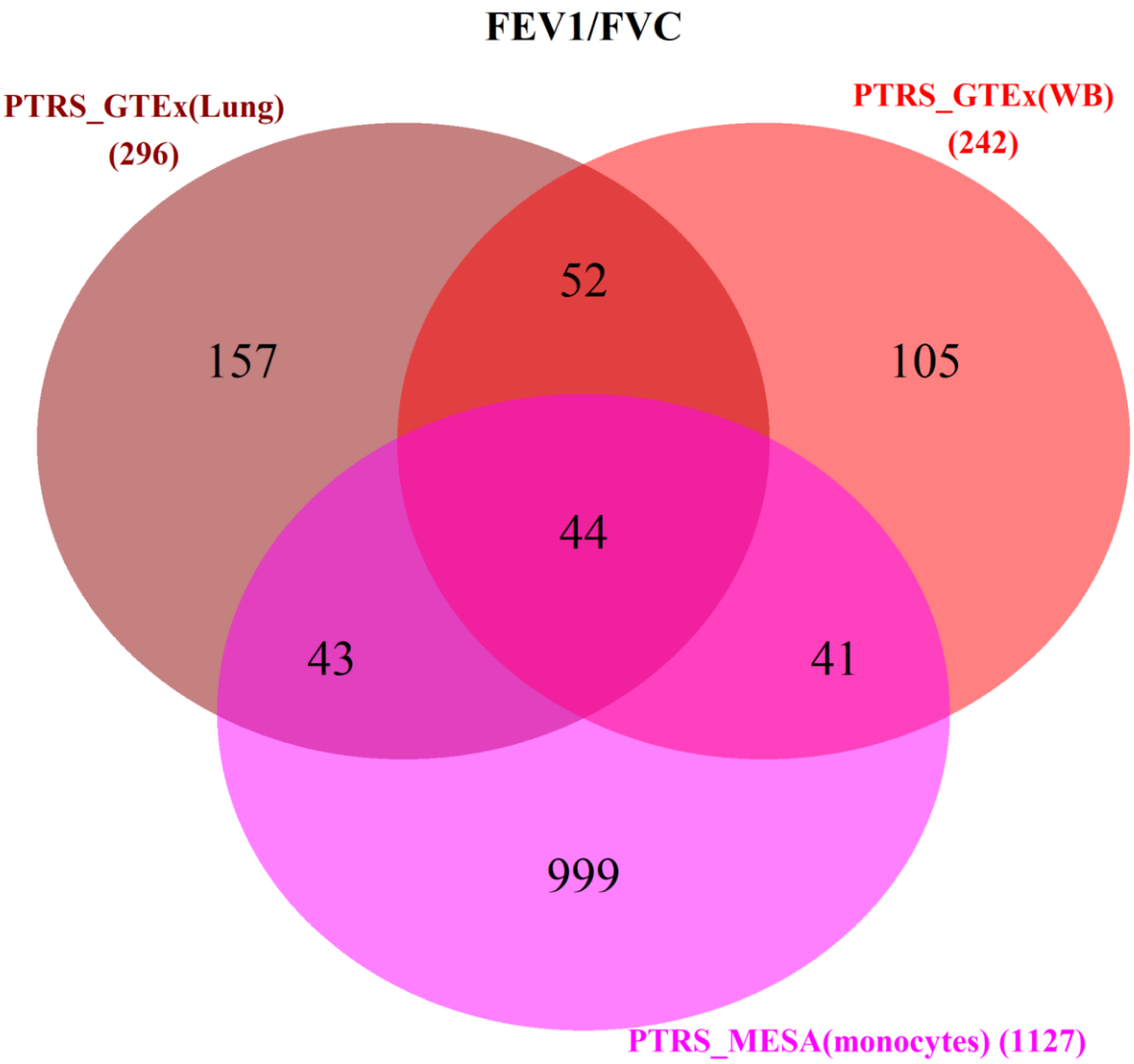

Figure S3: Venn diagram of the number of genes included in the best risk score candidate of each model for FEV1/FVC ratio. The best candidates of three models are PTRS_GTEx(Lung): Coloc0.1; PTRS_GTEx(WB): Coloc0.1; Coloc0.1 was derived by genes with regional colocalization probability greater than 0.1; PTRS_MESA(monocytes): TWAS-nominal that was derived by genes passing TWAS p-value threshold of 0.05. The number in the parenthesis shows the gene size of the best candidate.

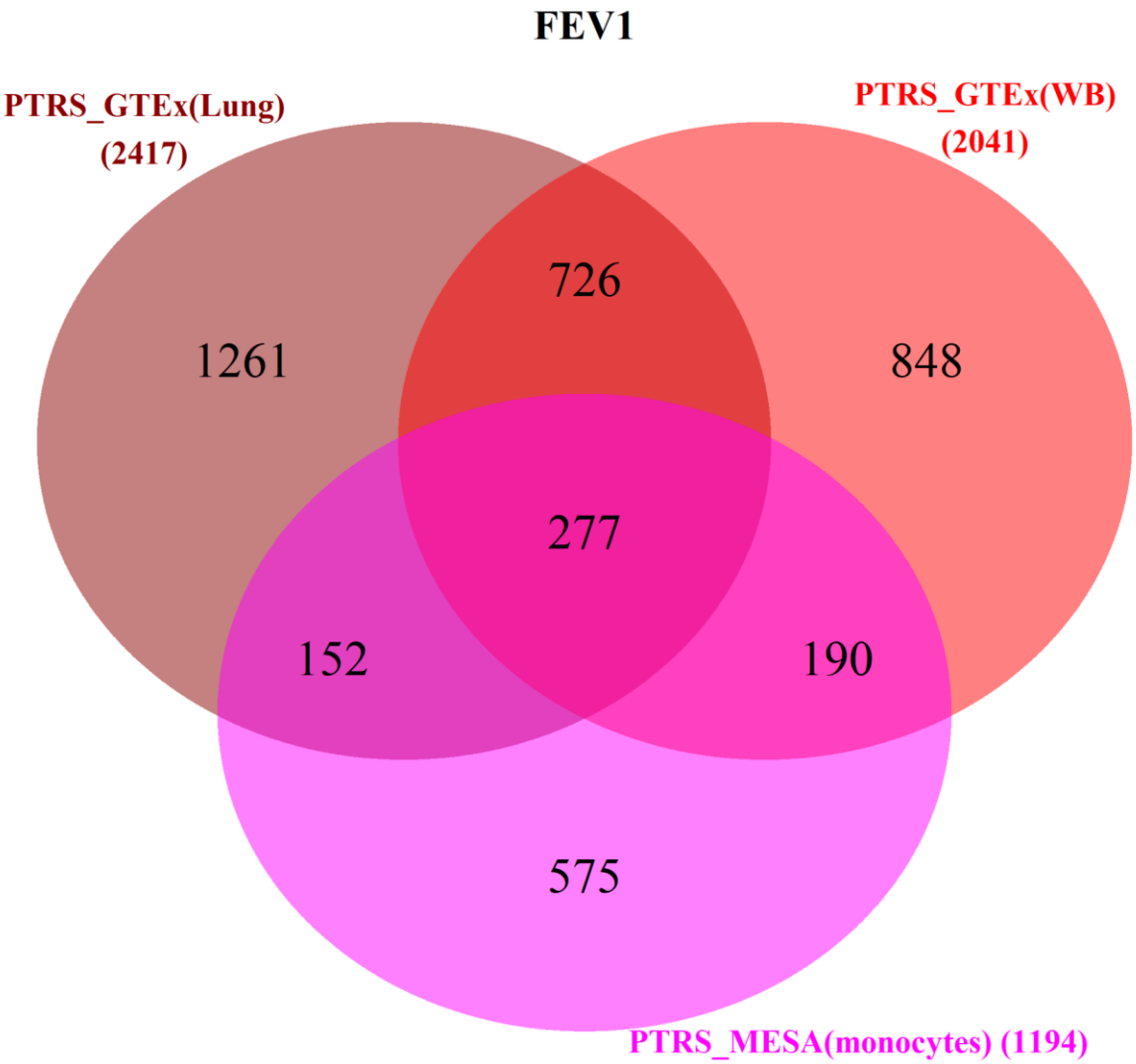

Figure S4: Venn diagram of the number of genes included in the best risk score candidate of each model for FEV1. The best candidates of three models are PTRS_GTEx(Lung): TWAS-nominal; PTRS_GTEx(WB): TWAS-nominal; PTRS_MESA(monocytes): TWAS-nominal. TWAS-nominal was derived by genes passing TWAS p-value threshold of 0.05. The number in the parenthesis shows the gene size of the best candidate.

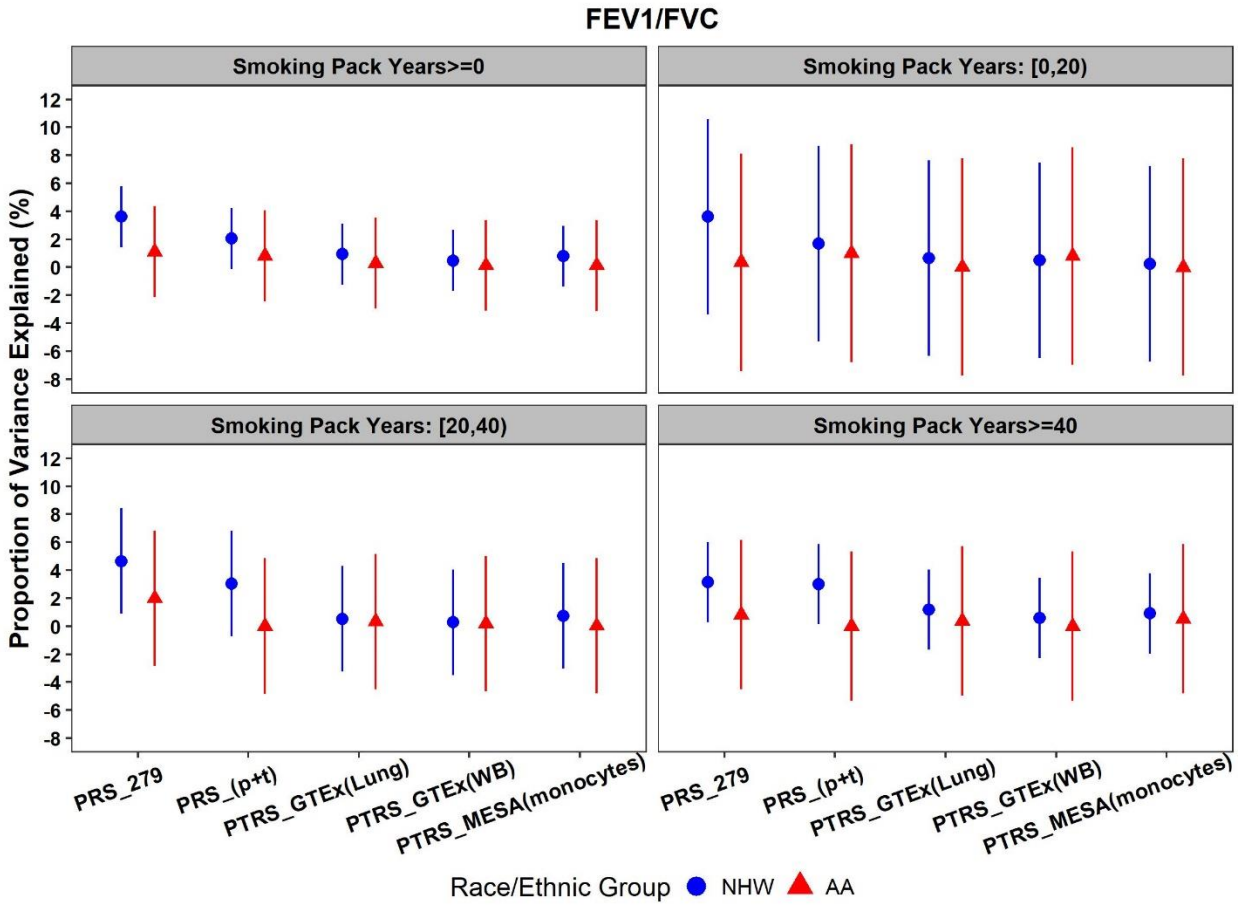

Figure S5: Prediction performance of the best risk scores with FEV1/FVC ratio in COPD-enriched studies. NHW, Non-Hispanic Whites; AA, African Americans; the risk scores used in the analyses were based on the prediction performance of each smoking stratum on population/family-based cohorts for FEV1/FVC ratio; proportion of variance explained (%) was estimated by 100*squared correlation between the observed phenotypes and the predicted phenotypes by risk score only. Data are shown as meta-analyzed results with error bars as 95% CIs of proportion of variance explained.

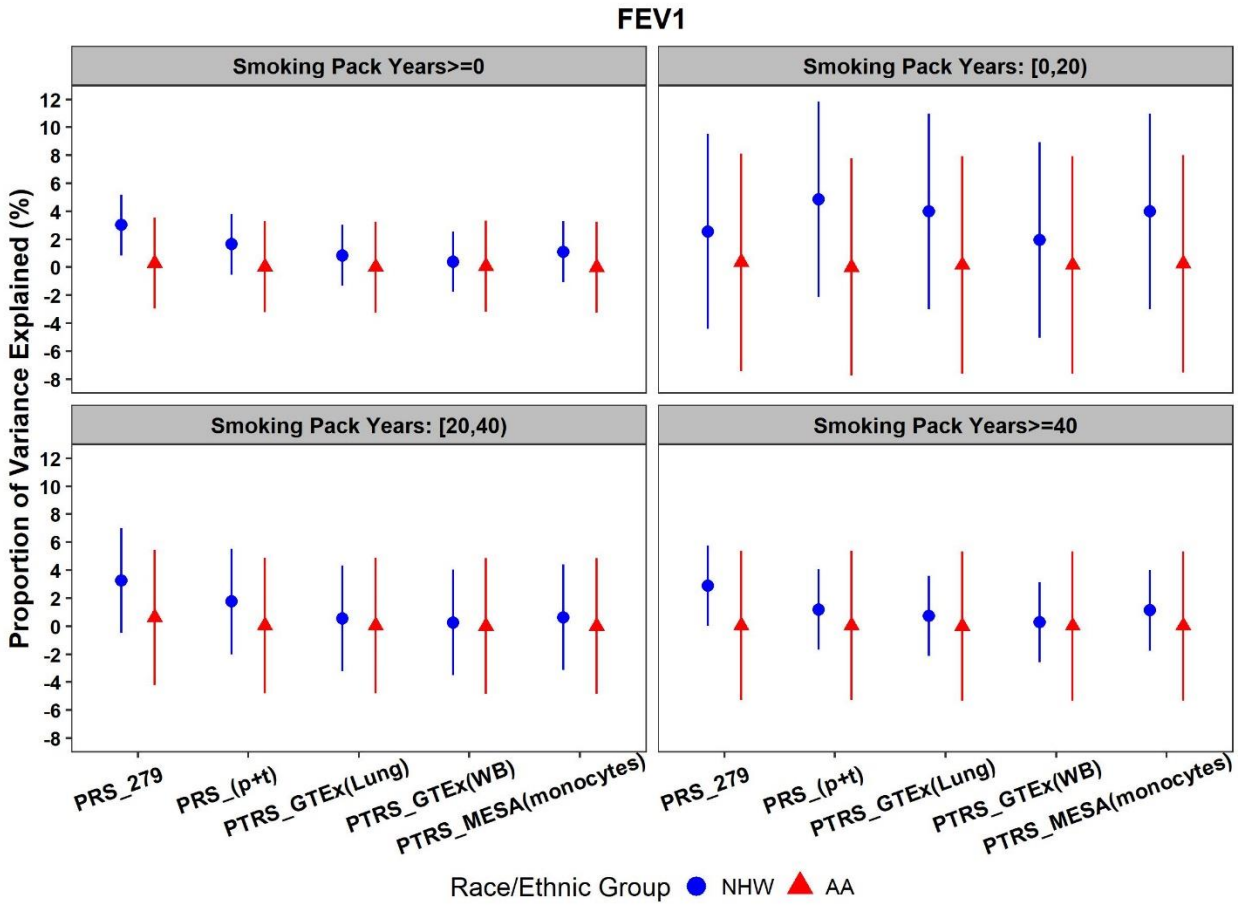

Figure S6: Prediction performance of the best risk scores with FEV1 in COPD-enriched studies. NHW, Non-Hispanic Whites; AA, African Americans; the risk scores used in the analyses were based on the prediction performance of each smoking stratum on population/family-based cohorts for FEV1; proportion of variance explained (%) was estimated by 100*squared correlation between the observed phenotypes and the predicted phenotypes by risk score only. Data are shown as meta-analyzed results with error bars as 95% CIs of proportion of variance explained.

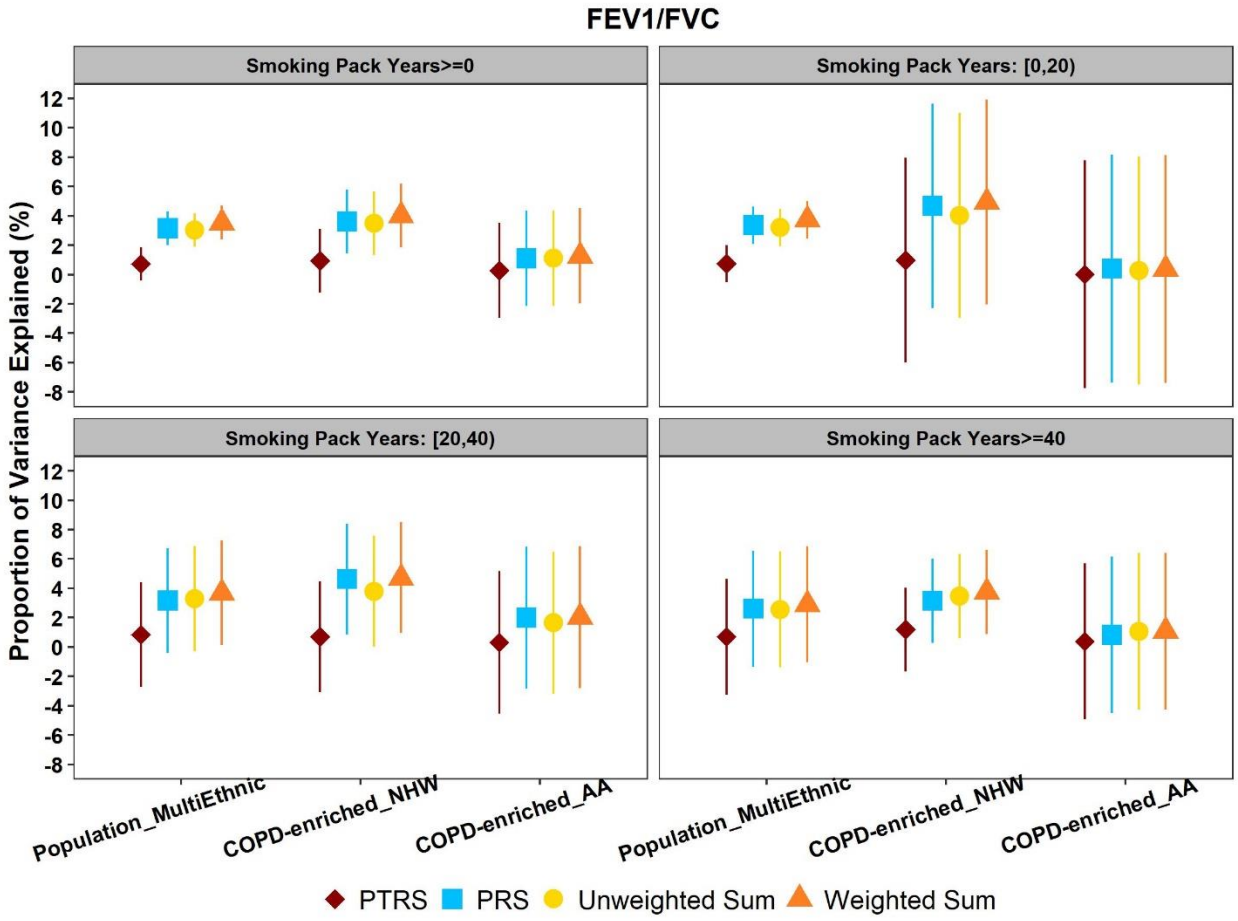

Figure S7: Prediction performance of the combined risk scores for FEV1/FVC ratio. The risk score candidates, PTRS_GTEx(Lung): 279-nearby and PRS_279: SNPs-279, were used for PTRS and PRS respectively in the analyses; Unweighted Sum and Weighted Sum refer to the direct summation and the weighted summation of PTRS and PRS respectively; Data are shown as proportion of variance explained (%) which was estimated by 100*squared correlation between the observed phenotypes and the predicted phenotypes by combined risk score only, error bars are 95% CIs; For COPD-enriched studies, the results were meta-analyzed. NHW, Non-Hispanic Whites; AA, African Americans; Population_MultiEthnic, multi-ethnic samples in population/family-based cohorts.

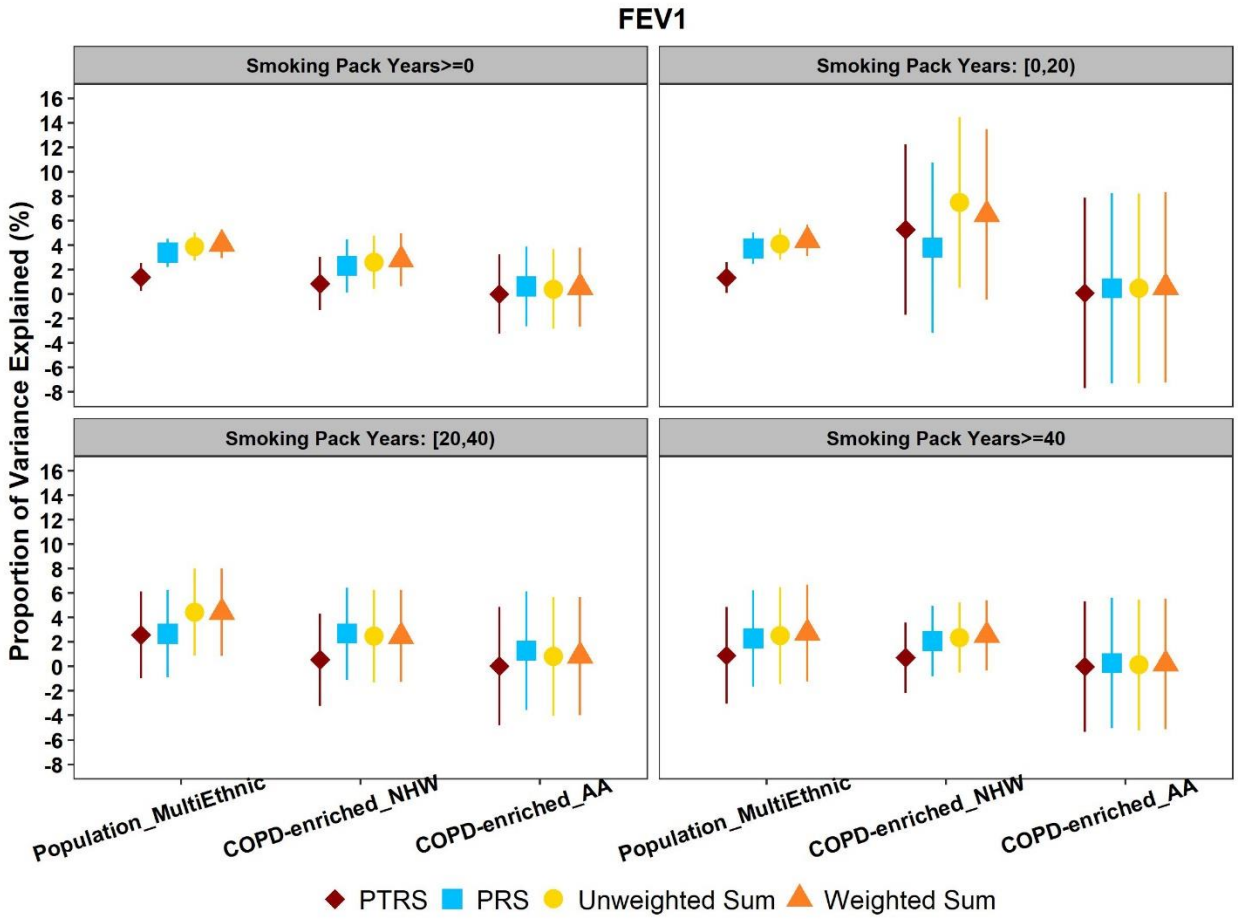

Figure S8: Prediction performance of the combined risk scores for FEV1. The risk score candidates, PTRS_GTEx(Lung): 279-nearby and PRS_279: SNPs-279, were used for PTRS and PRS respectively in the analyses; Unweighted Sum and Weighted Sum refer to the direct summation and the weighted summation of PTRS and PRS respectively; Data are shown as proportion of variance explained (%) which was estimated by 100*squared correlation between the observed phenotypes and the predicted phenotypes by combined risk score only, error bars are 95% CIs; For COPD-enriched studies, the results were meta-analyzed. NHW, Non-Hispanic Whites; AA, African Americans; Population_MultiEthnic, multi-ethnic samples in population/family-based cohorts.

Supplemental Tables

In this section we will include Table S1, Tables S14-15, and Tables S20-22. The Excel file will contain Tables S2-13 and Tables S16-19.

	FEV1/FVC			**FEV1**		
Candidate	PTRS_GTEx (Lung)	PTRS_GTEx (WB)	PTRS_MESA (monocytes)	PTRS_GTEx (Lung)	PTRS_GTEx (WB)	PTRS_MESA (monocytes)
279-nearby	126	102	75	126	102	75
279-1Mb	2797	2447	1400	2797	2447	1400
TWAS-nominal	2407	2065	1127	2417	2041	1194
TWAS-qval	1175	968	526	1164	928	513
TWAS-Bonf	271	219	143	206	200	142
Coloc0.1	296	242	289	301	260	352
Coloc0.2	185	147	207	182	161	230
Coloc0.3	126	107	147	128	102	177
Coloc0.4	96	72	119	97	85	149
Coloc0.5	72	50	98	56	50	131
Coloc0.6	55	32	80	43	38	111
Coloc0.7	40	26	61	22	20	82
Coloc0.8	35	16	42	8	14	67
Coloc0.9	19	8	25	2	2	46

Table S1: The number of genes included in each risk score candidate corresponding to three PTRS models for both FEV1/FVC ratio and FEV1. 279-nearyby and 279-1Mb, PTRS derived by genes nearby and within +/- 1Mb region of previously published 279 variants (for FEV_1/FVC ratio or FEV1) respectively; TWAS-nominal, TWAS-qval, and TWAS-Bonf, PTRS derived by genes passing TWAS p-value threshold of 0.05, q-value, and Bonferroni respectively; Coloc0.1 to Coloc0.9, PTRS derived by genes with regional colocalization probability ranging from 0.1 to 0.9.

Moderate-to-Severe COPD					
Model	Candidate	Score beta	Interaction beta	Interaction se	Interaction P-value
PRS_279	SNPs-279	0.3143	0.0012	0.0009	0.2121
PRS_(p+t)	5e-8_0.2	0.2455	0.0015	0.0009	0.0767
PTRS_GTEx (Lung)	279-nearby	0.1501	0.0001	0.0009	0.8872
PTRS_GTEx (WB)	TWAS-qval	0.0225	0.0019	0.0009	0.0355
PTRS_MESA (monocytes)	TWAS-nominal	0.1309	0.0014	0.0009	0.1119
Severe COPD					
PRS_279	SNPs-279	0.2002	0.0052	0.0015	0.0007
PRS_(p+t)	5e-8_0.2	0.2458	0.004	0.0014	0.0042
PTRS_GTEx (Lung)	279-nearby	0.1151	0.0025	0.0016	0.1143
PTRS_GTEx (WB)	TWAS-qval	0.0066	-0.004	0.0015	0.0074
PTRS_MESA (monocytes)	TWAS-nominal	0.1441	0.0026	0.0015	0.0789
FEV1/FVC					
PRS_279	SNPs-279	-1.2297	-0.0075	0.0021	0.0004
PRS_(p+t)	5e-8_0.2	-0.9944	-0.0078	0.0020	0.0001
PTRS_GTEx (Lung)	Coloc0.1	-0.5938	-0.0032	0.0021	0.1384

PTRS_GTEx (WB)	Coloc0.1	-0.4156	-0.0027	0.0022	0.2189
PTRS_MESA (monocytes)	TWAS-nominal	-0.5215	-0.0069	0.0022	0.0017
FEV1					
PRS_279	SNPs-279	-0.0703	-0.0003	0.0001	0.0288
PRS_(p+t)	5e-4_0.8	-0.0571	-0.0004	0.0001	0.0024
PTRS_GTEx (Lung)	TWAS-nominal	-0.0491	-0.0001	0.0001	0.4462
PTRS_GTEx (WB)	TWAS-nominal	-0.0272	-0.0002	0.0001	0.1013
PTRS_MESA (monocytes)	TWAS-nominal	-0.0397	-0.0002	0.0001	0.1249

Table S14: Risk score by smoking pack-years interaction analysis for the best performing risk score candidates in population/family-based cohorts. PRS_279, PRS derived by previously published 279 variants for FEV1/FVC ratio or FEV1; PRS_(p+t), PRS derived by pruning and thresholding, a range of p-value and pairwise correlation thresholds were used to create candidates (5e-4_0.8 and 5e-8_0.2) ; 279-nearyby, PTRS derived by genes nearby previously published 279 variants for FEV1/FVC ratio or FEV1; TWAS-nominal and TWAS-qval, PTRS derived by genes passing TWAS p-value threshold of 0.05 and q-value respectively; Coloc0.1, PTRS derived by genes with regional colocalization probability greater than 0.1.

		Moderate-to-Severe COPD			Severe COPD			FEV1/FVC and FEV1
Study	Smoking Pack-Years	N	Control	Case	N	Control	Case	N
COPDGene_NHW	>=0	5183	2122	3061	3729	2122	1607	6609
	[0,20)	559	400	159	467	400	67	715
	[20,40)	1721	939	782	1289	939	350	2245
	>=40	2903	783	2120	1973	783	1190	3649
COPDGene_AA	>=0	2504	1584	920	1999	1584	415	3258
	[0,20)	434	307	127	362	307	55	612
	[20,40)	1137	769	368	924	769	155	1450
	>=40	933	508	425	713	508	205	1196
SPIROMICS_NHW	>=0	1284	346	938	803	346	457	1535
	[0,20)	66	63	3	63	63	0	74
	[20,40)	368	143	225	253	143	110	457
	>=40	850	140	710	487	140	347	1004
SPIROMICS_AA	>=0	316	139	177	229	139	90	369
	[0,20)	24	23	1	23	23	0	25
	[20,40)	162	74	88	118	74	44	188
	>=40	130	42	88	88	42	46	156

Table S15: Sample size by smoker group for four traits in COPD-enriched studies. NHW, Non-Hispanic Whites; AA, African Americans.

Trait	Effect	Beta	se	P-value
Moderate-to-Severe COPD	PRS	0.3181	0.0238	8.50E-41
	PTRS	0.0571	0.0238	0.0163
	Interaction	0.0075	0.0212	0.7249
Severe COPD	PRS	0.3248	0.0502	9.90E-11
	PTRS	0.085	0.0511	0.0961
	Interaction	0.0049	0.0446	0.9124
FEV1/FVC	PRS	-1.2434	0.0426	9.43E-185
	PTRS	-0.4556	0.0425	9.71E-27
	Interaction	0.0187	0.0415	0.6523
FEV1	PRS	-0.0666	0.0028	3.68E-127
	PTRS	-0.0369	0.003	9.17E-36
	Interaction	0.0013	0.0026	0.6190

Table S20: Risk score interaction analysis in population/family-based cohorts. PRS, PRS_279: SNPs-279 that was derived by previously published 279 variants for FEV1/FVC ratio or FEV1; PTRS, PTRS_GTEx(Lung): 279-nearby that was derived by genes nearby previously published 279 variants for FEV1/FVC ratio or FEV1.

	Smoking Pack-Years≥0			Smoking Pack-Years: [0,20)		
	AUC	L95_AUC	U95_AUC	AUC	L95_AUC	U95_AUC
Clinical factors	0.807	0.798	0.815	0.742	0.729	0.754
PRS	0.579	0.567	0.59	0.589	0.574	0.605
PTRS	0.549	0.537	0.56	0.548	0.532	0.564
Unweighted Sum	0.58	0.568	0.591	0.586	0.571	0.602
Weighted Sum	0.582	0.571	0.593	0.592	0.576	0.608
Clinical+PRS	0.813	0.805	0.821	0.753	0.741	0.766
Clinical+PTRS	0.808	0.8	0.816	0.744	0.732	0.756
Clinical+Unweighted Sum	0.812	0.804	0.82	0.751	0.739	0.764
Clinical+Weighted Sum	0.813	0.805	0.822	0.754	0.741	0.766
	Smoking Pack-Years: [20,40)			Smoking Pack-Years≥40		
	AUC	L95_AUC	U95_AUC	AUC	L95_AUC	U95_AUC
Clinical factors	0.647	0.623	0.672	0.63	0.604	0.657
PRS	0.585	0.558	0.611	0.598	0.571	0.625
PTRS	0.547	0.52	0.574	0.542	0.515	0.569
Unweighted Sum	0.581	0.554	0.608	0.585	0.559	0.612
Weighted Sum	0.587	0.56	0.613	0.599	0.572	0.625
Clinical+PRS	0.667	0.642	0.691	0.659	0.634	0.684
Clinical+PTRS	0.651	0.626	0.676	0.635	0.61	0.661
Clinical+Unweighted Sum	0.662	0.637	0.686	0.651	0.625	0.676
Clinical+Weighted Sum	0.667	0.643	0.691	0.659	0.634	0.684

Table S21: Prediction accuracy for Moderate-to-Severe COPD in population/family-based cohorts from models using clinical risk factors alone (including age, sex, race, and smoking pack-years), risk score alone, and the combination of clinical risk factors and risk score. PRS, PRS_279: SNPs-279 that was derived by previously published 279 variants for FEV1/FVC ratio; PTRS, PTRS_GTEx(Lung): 279-nearby that was derived by genes nearby previously published 279 variants for FEV1/FVC ratio; Unweighted Sum and Weighted Sum refer to the direct summation and the weighted summation of PTRS and PRS respectively; L95_AUC and U95_AUC are lower and upper 95% CIs of AUC respectively.

	Smoking Pack-Years≥0			Smoking Pack-Years: [0,20)		
	AUC	L95_AUC	U95_AUC	AUC	L95_AUC	U95_AUC
Clinical factors	0.854	0.837	0.87	0.762	0.731	0.792
PRS	0.582	0.558	0.607	0.545	0.502	0.587
PTRS	0.554	0.53	0.579	0.527	0.486	0.568
Unweighted Sum	0.587	0.562	0.612	0.545	0.503	0.587
Weighted Sum	0.588	0.563	0.613	0.547	0.505	0.589
Clinical+PRS	0.853	0.836	0.87	0.761	0.731	0.792
Clinical+PTRS	0.853	0.836	0.87	0.761	0.731	0.792
Clinical+Unweighted Sum	0.853	0.836	0.87	0.761	0.731	0.791
Clinical+Weighted Sum	0.853	0.836	0.87	0.761	0.73	0.791
	Smoking Pack-Years: [20,40)			**Smoking Pack-Years≥40**		
	AUC	L95_AUC	U95_AUC	AUC	L95_AUC	U95_AUC
Clinical factors	0.703	0.657	0.748	0.693	0.658	0.727
PRS	0.659	0.607	0.711	0.617	0.577	0.657
PTRS	0.574	0.515	0.633	0.556	0.517	0.596
Unweighted Sum	0.639	0.583	0.695	0.612	0.573	0.651
Weighted Sum	0.66	0.608	0.712	0.621	0.581	0.66
Clinical+PRS	0.743	0.698	0.789	0.723	0.688	0.758
Clinical+PTRS	0.71	0.663	0.757	0.699	0.664	0.734
Clinical+Unweighted Sum	0.733	0.685	0.781	0.717	0.682	0.752
Clinical+Weighted Sum	0.743	0.698	0.789	0.723	0.688	0.758

Table S22: Prediction accuracy for Severe COPD in population/family-based cohorts from models using clinical risk factors alone (including age, sex, race, and smoking pack-years), risk score alone, and the combination of clinical risk factors and risk score. PRS, PRS_279: SNPs-279 that was derived by previously published 279 variants for FEV1/FVC ratio; PTRS, PTRS_GTEx(Lung): 279-nearby that was derived by genes nearby previously published 279 variants for FEV1/FVC ratio; Unweighted Sum and Weighted Sum refer to the direct summation and the weighted summation of PTRS and PRS respectively; L95_AUC and U95_AUC are lower and upper 95% CIs of AUC respectively.

Supplemental Methods

Phenotype harmonization

Phenotype harmonization, data management, sample-identity QC, and general study coordination, were provided by the TOPMed Data Coordinating Center (3R01HL-120393-02S1; contract HHSN268201800001I). Phenotype harmonization for pulmonary traits was contributed by the NHLBI Pooled Cohorts Study with funding from NIH/NHLBI R21 HL121457, R21 HL129924, K23 HL130627, R01 HL077612.

The NHLBI Pooled Cohorts Study (PCS)[1] harmonized and pooled data from nine large US epidemiologic cohorts that conducted lung function assessments over the last four decades. Additionally, data on self-administered questionnaires, with detailed questions regarding tobacco consumption, past medical history, medications, respiratory symptoms, lipids, renal biomarkers, etc., were also harmonized. Data on CLRD events was harmonized using either adjudicated CLRD hospitalizations or ICD data for all hospitalizations occurring over follow-up.

Among the cohorts included in the NHLBI PCS, the follow cohorts were also included in our TOPMed WGS analysis, for which we utilized harmonized data sets from NHLBI PCS: Atherosclerosis Risk in Communities (ARIC) study; Cardiovascular Health Study (CHS); Coronary Artery Risk Development in Young Adults (CARDIA); Framingham Heart Study (FHS); Hispanic Community Health Study/Study of Latinos (HCHS/SOL); Jackson Heart Study (JHS); and Multi-Ethnic Study of Atherosclerosis (MESA). For the purpose of phenotype harmonization in TOPMed, harmonized data sets for each of the participating cohorts were provided by the NHLBI Pooled Cohorts Study for analyses in the current WGS analysis. We note that a subset of the ARIC participants was later recruited into JHS. For TOPMed purposes, participants in the ARIC-JHS overlap group were not included as part of Exam 4 in ARIC. For JHS, ARIC participants were not excluded. For TOPMed studies not included in the NHLBI Pooled Cohorts Study (Cleveland Family Study (CFS); Genetic Epidemiology of COPD (COPDGene) and Sub-Populations and Intermediate Outcome Measures in COPD Study (SPIROMICS)), phenotype harmonization was conducted separately by each of the participating cohorts, following as closely as possible with the NHLBI Pooled Cohorts Study variable definitions and procedures.

Study descriptions: population/family-based cohorts

The Atherosclerosis Risk in Communities Study (ARIC)

The ARIC study is a population-based prospective cohort study of cardiovascular disease sponsored by the National Heart, Lung, and Blood Institute (NHLBI). ARIC included 15,792 individuals, predominantly European American and African American, aged 45-64 years at baseline (1987-89), chosen by probability sampling from four US

communities. Cohort members completed three additional triennial follow-up examinations, a fifth exam in 2011-2013, a sixth exam in 2016-2017, and a seventh exam in 2018-2019. The ARIC study has been described in detail previously (The ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) study: Design and objectives. American Journal of Epidemiology 1989; 129:687-702).

At each visit, spirometry testing protocols were standardized across the four ARIC field centers, calibration checks were performed daily, and the standardization of data collection and management was coordinated across field centers by a single pulmonary function reading center. Each participant's best FEV_1 and FVC of three acceptable maneuvers, based on the centralized expert review, was used for analysis[3]. For the current cross-sectional WGS analysis, data from the most recent spirometry exam were utilized for participants having multiple longitudinal measures.

The Coronary Artery Risk Development in Young Adults (CARDIA)

During 1985 -1986, CARDIA recruited 5,115 black and white men and women, aged 18 to 30 years, from the general population at Birmingham, Alabama; Chicago, Illinois; and Minneapolis, Minnesota; and from the membership of the Oakland Kaiser-Permanente Health Plan in Oakland, California. The participants were selected so that there would be approximately the same number of people in subgroups of race, gender, education (high school or less and more than high school) and age (18-24 and 25-30) in each of 4 centers. Detailed methods, instruments, and quality control procedures are described at the CARDIA website (http://www.cardia.dopm.uab.edu/ex_mt.htm) and in other published reports[4,5]. Spirometric pulmonary function testing was performed using the Collins survey 8-liter water-sealed spirometer and the Eagle II microprocessor (Warren E. Collins, Inc., Braintree, MA) in a sitting position with noseclips, as per the 1979 American Thoracic Society criteria[6]. Specifically, each subject performed a minimum of three trials with expirations recorded to the FVC plateau, which occurs after six seconds of expiration in adult males and was maintained for at least one second before terminating the forced expiratory maneuver. If, at the end of the three trials, there were at least three acceptable tracings, and with the maximum FVC and FEV1 reproduced to within 5% or 100 mL, whichever is greater, no more trials were performed. For the current cross-sectional WGS analysis, data from the most recent spirometry exam were utilized for participants having multiple longitudinal measures.

The Cardiovascular Health Study (CHS)

CHS is a population-based cohort study of risk factors for coronary heart disease and stroke in adults ≥65 years conducted across four field centers7. The original predominantly European ancestry cohort of 5,201 persons was recruited in 1989-1990 from random samples of the Medicare eligibility lists; subsequently, an additional

predominantly African-American cohort of 687 persons was enrolled in 1992-1993 for a total sample of 5,888. Blood samples were drawn from all participants at their baseline examination and DNA was subsequently extracted from available samples. Pulmonary function testing was conducted at the 1989-1990 visit and follow-up visits four and seven years later. The spirometry procedures for pulmonary function testing have been previously described[8,9]. Briefly, spirometry technicians were centrally trained and certified prior to recruitment of participants. A standard spirometry system, including a Collins Survey I water-seal spirometer (Collins Medical, Inc., Braintree, Massachusetts) and software from S&M Instruments (Doylestown, Pennsylvania), was used by technicians at all four recruitment centers. Stringent quality assurance procedures for spirometry testing exceeded ATS recommendations[8]. For the current cross-sectional WGS analysis, data from the most recent spirometry exam were utilized for participants having multiple longitudinal measures.

European ancestry and African American ancestry CHS participants that had been selected for inclusion in the second phase of the TOPMed sequencing program were included in our discovery analyses. CHS was approved by institutional review committees at each field center and individuals in the present analysis had available DNA and gave informed consent including consent to use of genetic information for the study of cardiovascular disease.

The Cleveland Family Study (CFS)

CFS is a family-based longitudinal study that includes participants with laboratory diagnosed sleep apnea, their family members and neighborhood control families followed between 1990 and 2006. Four examinations over 16 years provided measurements of sleep apnea with overnight polysomnography, anthropometry, and other related phenotypes, as detailed previously[10]. At each exam, forced vital capacity (FVC) and forced expiratory flow (FEV1) was obtained using a calibrated spirometer (Multi-Spiro). While seated, participants were encouraged to perform between 5-8 maneuvers to obtain 3 curves that met ATS standards for acceptability and reproducibility. For the current cross-sectional WGS analysis, data from the most recent spirometry exam were utilized for participants having multiple longitudinal measures.

The Framingham Heart Study (FHS)

The Original Cohort of the Framingham Study was established between 1948 and 1952 as a random sample of 5,209 adult residents of the town of Framingham, Massachusetts. Between 1971 and 1975, the Framingham Study was expanded to include a second generation, the Offspring Cohort, comprising 5,124 adults who were the offspring, or spouses of the offspring, of Original Cohort participants[11]. The Offspring

Cohort has returned for examinations approximately every 4 years since enrollment, and spirometry data are available for the 3rd, 5th, 6th, 7th, ,8th, and 9th examinations.

Spirometry for the Offspring Cohort 3rd examination (1983-87) was performed with a Collins Survey II spirometer interfaced with an Eagle II microprocessor (Warren E. Collins, Inc., Braintree, MA). Spirometry for the 5th (1991-95), 6th (1995-98), and 7th (1998-2001) examinations were performed with a Collins Survey II spirometer interfaced with a personal computer equipped with software developed by S & M Instruments (Doylestown, PA) and adapted for use in epidemiologic studies. Spirometry for the 8th (2005-08) and 9th (2011-14) examinations was performed with the Collins Comprehensive Pulmonary Laboratory (CPL) system with Collins 2000 Plus/SQL Software (Nspire Health, Inc., Longmont, CO). Spirometry was performed in accordance with contemporaneous guidelines of the American Thoracic Society. For the current cross-sectional WGS analysis, data from the first available spirometry exam were utilized for participants having multiple longitudinal measures.

The Hispanic Community Health Study/Study of Latinos (HCHS/SOL)

HCHS/SOL is a community-based cohort study of 16,415 self-identified Hispanic/Latino persons aged 18 to 74 years at baseline recruited from four U.S. communities (Bronx NY, Chicago IL, San Diego CA, Miami FL). The baseline clinic visit took place in 2008-2011 and included spirometry data collection. The study design, cohort recruitment[12,13] and baseline clinical examination[14] have been previously described. Institutional Review Boards at each field center approved study protocols, and written informed consent was obtained from all participants. For the current cross-sectional WGS analysis, only baseline spirometry data were utilized.

The Jackson Heart Study (JHS)

JHS is a large, population-based observational study evaluating the etiology of cardiovascular diseases and related disorders among African Americans residing in the three counties (Hinds, Madison, and Rankin) that make up the Jackson, Mississippi metropolitan area[15,16]. Data and biologic materials have been collected from 5,301 participants, including a nested family cohort of 1,498 members of 264 families. The age at enrollment for the unrelated cohort was 35-84 years; the family cohort included related individuals >21 years old. During a baseline examination (2000-2004) and two follow-up examinations (2005-2008 and 2009-2012), participants provided extensive medical and social history, had an array of physical and biochemical measurements and diagnostic procedures, and provided blood for genomic DNA[17]. The study population is characterized by a high prevalence of diabetes, hypertension, obesity, and related disorders. Annual follow-up interviews and cohort surveillance are ongoing. For the current cross-sectional WGS analysis, only baseline spirometry data were utilized.

The Multi-Ethnic Study of Atherosclerosis (MESA)

MESA is a longitudinal study of subclinical cardiovascular disease and risk factors that predict progression to clinically overt cardiovascular disease or progression of the subclinical disease[18]. Between 2000 and 2002, MESA recruited 6,814 men and women 45 to 84 years of age from Forsyth County, North Carolina; New York City; Baltimore; St. Paul, Minnesota; Chicago; and Los Angeles. Exclusion criteria were clinical cardiovascular disease, weight exceeding 136 kg (300 lb.), pregnancy, and impediment to long-term participation. The MESA Lung Study performed spirometry following the 2005 ATS/ERS guidelines in a subset of the MESA Study, as previously described[19]. All participants provided informed consent and the protocols of MESA were approved by the IRBs of collaborating institutions and the National Heart, Lung and Blood Institute. For the current cross-sectional WGS analysis, data from the earliest spirometry exam were utilized for participants having multiple longitudinal measures.

Study descriptions: COPD-enriched studies

Genetic Epidemiology of COPD (COPDGene)

COPDGene[20] is a multi-center observational cohort for epidemiologic and genetic study of over 10,000 subjects (2/3 non-Hispanic White and 1/3 African Americans) with at least 10 pack-years of cigarette smoking with and without COPD. All subjects underwent extensive phenotyping, including lung function, chest CT phenotypes (including emphysema and expiratory gas trapping). Pre- and post-bronchodilator spirometry measures were obtained using a standardized protocol and spirometer (ndd EasyOne Spirometer, Zurich, Switzerland). All study sites obtained local IRB approval to enroll participants and all subjects provided written informed consent. For the current cross-sectional WGS analysis, only baseline spirometry data were utilized.

Sub-Populations and Intermediate Outcome Measures in COPD Study (SPIROMICS)

SPIROMICS is a prospective cohort study that enrolled 3,200 participants into four strata (non-smokers, smokers without airflow obstruction, mild/moderate COPD, and severe COPD). Participants may be enrolled in concurrent observational studies, excluding the COPDGene Study[20], which facilitates combined analyses between SPIROMICS and COPDGene. The Institutional Review Boards/Ethics Committees of all the cooperating institutions have approved the study protocols. SPIROMICS participants were 40 to 80 years of age with a smoking history ≥20 pack-years with COPD (GOLD spirometric grades 1-4) and without COPD21,22. Participants were comprehensively characterized with annual pre- and post-bronchodilator lung function measures for up to 3 years, computed tomography scans, and standardized

questionnaires[21]. For the current cross-sectional WGS analysis, only baseline spirometry data were utilized.

Quality control of samples and variants included in TOPMed cohorts for analyses

For the pooled cohort, we first removed subjects who failed sample-level quality control. The filters included checking for pedigree errors, discrepancies between self-reported and genetic sex. Details regarding the quality control are described on the website (https://www.nhlbiwgs.org/topmed-whole-genome-sequencing-methods-freeze-8). In addition, only one subject from each pair of duplicates was kept. There were 29,381 subjects from population/family-based cohorts and 11,771 subjects from COPD-enriched studies that passed the filtering.

For site-level quality control, variants were removed based on Mendelian discordance, a support vector machine (SVM) quality filter and excess heterozygosity filter. Details are described on the online document, https://www.nhlbiwgs.org/topmed-whole-genome-sequencing-methods-freeze-8.

Acknowledgments for TOPMed whole genome sequencing

Whole genome sequencing (WGS) for the Trans-Omics in Precision Medicine (TOPMed) program was supported by the National Heart, Lung and Blood Institute (NHLBI). WGS for "NHLBI TOPMed: Atherosclerosis Risk in Communities (ARIC)" (phs001211) was performed at the Baylor College of Medicine Human Genome Sequencing Center (HHSN268201500015C and 3U54HG003273-12S2) and the Broad Institute for MIT and Harvard (3R01HL092577-06S1). WGS for "NHLBI TOPMed: Coronary Artery Risk Development in Young Adults (CARDIA)" (phs001612) was performed at the Baylor College of Medicine Human Genome Sequencing Center (HHSN268201600033I). WGS for "NHLBI TOPMed: The Cleveland Family Study (CFS)" (phs000954) was performed at the University of Washington Northwest Genomics Center (3R01HL098433-05S1 and HHSN268201600032I). WGS for "NHLBI TOPMed: Cardiovascular Health Study (CHS)" (phs001368) was performed at the Baylor College of Medicine Human Genome Sequencing Center (HHSN268201600033I) and Broad Institute Genomics Platform (HHSN268201600034I). WGS for "NHLBI TOPMed: Whole Genome Sequencing and Related Phenotypes in the Framingham Heart Study (FHS)" (phs000974) was performed at the Broad Institute Genomics Platform (3U54HG003067-12S2and 3R01HL092577-06S1). WGS for "NHLBI TOPMed: Hispanic Community Health Study/Study of Latinos (HCHS/SOL)" (phs001395) was performed at the Baylor College of Medicine Human Genome Sequencing Center (HHSN268201600033I). WGS for "NHLBI TOPMed: The Jackson Heart Study (JHS)" (phs000964) was performed at the University of Washington Northwest Genomics Center (HHSN268201100037C). WGS for "NHLBI TOPMed: Multi-

Ethnic Study of Atherosclerosis (MESA)" (phs001416) was performed at the Broad Institute Genomics Platform (3U54HG003067-13S1 and HHSN268201600034I). WGS for "NHLBI TOPMed: Genetic Epidemiology of COPD (COPDGene) in the TOPMed Program" (phs000951) was performed at the University of Washington Northwest Genomics Center (3R01 HL089856-08S1) and the Broad Institute Genomics Platform (HHSN268201500014C). WGS for "NHLBI TOPMed: SubPopulations and InteRmediate Outcome Measures in COPD Study (SPIROMICS)" (phs001927) was performed at Broad Institute Genomics Platform (HHSN268201600034I). Centralized read mapping and genotype calling, along with variant quality metrics and filtering were provided by the TOPMed Informatics Research Center (3R01HL-117626-02S1; contract HHSN268201800002I).

Study specific acknowledgments: population/family-based cohorts

The Atherosclerosis Risk in Communities Study

The Genome Sequencing Program (GSP) was funded by the National Human Genome Research Institute (NHGRI), the National Heart, Lung, and Blood Institute (NHLBI), and the National Eye Institute (NEI). The GSP Coordinating Center (U24 HG008956) contributed to cross-program scientific initiatives and provided logistical and general study coordination. The Centers for Common Disease Genomics (CCDG) program was supported by NHGRI and NHLBI, and whole genome sequencing was performed at the Baylor College of Medicine Human Genome Sequencing Center (UM1 HG008898 and R01HL059367). The Atherosclerosis Risk in Communities study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services (contract numbers HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I and HHSN268201700005I). The authors thank the staff and participants of the ARIC study for their important contributions. SJL is supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (ZO1 ES043012).

The Coronary Artery Risk Development in Young Adults Study

The Coronary Artery Risk Development in Young Adults study is conducted at the University of Alabama at Birmingham (HHSN268201800005I & HHSN268201800007I), Northwestern University (HHSN268201800003I), University of Minnesota (HHSN268201800006I), and Kaiser Foundation Research Institute (HHSN268201800004I). CARDIA was also partially supported by the Intramural Research Program of the National Institute on Aging (NIA) and an intra-agency agreement between NIA and NHLBI (AG0005).

The Cleveland Family Study

The Cleveland Family Study and SR were supported by NIH grants HL 046389, HL113338, and 1R35HL135818. BC is supported by the NIH grant K01 HL135405 and an American Thoracic Society Foundation Unrestricted Grant (Sleep) (http://foundation.thoracic.org).

The Cardiovascular Health Study

This Cardiovascular Health Study (CHS) research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, HHSN268200960009C, HHSN268201800001C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, 75N92021D00006; and NHLBI grants U01HL080295, U01HL130114, R01HL087652, R01HL105756, R01HL103612, R01HL085251, and R01HL120393 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The Framingham Heart Study

The Framingham Heart Study (FHS) acknowledges the support of contracts NO1-HC-25195, HHSN268201500001I, and 75N92019D00031 from the National Heart, Lung and Blood Institute and grant supplement R01 HL092577-06S1 for this research. We also acknowledge the dedication of the FHS study participants without whom this research would not be possible. Dr. Vasan is supported in part by the Evans Medical Foundation and the Jay and Louis Coffman Endowment from the Department of Medicine, Boston University School of Medicine.

The Hispanic Community Health Study/Study of Latinos

The authors thank the staff and participants of HCHS/SOL for their important contributions. The Hispanic Community Health Study/Study of Latinos is a collaborative study supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (HHSN268201300001I / N01-HC-65233), University of Miami (HHSN268201300004I / N01-HC-65234), Albert Einstein College of Medicine

(HHSN268201300002I / N01-HC-65235), University of Illinois at Chicago – HHSN268201300003I / N01-HC-65236 Northwestern Univ), and San Diego State University (HHSN268201300005I / N01-HC-65237). The following Institutes/Centers/Offices have contributed to the HCHS/SOL through a transfer of funds to the NHLBI: National Institute on Minority Health and Health Disparities, National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Neurological Disorders and Stroke, NIH Institution-Office of Dietary Supplements. The Genetic Analysis Center at the University of Washington was supported by NHLBI and NIDCR contracts (HHSN268201300005C AM03 and MOD03).

The Jackson Heart Study

The Jackson Heart Study (JHS) is supported and conducted in collaboration with Jackson State University (HHSN268201800013I), Tougaloo College (HHSN268201800014I), the Mississippi State Department of Health (HHSN268201800015I) and the University of Mississippi Medical Center (HHSN268201800010I, HHSN268201800011I and HHSN268201800012I) contracts from the National Heart, Lung, and Blood Institute (NHLBI) and the National Institute on Minority Health and Health Disparities (NIMHD). The authors also wish to thank the staffs and participants of the JHS.

The Multi-Ethnic Study of Atherosclerosis

MESA and the MESA SHARe projects are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts 75N92020D00001, HHSN268201500003I, N01-HC-95159, 75N92020D00005, N01-HC-95160, 75N92020D00002, N01-HC-95161, 75N92020D00003, N01-HC-95162, 75N92020D00006, N01-HC-95163, 75N92020D00004, N01-HC-95164, 75N92020D00007, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-001079, and UL1-TR-001420. The MESA Lung Study was funded by R01-HL077612 and R01-HL093081. Funding for genotyping was provided by NHLBI Contract N02-HL-64278. Also supported in part by NHLBI CHARGE Consortium Contract HL105756. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. Infrastructure for the CHARGE Consortium is supported in part by the National Heart, Lung, and Blood Institute (NHLBI) grant R01HL105756.

Study specific acknowledgments: COPD-enriched studies

Genetic Epidemiology of COPD (COPDGene)

The project described was supported by Award Number U01 HL089897 and Award Number U01 HL089856 from the National Heart, Lung, and Blood Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, and Blood Institute or the National Institutes of Health.

COPD Foundation Funding

COPDGene is also supported by the COPD Foundation through contributions made to an Industry Advisory Board that has included AstraZeneca, Bayer Pharmaceuticals, Boehringer-Ingelheim, Genentech, GlaxoSmithKline, Novartis, Pfizer, and Sunovion.

COPDGene® Investigators – Core Units

Administrative Center: James D. Crapo, MD (PI); Edwin K. Silverman, MD, PhD (PI); Barry J.Make, MD; Elizabeth A. Regan, MD, PhD

Genetic Analysis Center: Terri H. Beaty, PhD; Peter J. Castaldi, MD, MSc; Michael H. Cho,MD, MPH; Dawn L. DeMeo, MD, MPH; Adel El Boueiz, MD, MMSc; Marilyn G. Foreman,MD, MS; Auyon Ghosh, MD; Lystra P. Hayden, MD, MMSc; Craig P. Hersh, MD, MPH; Jacqueline Hetmanski, MS; Brian D. Hobbs, MD, MMSc; John E. Hokanson, MPH, PhD; Wonji Kim, PhD; Nan Laird, PhD; Christoph Lange, PhD; Sharon M. Lutz, PhD; Merry-Lynn McDonald, PhD; Dmitry Prokopenko, PhD; Matthew Moll, MD, MPH; Jarrett Morrow, PhD; Dandi Qiao, PhD; Elizabeth A. Regan, MD, PhD; Aabida Saferali, PhD; Phuwanat Sakornsakolpat, MD; Edwin K. Silverman, MD, PhD; Emily S. Wan, MD; Jeong Yun, MD, MPH

Imaging Center: Juan Pablo Centeno; Jean-Paul Charbonnier, PhD; Harvey O. Coxson, PhD; Craig J. Galban, PhD; MeiLan K. Han, MD, MS; Eric A. Hoffman, Stephen Humphries, PhD; Francine L. Jacobson, MD, MPH; Philip F. Judy, PhD; Ella A. Kazerooni, MD; Alex Kluiber; David A. Lynch, MB; Pietro Nardelli, PhD; John D. Newell, Jr., MD; Aleena Notary; Andrea Oh, MD; Elizabeth A. Regan, MD, PhD; James C. Ross, PhD; Raul San Jose Estepar, PhD; Joyce Schroeder, MD; Jered Sieren; Berend C. Stoel, PhD; Juerg Tschirren, PhD; Edwin Van Beek, MD, PhD; Bram van Ginneken, PhD; Eva van Rikxoort, PhD; Gonzalo Vegas Sanchez-Ferrero, PhD; Lucas Veitel; George R. Washko, MD; Carla G. Wilson, MS

PFT QA Center, Salt Lake City, UT: Robert Jensen, PhD

Data Coordinating Center and Biostatistics, *National Jewish Health, Denver, CO*: Douglas Everett, PhD; Jim Crooks, PhD; Katherine Pratte, PhD; Matt Strand, PhD; Carla G. Wilson, MS
Epidemiology Core, *University of Colorado Anschutz Medical Campus, Aurora, CO*: John E.Hokanson, MPH, PhD; Erin Austin, PhD; Gregory Kinney, MPH, PhD; Sharon M. Lutz, PhD; Kendra A. Young, PhD

Mortality Adjudication Core: Surya P. Bhatt, MD; Jessica Bon, MD; Alejandro A. Diaz, MD, MPH; MeiLan K. Han, MD, MS; Barry Make, MD; Susan Murray, ScD; Elizabeth Regan, MD; Xavier Soler, MD; Carla G. Wilson, MS

Biomarker Core: Russell P. Bowler, MD, PhD; Katerina Kechris, PhD; Farnoush Banaei-Kashani, Ph.D

COPDGene® Investigators – Clinical Centers

Ann Arbor VA: Jeffrey L. Curtis, MD; Perry G. Pernicano, MD

Baylor College of Medicine, Houston, TX: Nicola Hanania, MD, MS; Mustafa Atik, MD; Aladin Boriek, PhD; Kalpatha Guntupalli, MD; Elizabeth Guy, MD; Amit Parulekar, MD
Brigham and Women's Hospital, Boston, MA: Dawn L. DeMeo, MD, MPH; Craig Hersh, MD, MPH; Francine L. Jacobson, MD, MPH; George Washko, MD

Columbia University, New York, NY: R. Graham Barr, MD, DrPH; John Austin, MD; Belinda D'Souza, MD; Byron Thomashow, MD

Duke University Medical Center, Durham, NC: Neil MacIntyre, Jr., MD; H. Page McAdams, MD; Lacey Washington, MD

HealthPartners Research Institute, Minneapolis, MN: Charlene McEvoy, MD, MPH; Joseph Tashjian, MD

Johns Hopkins University, Baltimore, MD: Robert Wise, MD; Robert Brown, MD; Nadia N. Hansel, MD, MPH; Karen Horton, MD; Allison Lambert, MD, MHS; Nirupama Putcha, MD, MHS

Lundquist Institute for Biomedical Innovation at Harbor UCLA Medical Center, Torrance, CA: Richard Casaburi, PhD, MD; Alessandra Adami, PhD; Matthew Budoff, MD; Hans Fischer, MD; Janos Porszasz, MD, PhD; Harry Rossiter, PhD; William Stringer, MD

Michael E. DeBakey VAMC, Houston, *TX*: Amir Sharafkhaneh, MD, PhD; Charlie Lan, DO
Minneapolis VA: Christine Wendt, MD; Brian Bell, MD; Ken M. Kunisaki, MD, MS

Morehouse School of Medicine, Atlanta, GA: Eric L. Flenaugh, MD; Hirut Gebrekristos, PhD; Mario Ponce, MD; Silanath Terpenning, MD; Gloria Westney, MD, MS

National Jewish Health, Denver, CO: Russell Bowler, MD, PhD; David A. Lynch, MB

Reliant Medical Group, Worcester, MA: Richard Rosiello, MD; David Pace, MD

Temple University, Philadelphia, PA: Gerard Criner, MD; David Ciccolella, MD; Francis Cordova, MD; Chandra Dass, MD; Gilbert D'Alonzo, DO; Parag Desai, MD; Michael Jacobs, PharmD; Steven Kelsen, MD, PhD; Victor Kim, MD; A. James Mamary, MD; Nathaniel Marchetti, DO; Aditi Satti, MD; Kartik Shenoy, MD; Robert M. Steiner, MD; Alex Swift, MD; Irene Swift, MD; Maria Elena Vega-Sanchez, MD

University of Alabama, Birmingham, AL: Mark Dransfield, MD; William Bailey, MD; Surya P. Bhatt, MD; Anand Iyer, MD; Hrudaya Nath, MD; J. Michael Wells, MD

University of California, San Diego, CA: Douglas Conrad, MD; Xavier Soler, MD, PhD; Andrew Yen, MD

University of Iowa, Iowa City, IA: Alejandro P. Comellas, MD; Karin F. Hoth, PhD; John Newell, Jr., MD; Brad Thompson, MD

University of Michigan, Ann Arbor, MI: MeiLan K. Han, MD MS; Ella Kazerooni, MD MS; Wassim Labaki, MD MS; Craig Galban, PhD; Dharshan Vummidi, MD

University of Minnesota, Minneapolis, MN: Joanne Billings, MD; Abbie Begnaud, MD; Tadashi Allen, MD

University of Pittsburgh, Pittsburgh, PA: Frank Sciurba, MD; Jessica Bon, MD; Divay Chandra, MD, MSc; Joel Weissfeld, MD, MPH

University of Texas Health, San Antonio, San Antonio, TX: Antonio Anzueto, MD; Sandra Adams, MD; Diego Maselli-Caceres, MD; Mario E. Ruiz, MD; Harjinder Singh

Sub-Populations and Intermediate Outcome Measures in COPD Study (SPIROMICS)

The authors thank the SPIROMICS participants and participating physicians, investigators and staff for making this research possible. More information about the study and how to access SPIROMICS data is available at www.spiromics.org. The authors would like to acknowledge the University of North Carolina at Chapel Hill BioSpecimen Processing Facility for sample processing, storage, and sample disbursements (http://bsp.web.unc.edu/).

We would like to acknowledge the following current and former investigators of the SPIROMICS sites and reading centers: Neil E Alexis, MD; Wayne H Anderson, PhD; Mehrdad Arjomandi, MD; Igor Barjaktarevic, MD, PhD; R Graham Barr, MD, DrPH; Patricia Basta, PhD; Lori A Bateman, MSc; Surya P Bhatt, MD; Eugene R Bleecker, MD; Richard C Boucher, MD; Russell P Bowler, MD, PhD; Stephanie A Christenson, MD; Alejandro P Comellas, MD; Christopher B Cooper, MD, PhD; David J Couper, PhD; Gerard J Criner, MD; Ronald G Crystal, MD; Jeffrey L Curtis, MD; Claire M Doerschuk, MD; Mark T Dransfield, MD; Brad Drummond, MD; Christine M Freeman, PhD; Craig Galban, PhD; MeiLan K Han, MD, MS; Nadia N Hansel, MD, MPH; Annette T Hastie, PhD; Eric A Hoffman, PhD; Yvonne Huang, MD; Robert J Kaner, MD; Richard E Kanner, MD; Eric C Kleerup, MD; Jerry A Krishnan, MD, PhD; Lisa M LaVange, PhD; Stephen C Lazarus, MD; Fernando J Martinez, MD, MS; Deborah A Meyers, PhD; Wendy C Moore, MD; John D Newell Jr, MD; Robert Paine, III, MD; Laura Paulin, MD, MHS; Stephen P Peters, MD, PhD; Cheryl Pirozzi, MD; Nirupama Putcha, MD, MHS; Elizabeth C Oelsner, MD, MPH; Wanda K O'Neal, PhD; Victor E Ortega, MD, PhD; Sanjeev Raman, MBBS, MD; Stephen I. Rennard, MD; Donald P Tashkin, MD; J Michael Wells, MD; Robert A Wise, MD; and Prescott G Woodruff, MD, MPH. The project officers from the Lung Division of the National Heart, Lung, and Blood Institute were Lisa Postow, PhD, and Lisa Viviano, BSN; SPIROMICS was supported by contracts from the NIH/NHLBI (HHSN268200900013C, HHSN268200900014C, HHSN268200900015C, HHSN268200900016C, HHSN268200900017C, HHSN268200900018C, HHSN268200900019C, HHSN268200900020C), grants from the NIH/NHLBI (U01 HL137880 and U24 HL141762), and supplemented by contributions made through the Foundation for the NIH and the COPD Foundation from AstraZeneca/MedImmune; Bayer; Bellerophon Therapeutics; Boehringer-Ingelheim Pharmaceuticals, Inc.; Chiesi Farmaceutici S.p.A.; Forest Research Institute, Inc.; GlaxoSmithKline; Grifols Therapeutics, Inc.; Ikaria, Inc.; Novartis Pharmaceuticals Corporation; Nycomed GmbH; ProterixBio; Regeneron Pharmaceuticals, Inc.; Sanofi; Sunovion; Takeda Pharmaceutical Company; and Theravance Biopharma and Mylan.

References

1. Oelsner, E.C., Balte, P.P., Cassano, P.A., Couper, D., Enright, P.L., Folsom, A.R., Hankinson, J., Jacobs, D.R., Kalhan, R., Kaplan, R., et al. (2018). Harmonization of Respiratory Data From 9 US Population-Based Cohorts: The NHLBI Pooled Cohorts Study. Am. J. Epidemiol. *187*, 2265–2278.

2. (1989). The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. Am. J. Epidemiol. *129*, 687–702.

3. Mirabelli, M.C., Preisser, J.S., Loehr, L.R., Agarwal, S.K., Barr, R.G., Couper, D.J., Hankinson, J.L., Hyun, N., Folsom, A.R., and London, S.J. (2016). Lung function decline over 25 years of follow-up among black and white adults in the ARIC study cohort. Respir. Med. *113*, 57–64.

4. Hughes, G.H., Cutter, G., Donahue, R., Friedman, G.D., Hulley, S., Hunkeler, E., Jacobs, D.R., Liu, K., Orden, S., and Pirie, P. (1987). Recruitment in the Coronary Artery Disease Risk Development in Young Adults (Cardia) Study. Control. Clin. Trials *8*, 68S-73S.

5. Friedman, G.D., Cutter, G.R., Donahue, R.P., Hughes, G.H., Hulley, S.B., Jacobs, D.R., Liu, K., and Savage, P.J. (1988). CARDIA: study design, recruitment, and some characteristics of the examined subjects. J. Clin. Epidemiol. *41*, 1105–1116.

6. (1979). ATS statement--Snowbird workshop on standardization of spirometry. Am. Rev. Respir. Dis. *119*, 831–838.

7. Fried, L.P., Borhani, N.O., Enright, P., Furberg, C.D., Gardin, J.M., Kronmal, R.A., Kuller, L.H., Manolio, T.A., Mittelmark, M.B., and Newman, A. (1991). The Cardiovascular Health Study: design and rationale. Ann. Epidemiol. *1*, 263–276.

8. Enright, P.L., Kronmal, R.A., Higgins, M., Schenker, M., and Haponik, E.F. (1993). Spirometry reference values for women and men 65 to 85 years of age. Cardiovascular health study. Am. Rev. Respir. Dis. *147*, 125–133.

9. Enright, P.L., Kronmal, R.A., Higgins, M.W., Schenker, M.B., and Haponik, E.F. (1994). Prevalence and correlates of respiratory symptoms and disease in the elderly. Cardiovascular Health Study. Chest *106*, 827–834.

10. Larkin, E.K., Patel, S.R., Goodloe, R.J., Li, Y., Zhu, X., Gray-McGuire, C., Adams, M.D., and Redline, S. (2010). A Candidate Gene Study of Obstructive Sleep Apnea in European Americans and African Americans. Am. J. Respir. Crit. Care Med. *182*, 947–953.

11. (2017). An Investigation of Coronary Heart Disease in Families: The Framingham Offspring Study. Am. J. Epidemiol. *185*, 1093–1102.

12. Lavange, L.M., Kalsbeek, W.D., Sorlie, P.D., Avilés-Santa, L.M., Kaplan, R.C., Barnhart, J., Liu, K., Giachello, A., Lee, D.J., Ryan, J., et al. (2010). Sample design and

cohort selection in the Hispanic Community Health Study/Study of Latinos. Ann. Epidemiol. *20*, 642–649.

13. Barr, R.G., Avilés-Santa, L., Davis, S.M., Aldrich, T.K., Gonzalez, F., Henderson, A.G., Kaplan, R.C., LaVange, L., Liu, K., Loredo, J.S., et al. (2016). Pulmonary Disease and Age at Immigration among Hispanics. Results from the Hispanic Community Health Study/Study of Latinos. Am. J. Respir. Crit. Care Med. *193*, 386–395.

14. Sorlie, P.D., Avilés-Santa, L.M., Wassertheil-Smoller, S., Kaplan, R.C., Daviglus, M.L., Giachello, A.L., Schneiderman, N., Raij, L., Talavera, G., Allison, M., et al. (2010). Design and implementation of the Hispanic Community Health Study/Study of Latinos. Ann. Epidemiol. *20*, 629–641.

15. Taylor, H.A. (2005). The Jackson Heart Study: an overview. Ethn. Dis. *15*, S6-1–3.

16. Taylor, H.A., Wilson, J.G., Jones, D.W., Sarpong, D.F., Srinivasan, A., Garrison, R.J., Nelson, C., and Wyatt, S.B. (2005). Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study. Ethn. Dis. *15*, S6-4–17.

17. Wilson, J.G., Rotimi, C.N., Ekunwe, L., Royal, C.D.M., Crump, M.E., Wyatt, S.B., Steffes, M.W., Adeyemo, A., Zhou, J., Taylor, H.A., et al. (2005). Study design for genetic analysis in the Jackson Heart Study. Ethn. Dis. *15*, S6-30–37.

18. Bild, D.E., Bluemke, D.A., Burke, G.L., Detrano, R., Diez Roux, A.V., Folsom, A.R., Greenland, P., Jacob, D.R., Kronmal, R., Liu, K., et al. (2002). Multi-Ethnic Study of Atherosclerosis: objectives and design. Am. J. Epidemiol. *156*, 871–881.

19. Hankinson, J.L., Kawut, S.M., Shahar, E., Smith, L.J., Stukovsky, K.H., and Barr, R.G. (2010). Performance of American Thoracic Society-recommended spirometry reference values in a multiethnic sample of adults: the multi-ethnic study of atherosclerosis (MESA) lung study. Chest *137*, 138–145.

20. Regan, E.A., Hokanson, J.E., Murphy, J.R., Make, B., Lynch, D.A., Beaty, T.H., Curran-Everett, D., Silverman, E.K., and Crapo, J.D. (2010). Genetic epidemiology of COPD (COPDGene) study design. COPD *7*, 32–43.

21. Couper, D., LaVange, L.M., Han, M., Barr, R.G., Bleecker, E., Hoffman, E.A., Kanner, R., Kleerup, E., Martinez, F.J., Woodruff, P.G., et al. (2014). Design of the Subpopulations and Intermediate Outcomes in COPD Study (SPIROMICS). Thorax *69*, 491–494.

22. Vogelmeier, C.F., Criner, G.J., Martinez, F.J., Anzueto, A., Barnes, P.J., Bourbeau, J., Celli, B.R., Chen, R., Decramer, M., Fabbri, L.M., et al. (2017). Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report. GOLD Executive Summary. Am. J. Respir. Crit. Care Med. *195*, 557–582.