UC San Diego
UC San Diego Previously Published Works

Title

Race-Specific Spirometry Equations Do Not Improve Models of Dyspnea and Quantitative Chest CT Phenotypes.

Permalink

https://escholarship.org/uc/item/37z7r65r

Journal

Chest, 164(6)

Authors

Diaz, Alejandro

Baldomero, Arianne

Kinney, Greg

et al.

Publication Date

2023-12-01

DOI

10.1016/j.chest.2023.07.019

Copyright Information

Peer reviewed

Check for updates

Race-Specific Spirometry Equations Do Not Improve Models of Dyspnea and Quantitative Chest CT Phenotypes

Amy L. Non, PhD; Barbara Bailey, PhD; Surya P. Bhatt, MD; Richard Casaburi, MD; Elizabeth A. Regan, MD, PhD; Angela Wang, MD; Alfonso Limon, PhD; Chantal Rabay, BS; Alejandro A. Diaz, MD; Arianne K. Baldomero, MD; Greg Kinney, PhD; Kendra A. Young, PhD; Ben Felts, MS; Carol Hand, MS; and Douglas J. Conrad, MD

BACKGROUND: Race-specific spirometry reference equations are used globally to interpret lung function for clinical, research, and occupational purposes, but inclusion of race is under scrutiny.

RESEARCH QUESTION: Does including self-identified race in spirometry reference equation formation improve the ability of predicted FEV_1 values to explain quantitative chest CT abnormalities, dyspnea, or Global Initiative for Chronic Obstructive Lung Disease (GOLD) classification?

STUDY DESIGN AND METHODS: Using data from healthy adults who have never smoked in both the National Health and Nutrition Survey (2007-2012) and COPDGene study cohorts, race-neutral, race-free, and race-specific prediction equations were generated for FEV_1. Using sensitivity/specificity, multivariable logistic regression, and random forest models, these equations were applied in a cross-sectional analysis to populations of individuals who currently smoke and individuals who formerly smoked to determine how they affected GOLD classification and the fit of models predicting quantitative chest CT phenotypes or dyspnea.

RESULTS: Race-specific equations showed no advantage relative to race-neutral or race-free equations in models of quantitative chest CT phenotypes or dyspnea. Race-neutral reference equations reclassified up to 19% of Black participants into more severe GOLD classes, while race-neutral/race-free equations may improve model fit for dyspnea symptoms relative to race-specific equations.

INTERPRETATION: Race-specific equations offered no advantage over race-neutral/race-free equations in three distinct explanatory models of dyspnea and chest CT scan abnormalities. Race-neutral/race-free reference equations may improve pulmonary disease diagnoses and treatment in populations highly vulnerable to lung disease.

CHEST 2023; 164(6):1492-1504

KEY WORDS: ethnicity; pulmonary function test; race; reference equations; spirometry

---

AFFILIATIONS: From the Department of Anthropology (A. L. N. and C. R.) and Department of Medicine (A. W. and D. J. C.), University of California San Diego, La Jolla, CA; Department of Mathematics and Statistics (B. B. and B. F.), San Diego State University, San Diego, CA; Division of Pulmonary, Allergy, and Critical Care Medicine (S. P. B.), University of Alabama at Birmingham, Birmingham, AL; Rehabilitation Clinical Trials Center (R. C.), Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA; Division of Rheumatology and Department of Medicine (E. A. R.), National

Take-home Points

Study Question: Does including self-identified race in the formation of spirometry reference equations improve the ability of predicted FEV$_1$ values to explain quantitative chest CT abnormalities, dyspnea, or Global Initiative for Chronic Obstructive Lung Disease classification?

Results: Race-neutral and race-free equations reclassified up to 19% of Black individuals who smoke to worse Global Initiative for Chronic Obstructive Lung Disease classes in the COPDGene smoking cohort, with the greatest effects seen in individuals with mild smoking-related disease. The generated percent predicted FEV$_1$ values from race-neutral and race-free spirometry equations showed no significant improvement in model fit of dyspnea or quantitative chest CT phenotypes (emphysema, air trapping, airway wall thickness).

Interpretation: Race-neutral/free reference equations may improve pulmonary disease diagnoses and treatment in populations highly vulnerable to lung disease relative to race-specific equations.

Interpretation of spirometry results has traditionally relied upon reference equations to provide an estimate of "normal" lung function for an individual's age, gender, height—and controversially—race/ethnicity. These equations are used for clinical, research, and occupational purposes to diagnose pulmonary disease, assess disease progression, and explain radiographic abnormalities, as well as determine disability and evaluate fitness for higher risk jobs, and thus have enormous clinical and financial importance. The inclusion of race in these equations is based on large cross-sectional, population-wide studies that consistently show lower measures of lung function for some racial/ethnic minority groups, specifically up to 10% to 15% lower FEV$_1$ for Black individuals.[1,2] However,

the clinical value of race adjustments has increasingly been questioned.[3-6] Although recent studies found no prognostic benefit of race-specific equations compared with "race-neutral" equations in mortality or respiratory events,[7-11] others continue to defend the use of race in prediction equations.[12,13] Race-specific equations are recommended by the most recent US and European guidelines[14] and are still used in clinical care and pulmonary research worldwide. However, applying race-specific equations may mask developmental or acquired lung damage among racial and ethnic minority groups[15-17] and risks underdiagnosing damaged lungs in marginalized groups at high risk of respiratory disease,[18-21] thereby exacerbating racial health inequalities.

We examined how the percent predicted FEV$_1$ (ppFEV$_1$) values calculated from race-specific, race-neutral, and race-free reference equations differentially affect pulmonary phenotypes in two large cohorts of individuals who smoke. First, using a selected sample of healthy adults who have never smoked from both National Health and Nutrition Examination Survey (NHANES) (2007-2012) and COPDGene cohorts (e-Fig 1), we generated new race-free equations that entirely exclude race from model formation and race-specific prediction equations for FEV$_1$ and FVC. Second, these equations were compared with the Global Lung Initiative (GLI) race-specific equations,[2] the race-specific equations of Hankinson et al,[1] and the race-neutral GLI-Other (uses a universal race-correction) and the GLI-Global (weights racial groups in the reference population) equations.[2] Third, we applied these spirometry prediction equations and determined how they differentially: (1) affect the Global Initiative for Chronic Obstructive Lung Disease (GOLD) severity classification in both the NHANES and COPDGene smoking cohorts; and (2) model quantitative chest CT scan phenotypes and dyspnea in the COPDGene study participants. Our intent was to compare how the different reference equations model clinically important pulmonary phenotypes.

Jewish Health, Denver, CO; Oneirix Labs (A. L.), Carlsbad, CA; Division of Pulmonary and Critical Care Medicine (A. A. D.), Brigham and Women's Hospital, Harvard Medical School, Boston, MA; Pulmonary, Allergy, Critical Care and Sleep Medicine Section (A. K. B.), Minneapolis VA Health Care System, Minneapolis, MN; Department of Epidemiology (G. K. and K. A. Y.), Colorado School of Public Health, University of Colorado Anschutz Medical Campus, Aurora, CO; and Advanced Mathematical Computing (C. H.), San Diego, CA.

CORRESPONDENCE TO: Douglas J. Conrad, MD; email: dconrad@health.ucsd.edu

Study Design and Methods

Details on formation and characterization of asymptomatic nonsmoking (nh3700 and cg419 cohorts) and smoking (nh785 and COPDGene Phase I) cohorts are presented in e-Appendix 1, e-Figures 1 to 3, and Table 1.

Race-Specific and Race-Neutral Prediction Models

The predicted FEV$_1$ and lower limit of normal (LLN) values for those that never smoked from the GLI equations were obtained using the GLI website (https://gli-calculator.ersnet.org/index.html, version 2.0, April 2023). Predicted and LLN FEV$_1$ values by Hankinson et al[1] (NHANES III)

TABLE 1] Characteristics of the NHANES (N = 3,700) and COPDGene (N = 419) Healthy Cohorts Who Have Never Smoked by Race/Ethnicity

Characteristic	NHANES, All (n= 3,700)	NHANES, White (n = 1,420 [38.4%])	NHANES, Black (n = 762 [20.6%])	NHANES, Mexican American (n = 682 [18.4%])	NHANES, Other Hispanic (n = 473 [12.8%])	NHANES, Other/Mixed (n = 363 [9.8%])	COPDGene, All (n = 419)	COPDGene, White (n = 342 [81.6%])	COPDGene, Black (n = 77 [18.4%])
Age, y	51.0 (19.0)	52.0 (21.0)	52.0 (18.0)	49.0 (11.0)[a]	51.0 (20.0)	49.0 (17.0)[a]	59.1 (15.4)	61.4 (15.4)	55.1 (9.6)[a]
Female, n (%)	2,160 (58.4)	803 (56.5)	446 (58.5)	396 (58.1)	302 (63.8)	213 (58.7)	239 (57.0)	194 (56.7)	45 (58.4)
Height, cm	165.1 (14.4)	168.4 (15.1)	167.7 (13.4)	160.9 (12.7)[a]	160.3 (14.4)[a]	162.1 (12.9)[a]	168.6 (14.0)	168.7 (14.7)	168.0 (12.7)
Weight, kg	78.8 (25.1)	81.8 (26.1)	86.5 (27.0)[a]	75.4 (18.9)[a]	75.1 (22.7)[a]	66.1 (19.4)[a]	78.0 (23.1)	76.5 (23.9)	80.0 (19.2)
BMI, kg/m^2	28.7 (7.4)	28.6 (7.6)	30.4 (8.2)[a]	29.2 (6.0)	28.6 (6.5)	24.9 (5.5)[a]	27.1 (6.0)	26.8 (5.9)	28.3 (5.11)
FEV$_1$, L	2.83 (1.11)	3.07 (1.19)	2.57 (0.98)[a]	2.89 (1.01)[a]	2.70 (1.03)[a]	2.65 (0.94)[a]	2.83 (1.11)	2.92 (1.19)	2.62 (0.85)[b]
FVC, L	3.55 (1.41)	3.90 (1.51)	3.16 (1.16)[a]	3.58 (1.32)[a]	3.39 (1.25)[a]	3.30 (1.22)[a]	3.55 (1.38)	3.60 (1.45)	3.16 (1.20)[a]
FEV$_1$/FVC	0.80 (0.07)	0.79 (0.07)	0.81 (0.07)[a]	0.81 (0.06)[a]	0.80 (0.06)[a]	0.80 (0.07)[a]	0.80 (0.08)	0.80 (0.07)	0.82 (0.07)[a]
ppFEV$_1$ GLI[d]	99.0 (17.1)	99.8 (16.3)	98.0 (18.7)[c]	100.4 (16.2)	97.2 (16.4)[c]	96.8 (16.7)[c]	101.0 (18.7)	101.6 (18.6)	99.9 (16.7)
ppFEV$_1$ GLI-Other[e]	102.61 (20.0)	107.1 (17.5)	90.5 (17.4)[a]	107.7 (17.4)	104.4 (17.6)[c]	96.8 (16.7)[a]	106.6 (20.3)	109.0 (20.0)	92.3 (16.2)[a]
qCT parameters									
Pi10							1.77 (0.42)	1.75 (0.41)	1.94 (0.44)[c]
Air trapping (n = 321)	…	…	…	…	…	…	7.18 (8.90)	7.20 (8.70)	6.80 (10.01)
Percent emphysema	…	…	…	…	…	…	0.84 (1.94)	0.88 (2.16)	0.62 (1.19)[c]
mMRC score (0-4)	…	…	…	…	…	…	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)

Values of continuous variables are presented as medians (interquartile range). GLI = Global Lung Initiative; mMRC = modified Medical Research Council; NHANES = National Health and Nutrition Examination Survey; Pi10 = airway wall thickness estimate based on square root of wall area of a 10 mm lumen perimeter[22]; ppFEV$_1$ = percent predicted FEV$_1$; qCT = quantitative chest CT.

Indicates significant difference relative to White group at: [a]$P < .0001$, [b]$P < .001$, [c]$P < .05$, according to the Kruskal-Wallis test for comparison of continuous variables between (non-Hispanic) Black and White racial/ethnic groups in COPDGene, and analysis of variance with Tukey ad hoc comparisons for continuous variables between each racial/ethnic group relative to White participants in the NHANES data.

[d]Guideline-based application of GLI race/ethnic-specific reference equations. The GLI equations for White/European individuals were used to estimate ppFEV$_1$ for the NHANES Mexican American and Other Hispanic groups, following other studies (8). The GLI-predicted FEV$_1$ values for the NHANES group "Other/Mixed Race" used the GLI-Other equations.

[e]The GLI-Other equation was used to generate race-neutral estimates of ppFEV$_1$ for all racial/ethnic groups.

were calculated using published equations. For both never-smoking data sets (the NHANES data set of 3,700 healthy individuals [nh3700] and the COPDGene data set of 419 healthy individuals), multivariable linear quantile regression was used to generate predicted (median quantile) and LLN (fifth quantile) models and associated R1 values (Table 2). The R1 value is a measure of explained variability of the data in quantile regression and is used to compare models.[23] Predictors in the race-specific equations included age (years), height (centimeters), gender (male/female), and self-identified race/ethnicity. Predictors in race-free equations included only age, height, and gender (e-Tables 1-3). A similar approach was used to generate models for predicted and LLN values for log (FVC). This approach generated four race-specific and four race-neutral/race-free models for the predicted log (FEV_1) and log (FVC) from different source populations (e-Tables 1-3, Tables 2, 3). Identity and probability density plots of the differences between the predicted race-specific, race-neutral, and race-free models were used to explore the effect of race in the models.

GOLD Classification Changes

Each individual in both smoking cohorts was assigned a GOLD spirometry class (GOLD 1-4), the preserved ratio impaired spirometry class,[25] or GOLD 0 (ie, FEV_1/FVC ratio > 0.7 and $ppFEV_1 \geq 80\%$) using the different race-specific and race-neutral equations. The percentage of individuals who changed GOLD class from the GLI standard (race-specific) was calculated in the total data set and within each racial group for both NHANES

participants who formerly smoked and COPDGene participants who currently smoke.

Modeling Pulmonary Phenotypes

Sensitivity/Specificity Modeling: Measured FEV_1 values were classified as above or below the LLN to assess the sensitivity and specificity of each $ppFEV_1$ reference equation to model abnormal chest CT scan phenotypes in COPDGene phase I participants. Chest CT scan phenotypes were defined as abnormal if: (1) the percent emphysema was > 5%; (2) the percent air trapping was > 15%; or (3) the airway wall thickness estimate based on square root of wall area of a 10 mm lumen perimeter was > 2.5.[26] The sensitivity, specificity, negative predictive value, positive predictive value, and the area under the curve (AUC) were calculated for each model in the overall population and within each race. A parallel approach assessed the ability of the LLN of each model to predict a modified Medical Research Council dyspnea score (mMRC) > 1.

Logistic Regression Models of Abnormal Chest CT Scan Phenotypes and Dyspnea: Because complex demographic factors (smoking status and history, gender, FEV_1/FVC ratio, and scanner type) and the $ppFEV_1$ influence quantitative chest CT scan metrics,[27,28] multivariable logistic regression models were generated of abnormal chest CT scan phenotypes using these covariates and the $ppFEV_1$ values derived from each of the race-neutral and race-specific equations (e-Tables 4A, 4B). Models were compared by using Akaike and Bayesian information criteria. A parallel approach was used to model dyspnea and included covariates of FEV_1/FVC ratio, pack-year

TABLE 2] Models of Predicted log (FEV_1) Median and Fifth Percentile Quantile Regression Coefficients for NHANES (N = 3,700) and COPDGene (N = 419)

Model	β0 (Intercept)	β1 (Age in Years)	β2 (Female Gender)	β3 (Height in Centimeters)	β4 (Race)	R1
NHANES: log (FEV_1)						
Median predicted: AGH	−0.30931	-0.00966	−0.16027	0.01177	...	0.485
Median predicted: AGHR (White)	−0.42116	-0.00922	−0.14935	0.01253	Reference	0.547
AGHR (African American)					−0.17048	
AGHR (Mexican American)					−0.00454	
AGHR (Hispanic Other)					−0.0219	
AGHR (other/mixed)					−0.09983	
5th percentile: AGH	−0.37667	−0.01069	−0.16976	0.01091	...	...
5th percentile: AGHR (White)	−0.55486	−0.01094	−0.15568	0.01258	Reference	...
African American					−0.19925	
Mexican American					−0.00657	
Hispanic Other					−0.03935	
Other/Mixed					−0.09986	
COPDGene: log (FEV_1)						
Median predicted: AGH	−0.40813	−0.00958	−0.17036	0.01262	...	0.476
Median predicted: AGHR (Black)	−0.37885	−0.01077	−0.15455	0.01294	−0.16483	0.531
5th percentile: AGH	−0.04169	−0.01084	−0.17910	0.00946	...	...
5th percentile: AGHR (Black)	−0.14623	−0.01251	−0.15592	0.01097	−0.17822	...

The equation used to model the predicted log (FEV_1) is as follows: predicted (or lower limit of normal) FEV_1 (L)= $e^{(β0 + Age * β1 + Gender code term * β2 + Height * β3 + Race code term * β4)}$. In both cohorts, race is coded as Black (1) relative to NHW (0) as the reference group. In data from the NHANES data set of 3,700 healthy individuals, race is modeled with NHW (0) as the reference group, and other racial/ethnic groups were coded as 1 if present and multiplied by the corresponding race coefficient (β4). In both cohorts, gender code term for male subjects is 0 and 1 for female subjects. The log of FEV_1 (post-bronchodilator values were used when available for NHANES data and always for COPDGene data) in liters was used because it optimized the explained variability compared with modeling raw FEV_1 values. The racial terminology used is consistent with that of the published cohorts. AGH = models including age, gender, and height only; AGHR = models including age, gender, height, and race/ethnicity; NHANES = National Health and Nutrition Examination Survey.

TABLE 3] Summaries of Source Data, Covariates, and Nomenclature for Predicted FEV_1 Models

Predicted FEV_1 Model	Type	Race	Source Population	Source Sample Size by Race/Ethnicity	Source Age Range	Covariates	Reference
GLI race-specific	QR-Linear	Specific	International	Total: n = 74,187 White: n = 57,395 Black: n = 3,545 NE Asian: n = 4,992 SE Asian: n = 8,225	3-95 y	AGHR	3
Hankinson	OLS-Linear	Specific	NHANES-1999	Total n = 7,429 White: n = 2,281 Black: n = 2,508 Mexican American: n = 2,639	8-80 y	AGHR	2
cg419_AGHR	QR-Linear	Specific	COPDGene never-smoking cohort	Total: n = 419 White: n = 342 Black: n = 77	45-82 y	AGHR	Table 2
nh3700_AGHR	QR-Linear	Specific	NHANES 2007-2012 healthy individuals who formerly smoked	Total: n = 3,700 White: n = 1,420 Black: n = 762 Mexican American: n = 682 Other Hispanic: n = 473 Other: n = 363	35-79 y	AGHR	Table 2
GLI-Other	GAMLSS	Neutral	International	n = 74,187	3-95 y	AGH[a]	3
GLI-Global	GAMLSS	Neutral	International	n = 74,185	3-95 y	AGH[b]	24
cg419_AGH	QR-Linear	Free	COPDGene healthy never-smoking	n = 419	45-82 y	AGH	Table 2
nh3700_AGH	QR-Linear	Free	NHANES 2007-2012 healthy never-smoking cohort	n = 3,700	35-79 y	AGH	Table 2

Details about models and source populations used to develop each of the race-specific, race-neutral, and race-free models used in this study. GLI-Other and GLI-Global equations are race-neutral, but are not race-free, as they averaged race/ethnicity estimates across four major racial/ethnic groups. AGH = age, gender, and height; AGHR = age, gender, height, and race; cg419 = COPDGene data set of 419 healthy individuals; GAMLSS = General Additive Models for Location Scale and Shape; GLI = Global Lung Initiative; NE = northeast; nh3700 = National Health and Nutrition Examination Survey data set of 3,700 healthy individuals; NHANES = National Health and Nutrition Examination Survey; OLS = ordinary least squares regression; QR = quantitative regression; SE = southeast.
[a]GLI-Other was calculated by taking "its mean and CoV adjustments the corresponding adjustments for the four main ethnic groups, averaged over group and sex".[3]
[b]For GLI-Global, an inverse probability weight was applied for each of the four racial groups included in the data set.

smoking history, age, weight, height, 6-min walking distance, total lung capacity from the CT scan,[27] and the pre-/post-bronchodilator difference of FEV_1 and $ppFEV_1$ values calculated from the different race-neutral and race-specific equations (e-Tables 5A, 5B).

Random Forest Models: The random forest algorithm was used to compare models of the abnormal chest CT phenotypes and dyspnea using the same covariates as the logistic regression models

(e-Tables 4C, 4D, e-Tables 5C, 5D). The classification error rates were compared to assess model performance using the different race-specific and race-neutral $ppFEV_1$ values.

Software: All analyses used features in the base R (version 4.0.5) program (R Foundation for Statistical Computing). Additional packages included rspiro (v.2), quantreg (v5.55), pROC (v1.18.0), randomForest (v4.6-14), and rfPermute (v2.1.81).

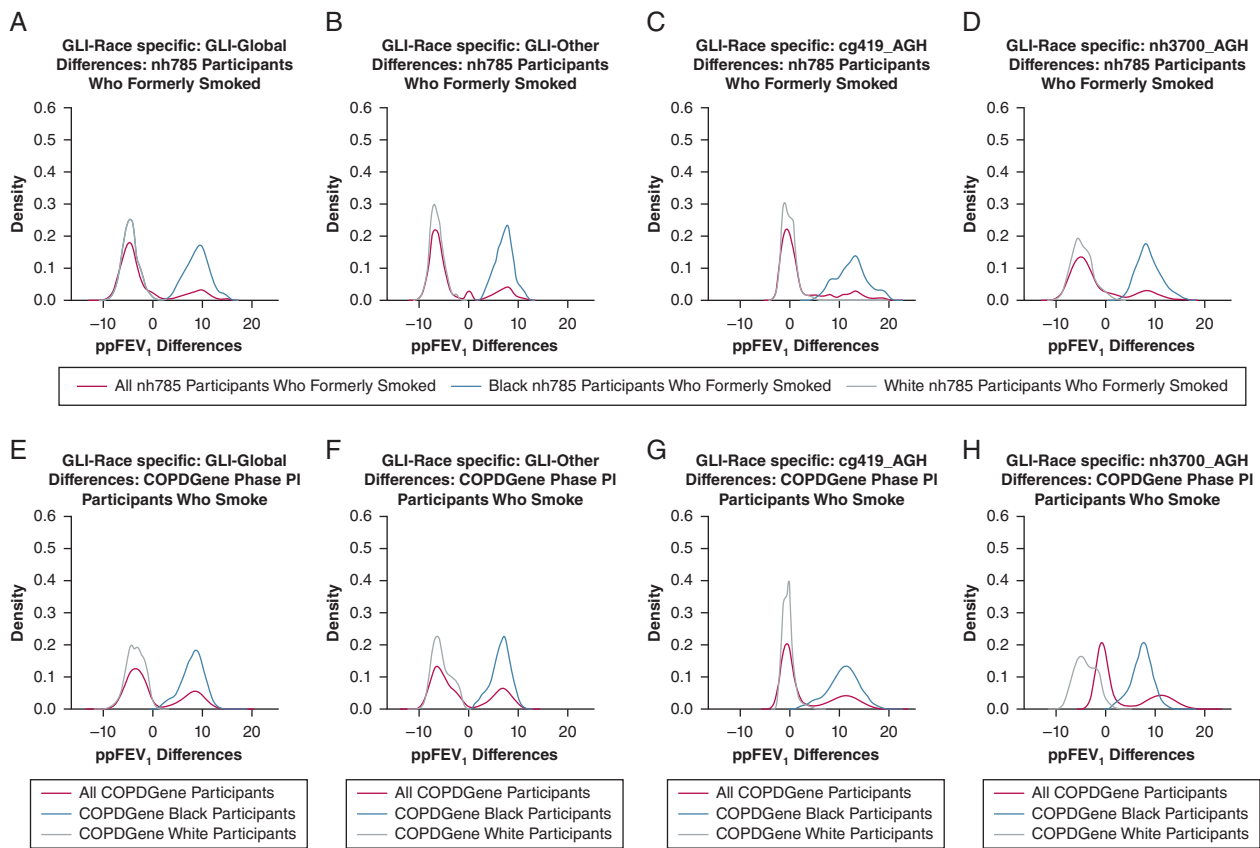## Results

The nh3700 never-smoking, healthy cohort consisted of 38% White, 21% Black, 18% Mexican-American, 13% other Hispanic, and 10% mixed racial or "other" race individuals (Table 1). Relative to White participants, Black participants had similar height but a higher median weight. All other groups had lower median weight and height than White participants. All racial/ethnic groups had a lower median FEV_1 and FVC, and a higher median FEV_1/FVC ratio, compared with White participants. White and Mexican-American individuals had higher GLI race-specific $ppFEV_1$ values than the other ethnic groups.

The COPDGene never-smoking, healthy cohort consisted of 18% Black and 82% White participants. Relative to White participants, Black participants were younger, had lower FEV_1 and FVC values, and had higher FEV_1/FVC ratios, consistent with other studies.[1,8] Black participants had significantly higher airway wall thickness estimate based on the square root of wall area of a 10 mm lumen perimeter and lower percent emphysema relative to White participants (e-Fig 2, Table 1). Compared with the nh3700 never-smoking cohort, the COPDGene healthy participants who had never smoked were older and less racially diverse, but they were otherwise similar in demographic, anthropometric, and spirometry assessments (Table 1). Compared with the COPDGene smoking cohort, the nh785 cohort of individuals who formerly smoked was younger, with higher BMI and higher FEV_1 and FVC values, suggesting less severe smoking-related disease (e-Fig 3).

Comparison of Models Among Those Who Never Smoked

There was a high correlation between predicted values generated using all race-specific equations for FEV_1 and for LLN (e-Fig 4). Probability density plots of the differences between the FEV_1 values generated using the different race-specific equations show minor differences in the predicted FEV_1 values between Black and White populations (e-Fig 5). These high correlations show the validity of the generated healthy data sets and the modeling approach. In contrast, race-neutral equations

generated higher predicted FEV_1 and LLN values than race-specific equations for the Black participants but unchanged or minimally shifted to lower values in the White participants of both healthy, never-smoking cohorts (e-Figs 6, 7). Race-free models generated even higher predicted FEV_1 and LLN values than those of GLI-Global (e-Fig 8). There were no differences between any racial groups in the density plots of predicted FEV_1 when using race-neutral equations. Including race in the FEV_1 prediction equations improved model fit as measured by slightly higher R1 values in both never-smoking, healthy populations, but it can also obscure measured differences in FEV_1 between White and Black individuals in both never-smoking, healthy populations (e-Fig 9, e-Table 1, Table 2). To a lesser extent, this effect was also seen between White and other/mixed race populations in the nh3700 population (e-Fig 10). Using alternative anthropometric measurements instead of height did not improve model fit or mitigate racial differences (e-Fig 11).

Race-Specific vs Race-Neutral Equations in Individuals Who Smoke:
The $ppFEV_1$ value generated from the new race-specific equations were within 3% of the estimates generated using GLI race-specific equations for the total COPDGene smoking population (e-Fig 12). In contrast, the differences between GLI race-specific and race-neutral/race-free $ppFEV_1$ values in both smoking cohorts created bimodal curves with the Black population shifted positively, corresponding to a lower $ppFEV_1$ by an average of approximately 7% to 11% (Figs 1A-1H); the curves in the White population shifted negatively, which would result in slightly higher $ppFEV_1$ values. Similar to values seen in the healthy never-smoking cohort, the race-free equations generated lower $ppFEV_1$ values than race-neutral equations in COPDGene participants who smoked (e-Fig 13).

Model Effects on GOLD Classification

To identify the differential effects of the $ppFEV_1$ equations on GOLD classification, we identified the number of individuals who were reclassified using the newly developed race-specific equations compared with the GLI race-specific equations. In the two smoking

A — GLI-Race specific: GLI-Global Differences: nh785 Participants Who Formerly Smoked

B — GLI-Race specific: GLI-Other Differences: nh785 Participants Who Formerly Smoked

C — GLI-Race specific: cg419_AGH Differences: nh785 Participants Who Formerly Smoked

D — GLI-Race specific: nh3700_AGH Differences: nh785 Participants Who Formerly Smoked

— All nh785 Participants Who Formerly Smoked   — Black nh785 Participants Who Formerly Smoked   — White nh785 Participants Who Formerly Smoked

E — GLI-Race specific: GLI-Global Differences: COPDGene Phase PI Participants Who Smoke

F — GLI-Race specific: GLI-Other Differences: COPDGene Phase PI Participants Who Smoke

G — GLI-Race specific: cg419_AGH Differences: COPDGene Phase PI Participants Who Smoke

H — GLI-Race specific: nh3700_AGH Differences: COPDGene Phase PI Participants Who Smoke

— All COPDGene Participants   — COPDGene Black Participants   — COPDGene White Participants

Average % GOLD reclassification compared with GLI race-specific reference equations

I

	Race-specific	Race-neutral/free
All nh785 Participants Who Smoke		
Worse GOLD Class	2.1	3.4
Better GOLD Class	0.9	3.9
White nh785 Participants Who Smoke		
Worse GOLD Class	2.4	0.2
Better GOLD Class	0.7	4.7
Black nh785 Participants Who Smoke		
Worse GOLD Class	1.4	_17.3_
Better GOLD Class	1.9	1.6

J

	Race-specific	Race-neutral/free
All COPDGene PI		
Worse GOLD Class	1.5	6.3
Better GOLD Class	0.9	4.9
White COPDGene PI		
Worse GOLD Class	1.9	0.2
Better GOLD Class	0.9	_7.3_
Black COPDGene PI		
Worse GOLD Class	0.6	_19.0_
Better GOLD Class	2.7	0

K

Black nh785 Participants Who Smoke GOLD % Reclassification	Hankinson	cg419_AGHR	nh3700_AGHR	GLI-Other	GLI-Global	cg419_AGH	nh3700_AGH
0 > PRISm	1.4	0.0	1.4	_6.3_	_9.9_	_16.2_	7.7
1 > 2	0.7	0.0	0.0	4.9	5.6	7.0	4.9
2 > 3	0.7	0.0	0.0	1.4	1.4	2.1	1.4

L

Black COPDGene PI GOLD % Reclassification	Hankinson	cg419_AGHR	nh3700_AGHR	GLI-Other	GLI-Global	cg419_AGH	nh3700_AGH
0 > PRISm	0.3	0.6	0.0	_9.9_	_12.8_	_19.1_	11.3
1 > 2	0.3	0.0	0.1	2.2	2.6	3.2	2.3
2 > 3	0.1	0.0	0.0	1.4	1.9	3.1	1.8
3 > 4	0.1	0.1	0.1	0.8	1.1	1.5	1.0

Figure 1 – A-L, Effects of race-specific vs race-neutral/free equations on predicted FEV_1 values and GOLD reclassification. Probability density of the differences in $ppFEV_1$ values in NHANES and COPDGene participants who smoked between GLI race-specific and GLI-Global (A, E), GLI Race-specific and GLI-Other (B, F); GLI race-specific and cg419_AGH (C, G), and GLI race-specific and nh3700_AGH (D, H). Differences in $ppFEV_1$ were calculated by subtracting each $ppFEV_1$ estimate derived from race-neutral or race-free equations from the estimate derived from the GLI standard (race-specific)

cohorts, the GOLD reclassification rate on average changed 1% to 3% when the calculated $ppFEV_1$ values using new race-specific reference equations were subtracted from those calculated by using GLI race-specific reference equations (Figs 1I, 1J). When stratified according to race, across all four race-neutral and race-free equations, an average of 17.3% of Black participants in the nh785 smoking cohort were reclassified to worse GOLD classes. In the more severely diseased COPDGene cohort (e-Fig 3), 19.0% of the Black participants were reclassified to worse GOLD classes using race-neutral/race-free equations. Most of these GOLD reclassifications were transitions from GOLD 0 (ie, $ppFEV_1 > 80$ and FEV_1/FVC ratio > 0.7) to the preserved ratio impaired spirometry class (Figs 1K, 1L).

Quantitative Chest CT Scan and Dyspnea

Three distinct models were used to compare the utility of the $ppFEV_1$ and respective LLN values from race-specific and race-neutral equations to model abnormal chest CT scan phenotypes and increased dyspnea (ie, mMRC score > 1). The sensitivity and specificity of each $ppFEV_1$ reference equation to model any abnormal chest CT scan findings were within 9%. Furthermore, the equations generated overlapping receiver-operating characteristic curves and thus similar AUC values. This univariate analysis showed no clinically relevant advantage of the race-specific equations over race-neutral equations to model chest CT scan abnormalities (e-Fig 14, Figs 2A, 2B).

The multivariable logistic regression models of abnormal chest CT scan phenotypes showed similar Akaike and Bayesian information criteria values using the race-specific and race-neutral equations in the total COPDGene phase I cohort as well as in the Black and White cohorts individually (e-Tables 4A, 4B, Fig 2C).

Finally, a supervised random forest approach was used to model the presence of any abnormal chest CT phenotype in the total COPDGene phase I population. Classification error rates differed minimally across equations in the total, Black, and White populations,
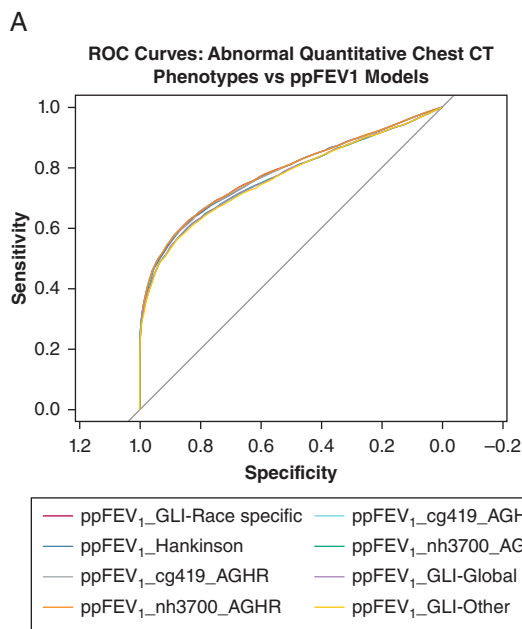
and they also differed minimally between each of the quantitative chest CT scan phenotypes. There were no instances in which the race-specific equations offered any significant decrease in classification errors over the race-neutral or race-free equations (e-Tables 4C, 4D, Fig 2D).

Findings from models assessing dyspnea revealed trends similar to the models of quantitative chest CT phenotypes. Specifically, the sensitivity/specificity, receiver-operating characteristic, and AUC values were nearly identical using race-specific and race-neutral/race-free equations and their respective LLNs (e-Fig 15, Figs 3A, 3B). In univariate analysis, the sensitivity and AUC values of the $ppFEV_1$ models were higher in White relative to Black COPDGene participants, regardless of whether race-specific or race-neutral/race-free $ppFEV_1$ models were used (e-Fig 15). The Akaike information criterion values from the multivariable logistic regression models of dyspnea were lower using the race-neutral or race-free vs race-specific $ppFEV_1$ values in the total COPDGene phase I population; this finding suggests improved model fit, but the clinical significance of this difference is unclear (e-Tables 5A, 5B, Fig 3C). Finally, classification error rates in the supervised random forest models were within 1% of each other using race-neutral/race-free vs race-specific equations, indicating no advantage in predicting dyspnea using race-specific equations in the total COPDGene phase I population or in analyses stratified according to race (e-Tables 5C, 5D, Fig 3D). The similarity in model fit when using GLI-Global vs race-free equations is unsurprising considering the very similar predicted FEV_1 distributions in identity and density plots of healthy populations (e-Figs 8, 16).

Discussion

This study investigates the role of race-specific, race-neutral, and race-free $ppFEV_1$ reference equations in evaluating disease severity and pulmonary phenotypes. Using three different modeling approaches, we show that

equation. Red lines = all COPDGene PI participants; blue lines = Black phase I participants; gray lines = White phase I participants. Average GOLD reclassification rates of the three race-specific (Hankinson, cg419AGHR, and nh3700AGHR) and three race-neutral/race-free (GLI-Other, cg419_AGH, and nh3700_AGH) prediction equations were subtracted from the standard GLI equation in the combined Black/White cohort, Black, and White participants in the nh785 smoker cohort (I) and the COPDGene PI cohort (J), reported as the average percent reclassified. GOLD 0 class is defined as FEV_1/FVC ratio > 0.7 and $ppFEV_1 > 80\%$. PRISm class is defined as FEV_1/FVC ratio < 0.7 and $ppFEV_1 < 80\%$. GOLD reclassification rates (percent reclassified of total Black population) are shown in Black nh785 participants who formerly smoked (K) and Black COPDGene participants who smoke (L) from the GLI race-specific equations by GOLD class transitions (rows) and models (columns). AGH = age, gender, and height; AGHR = age, gender, height, and race; cg419 = COPDGene dataset of 419 healthy individuals; COPDGene PI = COPDGene Study Phase I; GLI = Global Lung Initiative; GOLD = Global Initiative for Chronic Obstructive Lung Disease; nh3700 = National Health and Nutrition Examination Survey data set of 3,700 healthy individuals; NHANES = National Health and Nutrition Examination Survey; $ppFEV_1$ = percent predicted FEV_1; PRISm = preserved ratio impaired spirometry.

A

ROC Curves: Abnormal Quantitative Chest CT Phenotypes vs ppFEV1 Models

Legend:
- ppFEV₁_GLI-Race specific
- ppFEV₁_Hankinson
- ppFEV₁_cg419_AGHR
- ppFEV₁_nh3700_AGHR
- ppFEV₁_cg419_AGH
- ppFEV₁_nh3700_AGH
- ppFEV₁_GLI-Global
- ppFEV₁_GLI-Other

B

Univariate Sensitivity Analysis

ppFEV₁ Model	Sensitivity	Specificity	AUC	AUC (CI)
ppFEV₁ nh3700_AGHR	0.585	0.857	0.784	0.7753-0.7933
ppFEV₁ GLI-Race specific	0.567	0.879	0.782	0.7734-0.7915
ppFEV₁ cg419_AGHR	0.599	0.816	0.782	0.7730-0.7910
ppFEV₁ Hankinson	0.572	0.865	0.781	0.7720-0.7901
ppFEV₁ cg419_AGH	0.585	0.799	0.769	0.7601-0.7786
ppFEV₁ nh3700_AGH	0.567	0.860	0.768	0.7583-0.7769
ppFEV₁ GLI-Global	0.556	0.872	0.767	0.7576-0.7762
ppFEV₁ GLI-Other	0.556	0.862	0.766	0.7571-0.7757

C

Multivariable Logistic Regression AIC Values

ppFEV₁ Model	All	White	Black
ppFEV₁ cg419_AGH	8,886	5,594	3,260
ppFEV₁ nh3700_AGH	8,887	5,595	3,259
ppFEV₁ GLI-Global	8,890	5,595	3,263
ppFEV₁ GLI-Other	8,891	5,598	3,259
ppFEV₁ GLI-Race specific	8,892	5,598	3,260
ppFEV₁ nh3700_AGHR	8,893	5,594	3,260
ppFEV₁ cg419_AGHR	8,893	5,596	3,260
ppFEV₁ Hankinson	8,908	5,600	3,265

D

Random Forest Classification Errors

ppFEV₁ Model	All	White	Black
ppFEV₁ nh3700_AGH	24.1 (23.2-25.0)	23.1 (22.0-24.1)	28.3 (26.7-29.9)
ppFEV₁ cg419_AGH	24.3 (23.4-25.2)	23.0 (22.0-24.1)	28.4 (26.8-30.1)
ppFEV₁ cg419_AGHR	24.3 (23.4-25.2)	23.3 (22.3-24.4)	28.0 (26.4-29.6)
ppFEV₁ GLI-Global	24.3 (23.5-25.2)	22.9 (21.8-23.9)	28.0 (26.4-29.6)
ppFEV₁ GLI-Race specific	24.4 (23.5-25.3)	23.0 (22.0-24.1)	27.8 (26.3-29.5)
ppFEV₁ GLI-Other	24.4 (23.6-25.3)	23.0 (22.0-24.1)	27.9 (26.3-29.5)
ppFEV₁ nh3700_AGHR	24.6 (23.8-25.5)	22.8 (21.8-23.8)	28.0 (26.5-29.7)
ppFEV₁ Hankinson	25.1 (24.2-25.9)	23.4 (22.4-24.5)	28.2 (26.7-29.9)

Figure 2 – A, ROC curves of any abnormal quantitative chest CT phenotypes: (1) percent emphysema > 5%; (2) percent air trapping > 15%; and (3) airway wall thickness if the airway wall thickness estimate based on square root of wall area of a 10 mm lumen perimeter. B, Sensitivity, specificity, AUC and AUC confidence intervals (CI) of the ROC curve of the ppFEV₁ to predict any abnormal chest CT scan phenotype in the COPDGene phase I cohort. The sensitivity and specificity analyses used the lower limit of normal or the fifth percentile for each of the models. C, AIC from the multivariable logistic regression models of any abnormal quantitative CT phenotype with the following covariates: FEV₁/FVC ratio, smoking history (pack-years), scanner maker, smoking status, gender, and ppFEV₁. The AIC value generated from the models using the different race-specific and race-neutral equations for the ppFEV₁ are listed for the total cohort and also for the Black and White participants individually. D, Supervised random forest models of the abnormal chest CT scan phenotypes were generated using the same covariates as the logistic regression models. The classification error rates of the models using the different race-specific and race-neutral equations for the ppFEV₁ are listed for the total cohort and also for the Black and White individuals individually. The randomForest (v4.6-14) and rfPermute (v2.1.81) packages were used to obtain the classification error rates, mean decrease in accuracy, and P values. The default settings were used with ntree and nrep set to 500. AGH = age, gender, and height; AGHR = age, gender, height, and race; AUC = area under the curve; AIC = Akaike information criterion; cg419 = COPDGene data set of 419 healthy individuals; GLI = Global Lung Initiative; nh3700 = National Health and Nutrition Examination Survey data set of 3,700 healthy individuals; ppFEV₁ = percent predicted FEV₁; ROC = receiver-operating characteristic.

race-specific equations offer no advantage relative to either race-neutral or race-free equations in modeling quantitative chest CT scan phenotypes or dyspnea in two independent smoking cohorts. Specifically, we found that, compared with race-specific equations, race-neutral/race-free equations reclassified up to 19.0% of Black participants into more severe GOLD classes, and they may improve models of dyspnea. Use of race-neutral/race-free equations may result in additional pulmonary

disease diagnoses, as well as more aggressive treatment in populations highly vulnerable to lung disease.

Our findings are in line with other recent studies that found no prognostic benefit of race-specific over race-neutral spirometry reference equations.[7-11] Specifically, two studies found that the use of a race-neutral equation better predicted survival than race-specific equations using the NHANES III and 2007 to 2012

A

ROC Curves: mMRC > 1 vs ppFEV$_1$ Models

Legend:
- ppFEV$_1$ GLI-Race specific
- ppFEV$_1$ Hankinson
- ppFEV$_1$ cg419_AGHR
- ppFEV$_1$ nh3700_AGHR
- ppFEV$_1$ cg419_AGH
- ppFEV$_1$ nh3700_AGH
- ppFEV$_1$ GLI-Global
- ppFEV$_1$ GLI-Other

B
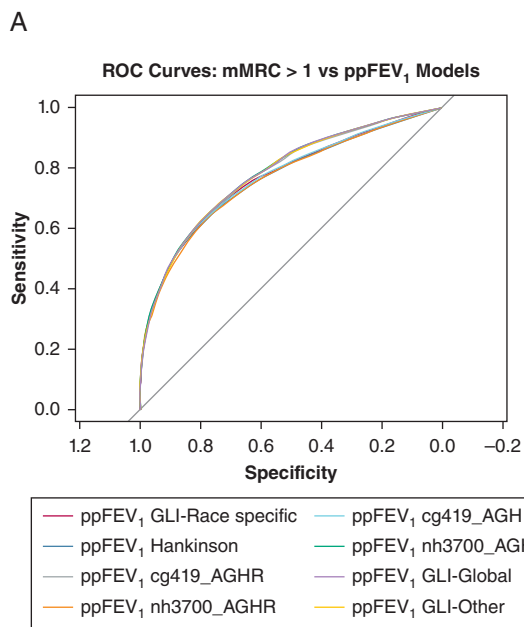
Univariate Sensitivity Analysis

ppFEV$_1$ Variable	Sensitivity	Specificity	AUC	AUC (CI)
ppFEV$_1$ nh3700_AGHR	0.767	0.657	0.783	0.7732-0.7922
ppFEV$_1$ GLI-Other	0.763	0.669	0.782	0.7726-0.7916
ppFEV$_1$ GLI-Global	0.556	0.872	0.782	0.7726-0.7916
ppFEV$_1$ cg419_AGH	0.793	0.604	0.780	0.7708-0.7899
ppFEV$_1$ GLI-Race specific	0.756	0.659	0.767	0.7573-0.7772
ppFEV$_1$ cg419_AGHR	0.784	0.599	0.766	0.7559-0.7758
ppFEV$_1$ nh3700_AGHR	0.761	0.626	0.763	0.7534-0.7734
ppFEV$_1$ Hankinson	0.751	0.633	0.762	0.7520-0.7721

C

Multivariable Logistic Regression AIC Values

ppFEV$_1$ Variable	All	White	Black
ppFEV$_1$ GLI-Other	8,903	5,400	3,421
ppFEV$_1$ nh3700_AGH	8,907	5,400	3,422
ppFEV$_1$ cg419_AGH	8,909	5,400	3,424
ppFEV$_1$ GLI-Global	8,909	5,399	3,424
ppFEV$_1$ GLI-Race specific	8,967	5,400	3,421
ppFEV$_1$ cg419_AGHR	8,980	5,399	3,424
ppFEV$_1$ nh3700_AGHR	8,985	5,399	3,423
ppFEV$_1$ Hankinson	8,989	5,401	3,422

D

Random Forest Classification Errors

Variable	All COPDGene Phase PI Participants	White COPDGene Phase PI Participants	Black COPDGene Phase PI Participants
	Classification Error % (CI)	Classification Error % (CI)	Classification Error % (CI)
ppFEV1 GLI-Race specific	24.6 (23.7-25.4)	21.6 (20.6-22.6)	31.4 (29.8-33.1)
ppFEV1 Hankinson	24.3 (23.4-25.2)	21.6 (20.6-22.6)	31.7 (30.0-33.4)
ppFEV1 cg419_AGHR	24.2 (23.3-25.0)	21.3 (20.2-22.3)	32.0 (30.3-33.7)
ppFEV1 nh3700_AGHR	24.4 (23.5-25.3)	21.6 (20.6-22.7)	31.5 (29.9-33.2)
ppFEV1 GLI-Other	24.1 (23.3-25.0)	21.6 (20.6-22.6)	31.5 (29.8-33.2)
ppFEV1 cg419_AGH	24.8 (23.9-25.7)	21.3 (20.3-22.4)	31.5 (29.9-33.2)
ppFEV1 nh3700_AGH	24.3 (23.4-25.2)	21.8 (20.8-22.9)	31.5 (29.9-33.2)
ppFEV1 GLI-Global	24.2 (23.3-25.1)	21.5 (20.5-22.5)	32.3 (30.7-34.1)

Figure 3 – A, ROC curves of dyspnea in COPDGene PI participants who smoked as predicted by ppFEV$_1$. Each colored line corresponds to a ROC curve using ppFEV$_1$ values derived from the listed race-specific and race-neutral model equations. B, Sensitivity, specificity, and AUC of the ROC curve of the ppFEV$_1$ to predict dyspnea (ie, mMRC score > 1). C, AIC from the multivariable logistic regression models of dyspnea (ie, mMRC score > 1), with the following covariates: FEV$_1$/FVC ratio, smoking history (pack-y), scanner maker, smoking status, gender, and the ppFEV$_1$. D, Classification error rates are presented from supervised random forest models of dyspnea (ie, mMRC score > 1), which were generated using the same covariates as the logistic regression models, for the total COPDGene PI cohort, and for White and Black participants individually. AGH = age, gender, and height; AGHR = age, gender, height, and race; AUC = area under the curve; AIC = Akaike information criterion; COPDGene PI = COPDGene phase I; cg419 = COPDGene data set of 419 healthy individuals; GLI = Global Lung Initiative; mMRC = modified Medical Research Council; nh3700 = National Health and Nutrition Examination Survey data set of 3,700 healthy individuals; ppFEV$_1$ = percent predicted FEV$_1$; ROC = receiver-operating characteristic.

data sets, consistent with earlier findings.[9,29] Similarly, Baugh et al[7] found that race-neutral equations improved prediction of respiratory symptom burden in individuals who smoke at risk for COPD. Another prospective study identified no benefits of race-specific equations for predicting symptom burden or mortality from chronic lower respiratory disease in a mixed-race population.[8] Finally, Liu et al[10] noted that race-specific equations relative to race-neutral equations may be underdiagnosing emphysema among Black participants.

Others justify the ongoing use of race/ethnicity to determine spirometry reference values, citing risk of discriminatory hiring practices, denial of health insurance,[13] or risk of overdiagnosis of lung disease in Black individuals and underdiagnosis in White individuals.[12] We believe that these risks are of lesser potential harm than the risk of underdiagnosing the group with the highest rate of respiratory illness, which already receives poorer pulmonary care.[30,31] There is an urgent moral obligation to reduce risk to the group most

historically harmed and also most at risk for lung disease. In terms of discriminatory hiring, our primary concern is for the health of the individuals who may be at risk of exacerbating illness if granted hazardous jobs. All patients, regardless of race or baseline lung function, should be evaluated prior to being granted a hazardous job, and followed up longitudinally with risk counseling. Clinicians should also not rely on a single predicted spirometry value for risk assessment and instead gather longitudinal measures and contextualize clinical findings with lifetime exposure data, particularly for patients with borderline values.

Our data, along with the growing base of similar literature,[7-10] support discontinuation of use of race-specific equations, particularly in clinical settings. Although including race or genetic ancestry may minimally improve model fit,[32] which may be of interest in certain epidemiologic research contexts, the use of race-specific equations has no demonstrable *clinical benefit* in improving diagnosis or classification of lung disease. Conversely, there is risk of clinical harm in continuing the use of race-specific equations for Black and other/mixed race patients who may be underdiagnosed for lung disease, as also shown in other studies.[7,20,21] In particular, we found that a shift away from race-specific equations has a bigger impact on the mild end of the spectrum of GOLD classification, which may enable clinicians to catch more early-stage disease among Black patients, potentially preventing more disease progression. Moreover, these equations reinforce false assumptions about genetic differences between groups, while obscuring the role of environmental factors.[4]

The assumption that Black individuals have innately lower lung capacity dates back to slavery era observations.[33] This assumption has carried through to modern medicine where biological differences are prioritized over social or environmental factors. Mounting evidence shows that racial/ethnic minority groups are disproportionately exposed to respiratory toxins via air pollution,[34] occupational hazards,[35] and harmful prenatal and childhood exposures, including preterm birth, very low birth weight, *in utero* smoke exposure, and childhood respiratory illnesses.[15,16,36,24] Social stressors, such as community or family violence,[37-39] and socioeconomic disadvantage[19] have also been linked with worse lung function in early life and are likely interacting with genetic and epigenetic effects. These exposures are influenced by structural inequalities that shape living conditions among marginalized groups. Until we see strong and specific genetic evidence for innate racial differences in lung function, we believe it is not justified to use different criteria to diagnose non-White individuals. These social/environmental factors contribute to reduced lung function among minority racial/ethnic groups, and race-specific equations can mask their damaging effects. Use of ancestry instead of race in these equations would not solve this problem, as ancestral alleles track with the same environmental factors as race.[7] Even if there were ancestral anthropometric differences, the consequences of these variations are not clearly linked to disease or dysfunction. Furthermore, not all individuals are likely to fit the average expectation for the group,[5] particularly when the racial/ethnic populations used to create these equations originally were relatively small and not necessarily all healthy.[1]

Interpretation

As this debate continues, race-specific equations are still recommended by the most recent US and European guidelines[14] and are used in clinical care and pulmonary research worldwide. Although recent studies recommend the race-neutral GLI-Other equation for universal use,[7,8] and the latest race-neutral GLI "Global" equation (based on an equally weighted and more balanced racial/ethnic dataset than GLI-Other),[40] these equations are not race free as GLI-Other applies a universal race correction, and GLI-Global uses weights to balance racial/ethnic diversity. These approaches both still consider race in their formation, and thus assume racial differences in respiratory physiology in healthy individuals, as opposed to the race-free approach. Our findings suggest that race-free equations may serve equally well, at least for adults aged > 35 years.

The primary strengths of our study are its careful selection of healthy individuals for developing race-free reference equations by eliminating asymptomatic individuals with abnormal airway physiology and the consistent findings across two distinct and relatively large data sets with variable racial/ethnic proportions, using multiple modeling techniques. However, our findings should be interpreted in light of certain limitations. The Black population sample in the COPDGene data was relatively small, and the available data sets were not representative of all racial/ethnic groups. Socioeconomic and environmental factors were outside the scope of our study but are clearly important variables to examine in future studies.[19,35,41,42] Although our analyses were cross-sectional, and thus cannot

address directions of effect, we always used a distinct data set to test the fit of equations developed in an independent data set. Future studies should also use prospective longitudinal data to improve the predictive potential of these equations.

With rising awareness of structural racism and misconceptions about race in medicine,[43-46] this is a critical moment for pulmonary clinicians to reconsider the value of continuing to use race when interpreting spirometry measures. The effect of adding race as a covariate only marginally improves the fit for some models with the risk of introducing bias driven by environment and social factors. In light of these concerns, along with the large amount of unexplained variability and the dynamic nature of self-identified race, we maintain that continued use of race-specific equations is not justified. The findings presented here contribute to the growing literature that we hope will be considered when revisions to clinical guidelines are made.

References

1. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general US population. *Am J Respir Crit Care Med*. 1999;159(1): 179-187.

2. Quanjer PH, Stanojevic S, Cole TJ, et al. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the Global Lung Function 2012 equations. *Eur Respir J*. 2012;40(6): 1324-1343.

3. Anderson MA, Malhotra A, Non AL. Could routine race-adjustment of spirometers exacerbate racial disparities in COVID-19 recovery? *Lancet Respir Med*. 2021;9(2):124-125.

4. Bhakta NR, Kaminsky DA, Bime C, et al. Addressing race in pulmonary function testing by aligning intent and evidence with practice and perception. *Chest*. 2022;161(1):288-297.

5. Braun L. *Breathing Race into the Machine: The Surprising Career of the Spirometer from Plantation to Genetics*. Minneapolis, MN: University of Minnesota Press; 2014.

6. Beaverson S, Ngo VM, Pahuja M, Dow A, Nana-Sinkam P, Schefft M. Things We Do for No Reason™: race adjustments in calculating lung function from spirometry measurements. *J Hosp Med*. 2023;18(9): 845-847.

7. Baugh AD, Shiboski S, Hansel NN, et al. Reconsidering the utility of race-specific lung function prediction equations. *Am J Respir Crit Care Med*. 2022;205(7): 819-829.

8. Elmaleh-Sachs A, Balte P, Oelsner EC, et al. Race/ethnicity, spirometry reference equations and prediction of incident clinical events: the Multi-Ethnic Study of Atherosclerosis (MESA) Lung Study. *Am J Respir Crit Care Med*. 2022;205(6): 700-710.

9. McCormack MC, Balasubramanian A, Matsui EC, Peng RD, Wise RA, Keet CA. Race, lung function, and long-term mortality in the National Health and Nutrition Examination Survey III. *Am J Respir Crit Care Med*. 2022;205(6): 723-724.

10. Liu GY, Khan SS, Colangelo LA, et al. Comparing racial differences in emphysema prevalence among adults with normal spirometry: a secondary data analysis of the CARDIA Lung Study. *Ann Intern Med*. 2022;175(8):1118-1125.

11. Ekström M, Mannino D. Race-specific reference values and lung function impairment, breathlessness and prognosis: analysis of NHANES 2007-2012. *Respir Res*. 2022;23(1):271.

12. Miller MR, Graham BL, Thompson BR. Race/ethnicity and reference equations for spirometry. *Am J Respir Crit Care Med*. 2022;206(6):790-792.

13. Townsend MC, Cowl CT. US occupational historical perspective on race and lung function. *Am J Respir Crit Care Med*. 2022;206(6):789-790.

14. Stanojevic S, Kaminsky DA, Miller MR, et al. ERS/ATS technical standard on interpretive strategies for routine lung function tests. *Eur Respir J*. 2022;60(1): 2101499.

15. Barker DJ, Godfrey KM, Fall C, Osmond C, Winter PD, Shaheen SO. Relation of birth weight and childhood respiratory infection to adult lung function and death from chronic obstructive airways disease. *BMJ*. 1991;303(6804):671-675.

16. Lovasi GS, Diez Roux AV, Hoffman EA, Kawut SM, Jacobs DR, Barr RG. Association of environmental tobacco smoke exposure in childhood with early emphysema in adulthood among nonsmokers: the MESA-Lung Study. *Am J Epidemiol*. 2010;171(1):54-62.

17. Martinez FD. Early-life origins of chronic obstructive pulmonary disease. *N Engl J Med*. 2016;375(9):871-878.

18. Fuller-Thomson E, Chisholm RS, Brennenstuhl S. COPD in a population-based sample of never-smokers: interactions among sex, gender, and race. *Int J Chronic Dis*. 2016;2016:5862026.

19. Rocha V, Soares S, Stringhini S, Fraga S. Socioeconomic circumstances and respiratory function from childhood to early adulthood: a systematic review and meta-analysis. *BMJ Open*. 2019;9(6): e027528.

20. Mamary AJ, Stewart JI, Kinney GL, et al. Race and gender disparities are evident in COPD underdiagnoses across all severities of measured airflow obstruction. *Chronic Obstr Pulm Dis*. 2018;5(3):177-184.

21. Moffett AT, Eneanya ND, Halpern SD, Weissman GE. The impact of race correction on the interpretation of pulmonary function testing among black patients. *Am J Respir Care Med*. 2021;203: A1030.

22. Hoesein FAAM, de Jong PA, Lammers J-WJ, et al. Airway wall thickness associated with forced expiratory volume in 1 second decline and development of airflow limitation. *Eur Respir J*. 2015;45(3): 644-651.

23. Koenker R, Machado JAF. Goodness of fit and related inference processes for quantile regression. *J Am Stat Assoc*. 1999;94(448):1296-1310.

24. Chan JYC, Stern DA, Guerra S, Wright AL, Morgan WJ, Martinez FD. Pneumonia in childhood and impaired lung function in adults: a longitudinal study. *Pediatrics*. 2015;135(4):607-616.

25. Han MK, Agusti A, Celli BR, et al. From GOLD 0 to pre-COPD. *Am J Respir Crit Care Med*. 2021;203(4):414-423.

26. Lowe KE, Regan EA, Anzueto A, et al. COPDGene® 2019: redefining the diagnosis of chronic obstructive pulmonary disease. *Chronic Obstr Pulm Dis*. 2019;6(5):384-399.

27. Zach JA, Newell JD, Schroeder J, et al. Quantitative computed tomography of the lungs and airways in healthy nonsmoking adults. *Invest Radiol*. 2012;47(10):596-602.

28. Hoffman EA, Ahmed FS, Baumhauer H, et al. Variation in the percent of emphysema-like lung in a healthy, nonsmoking multiethnic sample. The MESA Lung Study. *Ann Am Thorac Soc*. 2014;11(6):898-907.

29. Gaffney AW, McCormick D, Woolhandler S, Christiani DC, Himmelstein DU. Prognostic implications of differences in forced vital capacity in black and white US adults: findings from NHANES III with long-term mortality follow-up. *EClinicalMedicine*. 2021;39: 101073.

30. Odonkor CA, Esparza R, Flores LE, et al. Disparities in health care for black patients in physical medicine and rehabilitation in the United States: a narrative review. *PM R*. 2021;13(2): 180-203.

31. Lederer DJ, Benn EKT, Barr RG, et al. Racial differences in waiting list outcomes in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2008;177(4): 450-454.

32. Witonsky J, Elhawary JR, Eng C, Rodriguez-Santana J, Borrell LN, Burchard EG. Genetic ancestry to improve precision of race/ethnicity-based lung function equations in children. *Am J Respir Crit Care Med*. 2022;205(6): 725-727.

33. Braun L. Race, ethnicity and lung function: a brief history. *Can J Respir Ther*. 2015;51(4):99-101.

34. Nardone A, Casey JA, Morello-Frosch R, Mujahid M, Balmes JR, Thakur N. Associations between historical residential redlining and current age-adjusted rates of emergency department visits due to asthma across eight cities in California: an ecological study. *Lancet Planet Health*. 2020;4(1):e24-e31.

35. Ekenga CC, Yeung CY, Oka M. Cancer risk from air toxics in relation to neighborhood isolation and sociodemographic characteristics: a spatial analysis of the St. Louis metropolitan area, USA. *Env Res*. 2019;179(pt B):108844.

36. Anand D, Stevenson CJ, West CR, Pharoah POD. Lung function and respiratory health in adolescents of very low birth weight. *Arch Dis Child*. 2003;88(2):135-138.

37. Wright RJ. Perinatal stress and early life programming of lung structure and function. *Biol Psychol*. 2010;84(1):46-56.

38. Kariuki SM, Gray DM, Newton CRJC, et al. Association between maternal psychological adversity and lung function in South African infants: a birth cohort study. *Pediatr Pulmonol*. 2020;55(1): 236-244.

39. Suglia SF, Ryan L, Laden F, Dockery DW, Wright RJ. Violence exposure, a chronic psychosocial stressor, and childhood lung function. *Psychosom Med*. 2008;70(2): 160-169.

40. Bowerman C, Bhatka NR, Brazzale D, et al. A race-neutral approach to the interpretation of lung function measurements. *Am J Respir Crit Care Med*. 2023;207(6):768-774.

41. Hegewald MJ, Crapo RO. Socioeconomic status and lung function. *Chest*. 2007;132(5):1608-1614.

42. Johannessen A, Eagan TML, Omenaas ER, Bakke PS, Gulsvik A. Socioeconomic risk factors for lung function decline in a general population. *Eur Respir J*. 2010;36(3):480-487.

43. Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight—reconsidering the use of race correction in clinical algorithms. *N Engl J Med*. 2020;383(9): 874-882.

44. Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. *Lancet*. 2017;389(10077):1453-1463.

45. Deyrup A, Graves JL. Racial biology and medical misconceptions. *N Engl J Med*. 2022;386(6):501-503.

46. Cerdeña JP, Grubbs V, Non AL. Racialising genetic risk: assumptions, realities, and recommendations. *Lancet*. 2022;400(10368):2147-2154.

47. Regan EA, Hokanson JE, Murphy JR, et al. Genetic epidemiology of COPD (COPDGene) study design. *COPD*. 2010;7(1):32-43.