Original Contribution

Similar Relation of Age and Height to Lung Function Among Whites, African Americans, and Hispanics

Elizabeth M. Kiefer, John L. Hankinson, and R. Graham Barr*

* Correspondence to Dr. R. Graham Barr, Columbia University Medical Center, 630 West 168th Street, PH 9 East Room 105, New York, NY 10032 (e-mail: rgb9@columbia.edu).

Current guidelines recommend separate spirometry reference equations for whites, African Americans, and Mexican Americans, but the justification for this recommendation is controversial. The authors examined the statistical justification for race/ethnic-specific reference equations in adults in the Third National Health and Nutrition Examination Survey (1988–1994) and the Multi-Ethnic Study of Atherosclerosis Lung Study (2000–2006). Spirometry was measured following American Thoracic Society guidelines. "Statistical justification" was defined as the presence of effect modification by race/ethnicity among never-smoking participants without respiratory disease or symptoms and was tested with interaction terms for race/ethnicity ($\times$ age and height) in regression models. There was no evidence of effect modification by race/ethnicity for forced expiratory volume in 1 second, forced vital capacity, or the forced expiratory volume in 1 second/forced vital capacity ratio among white, African-American, and Mexican-American men or women on an additive scale or a log scale. Interaction terms for race/ethnicity explained less than 1% of variability in lung function. The mean lung function for a given age, gender, and height was the same for whites and Mexican Americans but was lower for African Americans. Findings were similar in the Multi-Ethnic Study of Atherosclerosis Lung Study. The associations of age and height with lung function are similar across the 3 major US race/ethnic groups. Multiethnic rather than race/ethnic-specific spirometry reference equations are applicable for the US population.

African Americans; age groups; body height; European continental ancestry group; Hispanic Americans; respiratory function tests; spirometry

Abbreviations: ATS, American Thoracic Society; ERS, European Respiratory Society; FEV$_1$, forced expiratory volume in 1 second; FEV$_6$, forced expiratory volume in 6 seconds; FVC, forced vital capacity; MESA, Multi-Ethnic Study of Atherosclerosis; NHANES, National Health and Nutrition Examination Survey; NHANES III, Third National Health and Nutrition Examination Survey.

Editor's note: *An invited commentary on this article appears on page 388, and the authors' response is published on page 391.*

The American Thoracic Society (ATS) and the European Respiratory Society (ERS) recommend that predicted spirometry reference values be obtained from reference equations derived from samples of similar race/ethnicity whenever possible. This recommendation is based on observations that there are race/ethnic differences in lung function not fully accounted for by sitting height, anthropo-metric, or environmental factors (1–5). However, these race/ethnic differences have been translated into separate spirometry reference equations without clear statistical and scientific justification.

With recent renewed interest in deriving a new set of internationally applicable reference equations, the approach to race/ethnicity will be an important, challenging component of this enterprise. In the United States, the recommended equations were derived in the Third National Health and Nutrition Examination Survey (NHANES III) separately for white, African-American, and Mexican-American participants (6). However, separate spirometry reference equations

for different race/ethnic groups can be problematic for several reasons.

First, they are derived from subgroups rather than full cohorts with a concomitant loss of precision. For example, the sample size of healthy African-American women used to derive the current race/ethnic-specific reference equations was 701 compared with 2,402 women in the full NHANES III sample (6).

Second, it is unclear how to apply these equations in persons of more than one race/ethnicity, which includes a growing proportion of the world population. For example, when the US Census Bureau first allowed reporting of more than one race in 2000, 7 million people self-identified as multiracial (7). This proportion has likely grown in the last decade. Furthermore, in the United States, race (e.g., white or African American) is also newly classified separately from ethnicity (e.g., Hispanic or non-Hispanic), which means that the NHANES III reference equation groups of non-Hispanic whites, African Americans, and Hispanics no longer correspond to US Census definitions of race/ethnicity (8).

Finally, differences in the relation of age and height to lung function by race/ethnicity—as opposed to differences in mean lung volumes by race/ethnicity—are sparsely reported in the literature.

The primary objective of this paper was to investigate statistically the role of race/ethnicity in the derivation of reference equations in a multiethnic population in order to provide greater understanding of the role of race/ethnicity in spirometry reference equations. Separate equations by race/ethnicity would be justified statistically and scientifically if the relations of age and height to lung function were modified by race/ethnicity—in other words, if the slopes of the relation of age and height to the forced expiratory volume in 1 second (FEV_1), for example, differed by race/ethnicity. We therefore tested for effect modification by race/ethnicity in NHANES III, a representative sample of the US population, and examined the amount of variability explained by interaction terms for race/ethnicity. We replicated these analyses in a large multiethnic cohort, the Multi-Ethnic Study of Atherosclerosis (MESA) Lung Study.

MATERIALS AND METHODS

Study samples

The NHANES III study sample included all adult patients who were lifelong nonsmokers without respiratory symptoms or disease and with valid spirometry (at least 2 acceptable maneuvers), following the criteria used in the original NHANES III reference paper (Figure 1) (6). The NHANES III design and methods have been published previously (9). Women and men aged 23 years or greater were included to ensure that participants had all reached peak lung age (10). The very small number of Hispanics of non-Mexican origin were excluded, as previously reported for spirometry analyses in NHANES III (6, 11).

MESA is a multicenter, prospective cohort studying the prevalence and progression of subclinical cardiovascular disease in individuals without clinical cardiovascular dis-

ease (12). In 2000–2002, MESA recruited 6,814 men and women aged 45–84 years from 6 US communities. Excluded were those with clinical cardiovascular disease, pregnancy, weight greater than 300 pounds, and other factors precluding long-term participation (12). The MESA Lung Study randomly sampled 3,965 MESA participants who consented to genetic analyses, underwent baseline measures of endothelial function, and attended an examination during 2004–2006 (99%, 89%, and 91% of the MESA cohort, respectively). Chinese Americans were oversampled to improve precision of estimates in that group. The NHANES III exclusions of smoking or respiratory symptoms or disease were applied to the MESA Lung Study, yielding 1,068 adult nonsmokers without respiratory symptoms or disease.

Spirometry

Spirometry in NHANES III was measured according to and exceeding ATS 1994 guidelines (6). A minimum of 5 acceptable maneuvers were performed in the standing position on a dry-rolling sealed spirometer, by using software with automated quality checks (Occupational Marketing, Inc., Houston, Texas). Spirometry in the MESA Lung Study was measured according to and exceeding the ATS/ERS 2005 guidelines (13). The methods and equipment in the MESA Lung Study were identical to those of NHANES III, except that in the former study at least 3 acceptable maneuvers were performed in the seated position with less stringent end-of-test criteria in order to strictly follow the ATS/ERS 2005 recommendation. Spirometry in the MESA Lung Study has previously been shown to yield results consistent with those of the National Health and Nutrition Examination Survey (NHANES) (14). All participants' recorded spirometry maneuvers in both samples were reviewed for quality by 1 author (J. L. H.).

Race/ethnic groups

Race/ethnicity was self-reported in NHANES III as non-Hispanic white, non-Hispanic black, Mexican American, or other (6). In MESA, race/ethnicity was self-reported as white, African American, Chinese American, and Hispanic of Mexican origin or non-Mexican origin (Dominican, Puerto Rican, Cuban, or other, based on the country of origin of the participant or the participant's family) (12).

Other information

"Smoking history" in NHANES III was defined as an affirmative answer to ever smoking (cigarette, cigar, or pipe) and current smoking (within 5 days) and, in the MESA Lung Study, as a lifetime history of greater than 100 cigarettes, 20 cigars, or 20 pipes of tobacco. Respiratory diagnoses and symptoms were assessed in both studies with standard questionnaires. Height, weight, and body mass index were measured at the time of spirometry by using standard methods and calibrated scales (15).

Statistical analysis

Three main statistical evaluations were performed. Linear regression was used to model the relations of age, age^2, and
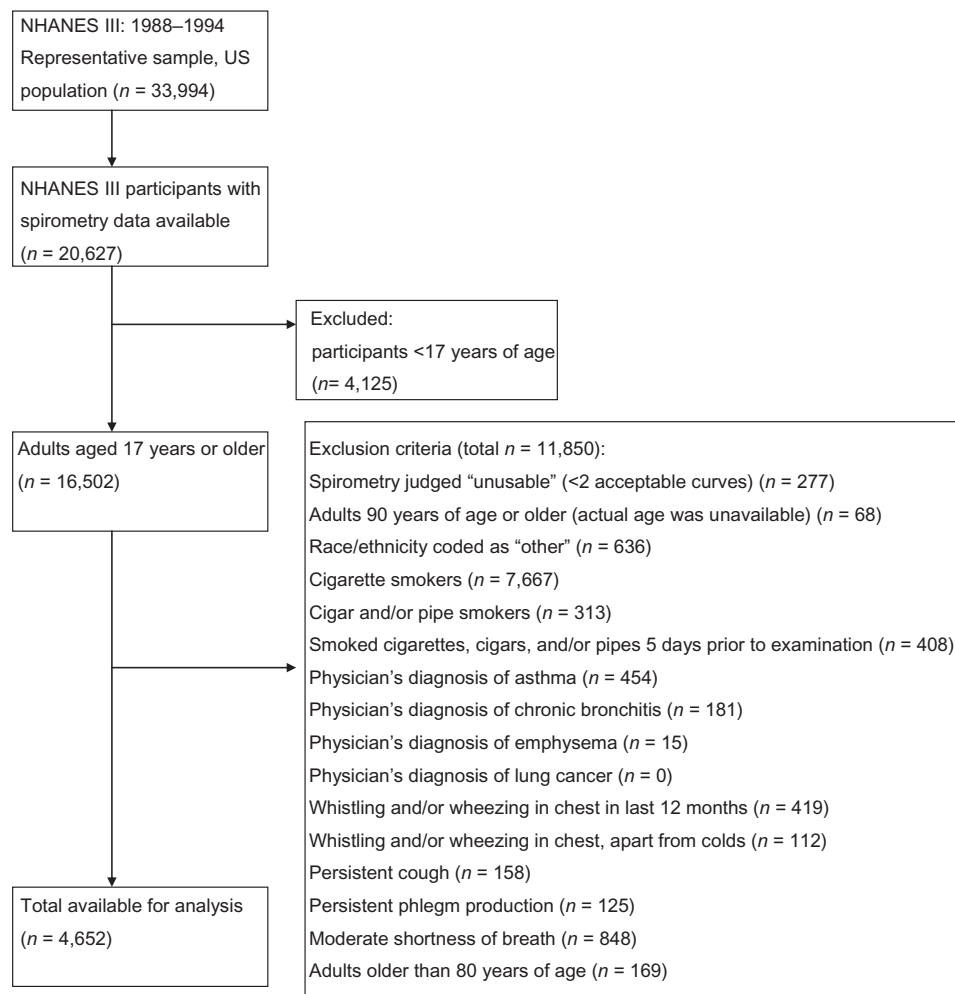
Figure 1. Number of adults excluded according to exclusion criteria, NHANES III, 1988–1994. NHANES III, Third National Health and Nutrition Examination Survey.

height^2 (the dependent variables traditionally used in this literature), as well as sex and race/ethnicity, to FEV_1, forced vital capacity (FVC), and the FEV_1/FVC ratio. We also examined FEV_6 because of its increasing use as a surrogate for the FVC. We tested first for effect modification, or differences in the relation of age and height to lung function, with interaction terms for race/ethnicity. We then evaluated the variability explained by those interaction terms in regression models. Finally, we examined the precision of predicted estimates for lung function.

Outcomes were normally distributed, and model assumptions were met. Diagnostics were performed, showing no major outliers. Likelihood ratio tests were used to compare fully saturated models, which included all main effect terms and interaction terms for each race/ethnicity, with a model without interaction terms for a specific race/ethnicity. Interaction terms by race/ethnicity were created by multiplying race/ethnicity by other predictors of spirometry in the equation (e.g., age × African American, age^2 × African American, height^2 × African American). Nonsignificant ($P \geq 0.05$) interaction terms and main effect terms were

dropped; least significant terms were dropped first. We tested for interactions on both an additive and multiplicative scale. Models were constructed first on an additive scale and then on a natural log (multiplicative) scale by using the natural log of each spirometry measure in regression models. We tested for model fit based on r-squared results in regression analysis and compared model results.

Reference equations for the FEV_1 and FVC were derived separately for men and women, because the interaction terms by gender were significant ($P < 0.05$) on both additive and log scales. The relations of age and height to the FEV_1/FVC ratio did not differ significantly by gender in NHANES III on either scale, and therefore one equation was derived for the FEV_1/FVC ratio for both men and women.

The proportion of the total variability explained by age, age^2, and height^2 terms, race/ethnic terms, and interaction terms of race/ethnicity was calculated by dividing the regression sum of squares by the total sum of squares.

The precision of predicted estimates for lung function using the equations derived here was compared with the precision of predicted estimates using the currently recommended separate

race/ethnic equations. The mean width of the 95% confidence interval for predicted values given average age and height was calculated by using the healthy men and women in the NHANES III derivation cohort.

The same analytical methods were used in the MESA Lung Study. All analyses were performed by using SAS, version 9.1, software (SAS Institute, Inc., Cary, North Carolina).

The institutional review boards at all participating centers approved the study.

RESULTS

Sample characteristics

Table 1 shows the characteristics of healthy adults in NHANES III and the MESA Lung Study samples, stratified by race/ethnicity. Overall, whites had higher mean spirometry values compared with other race/ethnic groups in NHANES III. Whites were older and Mexican Americans were shorter on average compared with other groups. African-American women had a higher mean body mass index compared with white women and Mexican-American women. Participants in the MESA Lung Study were older and had higher body mass indexes and lower spirometry values compared with NHANES III participants. In the MESA Lung Study sample, whites had higher mean spirometry values compared with other race/ethnic groups, including Chinese Americans and Hispanics of non-Mexican origin.

Effect modification for the FEV_1 and FVC by race/ethnicity in NHANES III

There was no evidence of effect modification for the FEV_1 in NHANES III among women on an additive scale; the relation of age and height to the FEV_1 did not differ between Mexican-American and white women ($P_{\text{interaction}} = 0.58$) or between African-American and white women ($P_{\text{interaction}} = 0.96$). On a log scale, interaction terms for African-American and Mexican-American women were also not significant ($P_{\text{interaction}} = 0.45$ and 0.38, respectively).

The main effects for age, age^2, $height^2$, and African-American race/ethnicity remained significant main effect terms in the prediction equation (Table 2). Plots of final reference equations with and without the nonsignificant linear dependent variables age and height did not reveal significant changes in the shape or nature of the curves. The mean FEV_1 for African-American women differed by 458.8 (95% confidence interval: -491.5, -426.2) mL ($P < 0.001$) from that of white women of the same age and height. The main effect term for Mexican-American women was not significant at -23.33 (95% confidence interval: -61.74, 15.08) mL ($P = 0.23$); thus, the mean FEV_1 was similar for Mexican-American and white women of the same age and height.

Findings were similar for FVC. There was no evidence of interaction by African-American and Mexican-American race/ethnicity on an additive or log scale; r-squared values were similar on an additive and log scale for FEV_1

($r^2 = 0.65$ and 0.66, respectively) and FVC ($r^2 = 0.60$ and 0.59, respectively). Table 2 shows the single equation derived for FEV_1 and the single equation for FVC for white, African-American, and Mexican-American women, presented on an additive scale, consistent with the 1999 NHANES equations.

Among men for the FEV_1, there was no evidence of effect modification on an additive scale or a log scale by Mexican-American race/ethnicity ($P = 0.78$ and 0.55, respectively), signifying that the relation of age and height to FEV_1 did not differ between whites and Mexican Americans. There was no evidence of effect modification by African-American race/ethnicity on an additive scale or a log scale ($P = 0.11$ and 0.58, respectively) (Table 3). Similar to the findings among women, the main effect term for Mexican-American men was not significant in either the additive or multiplicative model and was dropped from the final regression equation, but the main effect term remained significant for African Americans. Thus, the predicted FEV_1 for Mexican-American men did not differ from that for white men of the same age and height, but the predicted FEV_1 for African Americans was approximately 600 mL less compared with that for white men of the same age and height. Findings were similar for FVC. Among men, the r-squared values were similar on an additive and log scale for FEV_1 ($r^2 = 0.61$ and 0.59, respectively) and FVC ($r^2 = 0.58$ and 0.56, respectively). Table 3 shows the single reference equation for FEV_1 and the FVC on an additive scale, consistent with the 1999 NHANES equations.

Results for FEV_6, which has been proposed by some as a surrogate for FVC, are shown in Web Tables 1 and 2 posted on the *Journal*'s Web site, http://aje.oupjournals.org, for NHANES III. Interaction terms for African-American and Mexican-American race/ethnicity were not significant for women on either scale or for men on a log scale. However, interaction terms for African-American men were significant on an additive scale ($P = 0.02$).

Effect modification for the FEV_1/FVC ratio in NHANES III

The relation of age and height to the FEV_1/FVC ratio did not differ for Mexican Americans or African Americans on an additive ($P = 0.36$ and $P = 0.37$, respectively) or multiplicative ($P = 0.31$ and $P = 0.26$, respectively) scale. The final model included terms for age, $height^2$, and African-American race/ethnicity (Table 4).

Variance explained by interaction terms for race/ethnicity in NHANES III

The total variability in FEV_1, FVC, FEV_6, and the FEV_1/FVC ratio explained by age, $height^2$, race/ethnicity, and effect modification by race/ethnicity is shown in Table 5. Age and $height^2$ accounted for about 40%–50% of the total variability in lung function, and the African-American race accounted for an additional 10%–18%. In contrast, interaction terms for race/ethnicity explained less than 1% of the variability in lung function after adjustment.

Table 1. Characteristics of Never-Smoking Participants Without Respiratory Diseases or Symptoms in the Third National Health and Nutrition Examination Survey (1988–1994) and the Multi-Ethnic Study of Atherosclerosis Lung Study (2000–2006)[a]

| | Third National Health and Nutrition Examination Survey | | | | | | Multi-Ethnic Study of Atherosclerosis | | | | | | | | |
| | White | | African American | | Mexican American | | White | | African American | | Chinese American | | Hispanic (Mexican) American | | Hispanic (Non-Mexican) American | |
	No.	Mean (SD)	No.	Mean (SD)	No.	Mean (SD)	No.	Mean (SD)	No.	Mean (SD)	No.	Mean (SD)	No.	Mean (SD)	No.	Mean (SD)
Men	449		385		450		118		85		113		42		51	
Age, years		45.8 (17.3)		37.7 (13.7)		38.5 (14.0)		64.2 (9.8)		64.5 (10.6)		64.3 (9.5)		63.1 (10.4)		62.6 (9.6)
Height, cm		176.7 (6.8)		176.8 (7.1)		169.6 (6.6)		175.6 (7.5)		174.1 (7.1)		167.4 (5.4)		168.5 (5.2)		169.3 (7.3)
Weight, kg		83.2 (16.0)		85.1 (17.2)		76.8 (13.1)		83.0 (12.1)		85.9 (14.0)		67.1 (9.8)		80.8 (9.3)		76.9 (12.0)
BMI, kg/m^2		26.6 (4.7)		27.2 (5.0)		26.7 (4.2)		26.8 (3.1)		28.1 (3.7)		23.8 (2.8)		28.4 (3.3)		26.7 (3.5)
FEV$_1$, mL		3,949 (846)		3,594 (693)		3,853 (714)		3,257 (658)		2,693 (526)		2,685 (516)		3,177 (666)		2,968 (632)
FVC, mL		5,016 (937)		4,378 (784)		4,725 (792)		4,331 (802)		3,519 (648)		3,510 (620)		4,070 (765)		3,850 (848)
FEV$_1$/FVC ratio, %		79 (7)		82 (7)		81 (6)		75 (1)		77 (1)		77 (1)		78 (1)		77 (1)
Women	902		701		799		152		125		230		74		76	
Age, years		50.3 (17.2)		41.1 (15.1)		41.3 (15.3)		67.1 (10.7)		66.2 (9.2)		64.9 (9.7)		65.6 (9.4)		65.0 (10.2)
Height, cm		162.0 (6.7)		162.8 (6.3)		156.3 (6.0)		161.7 (6.7)		161.5 (6.3)		155.5 (6.1)		155.4 (6.5)		155.8 (6.4)
Weight, kg		67.9 (14.4)		76.5 (19.9)		66.4 (15.4)		67.6 (11.9)		72.6 (10.8)		57.7 (8.9)		69.0 (10.9)		66.4 (11.6)
BMI, kg/m^2		25.9 (5.5)		28.9 (7.2)		27.2 (6.2)		25.8 (3.9)		27.8 (3.6)		23.7 (3.4)		28.4 (3.5)		27.2 (4.1)
FEV$_1$, mL		2,730 (651)		2,509 (576)		2,727 (568)		2,221 (466)		1,960 (405)		1,856 (424)		2,163 (469)		1,992 (447)
FVC, mL		3,403 (719)		3,021 (649)		3,286 (616)		2,900 (604)		2,475 (514)		2,394 (496)		2,744 (546)		2,529 (499)
FEV$_1$/FVC ratio, %		80 (7)		83 (7)		83 (7)		77 (1)		79 (1)		77 (1)		79 (1)		79 (1)

Abbreviations: BMI, body mass index; FEV$_1$, forced expiratory volume in 1 second; FVC, forced vital capacity; SD, standard deviation.
[a] Table shows the mean (standard deviation) for each group.

Table 2. *P* Values for Main Effect Terms for the Forced Expiratory Volume in 1 Second and the Forced Vital Capacity Among Women in the Third National Health and Nutrition Examination Survey Sample (1988–1994), With Final Reference Equations on an Additive Scale

	P Values for Main Effect Terms					
	Intercept	Age	Age^2	Height^2	African American	Mexican American
FEV$_1$ model, mL						
Full model without interaction terms for race/ethnicity	<0.001	0.0011	<0.001	<0.001	0.23
Without main effect term for Mexican-American ethnicity	<0.001	0.0013	<0.001	<0.001	
FVC model, mL						
Full model without interaction terms for race/ethnicity	0.91	<0.0001	<0.0001	<0.0001	0.077
Without main effect term for age			<0.0001	<0.0001	<0.0001	0.077
Without main effect term for Mexican-American ethnicity			<0.0001	<0.0001	<0.0001	

	Final Reference Equation									
	Intercept		Age		Age^2		Height^2		African American	
	Beta Coefficient	95% CI	Beta Coefficient	95% CI	Beta Coefficient	95% CI	Beta Coefficient	95% CI	Beta Coefficient	95% CI
FEV$_1$ model, mL	836.0	621.9, 1,050.0	−13.67	−19.69, −7.658	−0.0993	−0.1598, −0.0388	0.1086	0.1019, 0.1154	−458.8	−491.5, −426.2
FVC model, mL	−66.12	−146.0, 278.2			−0.2083	−0.2191, −0.1976	0.1488	0.1408, 0.1569	−596.3	−635.7, −556.9

Abbreviations: CI, confidence interval; FEV$_1$, forced expiratory volume in 1 second; FVC, forced vital capacity.

Table 3. *P* Values for Main Effect Terms for the Forced Expiratory Volume in 1 Second and the Forced Vital Capacity Among Men in the Third National Health and Nutrition Examination Survey Sample (1988–1994), With Final Reference Equations on an Additive Scale

	P Values for Main Effect Terms					
	Intercept	Age	Age^2	Height^2	African American	Mexican American
FEV$_1$ model, mL						
Full model without interaction terms for race/ethnicity		0.016	0.012	<0.001	<0.0001	0.57
Without main effect term for Mexican-American ethnicity		0.017	0.011	<0.0001	<0.0001	
FVC model, mL						
Full model without interaction terms for race/ethnicity		0.36	0.0076	<0.0001	<0.0001	0.50
Without main effect term for Mexican-American ethnicity		0.35	0.0086	<0.0001	<0.0001	
Without main effect term for age			<0.0001	<0.0001	<0.0001	

	Final Reference Equation									
	Intercept		Age		Age^2		Height^2		African American	
	Beta Coefficient	95% CI	Beta Coefficient	95% CI	Beta Coefficient	95% CI	Beta Coefficient	95% CI	Beta Coefficient	95% CI
FEV$_1$ model, mL	879.0	486.3, 1,271.7	−13.72	−24.96, −2.483	−0.1514	−0.2675, −0.0353	0.1301	0.1198, 0.1404	−599.8	−659.4, −540.3
FVC model, mL	−74.08	−446.9, 298.7			−0.2476	−0.2686, −0.2266	0.1814	0.1693, 0.1935	−831.3	−901.4, −761.1

Abbreviations: CI, confidence interval; FEV$_1$, forced expiratory volume in 1 second; FVC, forced vital capacity.

Table 4. *P* Values for the Interaction and Main Effect Terms for the Forced Expiratory Volume in 1 Second to Forced Vital Capacity Ratio (in %) Among Men and Women in the Third National Health and Nutrition Examination Survey Sample (1988–1994), With Final Reference Equation

	P Values for Main Effect Terms						
	Intercept	Age	Age^2	Height^2	Sex	African American	Mexican American
FEV$_1$/FVC ratio model							
Full model without race/ethnic or sex interaction terms		<0.0001	0.36	<0.0001	0.17	<0.0001	0.16
Without main effect term for age^2		<0.0001		<0.0001	0.19	<0.0001	0.17
Without main effect term for sex		<0.0001		<0.0001		<0.0001	0.31
With eff for An eth		<0.0001		<0.0001		<0.0001	

	Final Reference Equation									
	Intercept		Age		Age^2		Height^2		African American	
	Beta Coefficient	95% CI	Beta Coefficient	95% CI	Beta Coefficient	95% CI	Beta Coefficient	95% CI	Beta Coefficient	95% CI
FEV$_1$/FVC ratio	101.4	99.57, 103.2	−0.21	−0.23, −0.20			−0.00041	−0.00047, −0.00035	1.2826	0.8493, 1.716

Abbreviations: CI, confidence interval; FEV$_1$, forced expiratory volume in 1 second; FVC, forced vital capacity.

Table 5.   Percentage of Variability in Lung Function in the Third National Health and Nutrition Examination Survey (1988–1994) Explained by Age and Height, Race/Ethnicity, and Interaction Terms for Race/Ethnicity

Variable	Men				Women			
	df	Regression Sum of Squares	Total Sum of Squares	% of Variability Explained by Variable	df	Regression Sum of Squares	Total Sum of Squares	% of Variability Explained by Variable
FEV_1								
Age and height	3	417,470,981	862,406,448	48	3	569,411,839	1,048,891,848	54
Race/ethnicity	2	105,279,869	862,406,448	12	2	110,732,577	1,048,891,848	11
Interaction by race/ethnicity[a]	6	4,596,765	862,406,448	0.53	6	861,943	1,048,891,848	0.08
FVC								
Age and height	3	444,690,627	1,113,558,707	40	3	555,060,439	1,270,794,743	44
Race/ethnicity	2	197,778,933	1,113,558,707	18	2	188,336,757	1,270,794,743	15
Interaction by race/ethnicity[a]	6	6,288,087	1,113,558,707	0.56	6	1,346,520	1,270,794,743	0.11
FEV_6								
Age and height	3	508	1,139	45	3	625	1,311	48
Race/ethnicity	2	187	1,139	16	2	184	1,311	14
Interaction by race/ethnicity	6	7.307	1,139	0.64	6	0.9104	1,311	0.07
FEV_1/FVC ratio								
Age and height	4	74,192	246,600	30	4	74,192	246,600	30
Race/ethnicity	2	1,856	246,600	0.75	2	1,856	246,600	0.75
Interaction by race/ethnicity[a]	8	558	246,600	0.23	8	558	246,600	0.23

Abbreviations: FEV_1, forced expiratory volume in 1 second; FEV_6, forced expiratory volume in 6 seconds; FVC, forced vital capacity.

[a] A significant interaction by race/ethnicity implies differences in the relation of age and height to lung function by race/ethnicity and, if high, would suggest the need for race/ethnic-specific equations.

Precision of predicted values

The width of the 95% confidence interval for predicted values of lung function for average age and height was narrower with the equations derived in the full sample compared with race/ethnic-specific reference equations. For example, the predicted values of the FEV_1 based on the equation derived from the full healthy sample of 2,402 women in NHANES III (Table 2) had a mean width of the 95% confidence interval of 63 mL. By using race/ethnic-specific equations, however, the mean width of the 95% confidence interval was 92 mL for white women, 107 mL for African-American women, and 91 mL for Mexican-American women. Among men, the predicted values of the FEV_1 based on the equation derived from the full healthy sample had a mean width of the 95% confidence interval of 115 mL, compared with 179 mL for white men, 193 mL for African-American men, and 155 mL for Mexican-American men, when using race/ethnic-specific equations.

Sensitivity analysis

To address the issue of selection bias by race/ethnicity or gender in our healthy derivation sample, we conducted a sensitivity analysis among all NHANES III individuals, aged 23–80 years with acceptable spirometry, which included

8,450 women and 7,177 men. Among women, there were no significant interaction terms for Mexican- or African-American race/ethnicity. Among men, however, there were statistically significant interaction terms for African-American race/ethnicity for FEV_1 and FVC on an additive and a log scale, which likely reflects the differential amount of smoking as a cause of reduced lung function in African-American men.

Comparison of predicted values

We show a plot of the predicted FEV_1 versus age for white women, calculated by using the original 1999 NHANES derivation equation for white women, and the equation derived here on an additive scale (Figure 2). We show a similar plot for white men (Figure 3).

Replication in the MESA Lung Study

Similar to NHANES III, the relations of age and height to lung function in the MESA Lung Study differed significantly by gender for FEV_1 and FVC ($P_{interaction} < 0.05$ for both) but not the FEV_1/FVC ratio ($P_{interaction} = 0.39$).

In both men and women in the MESA Lung Study, which included Chinese Americans and Hispanics of non-Mexican origin, there was no evidence for effect modification by race/ethnicity on both an additive and a multiplicative scale
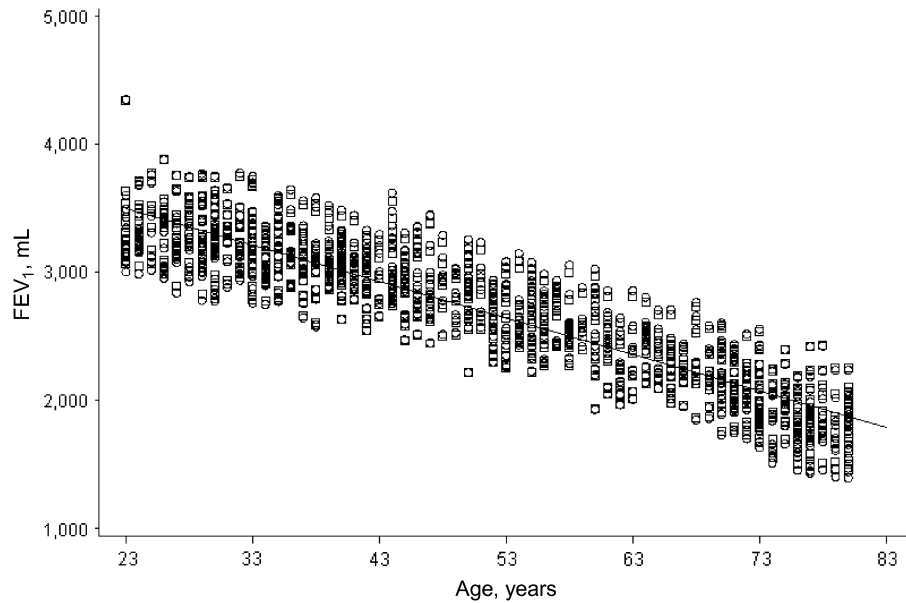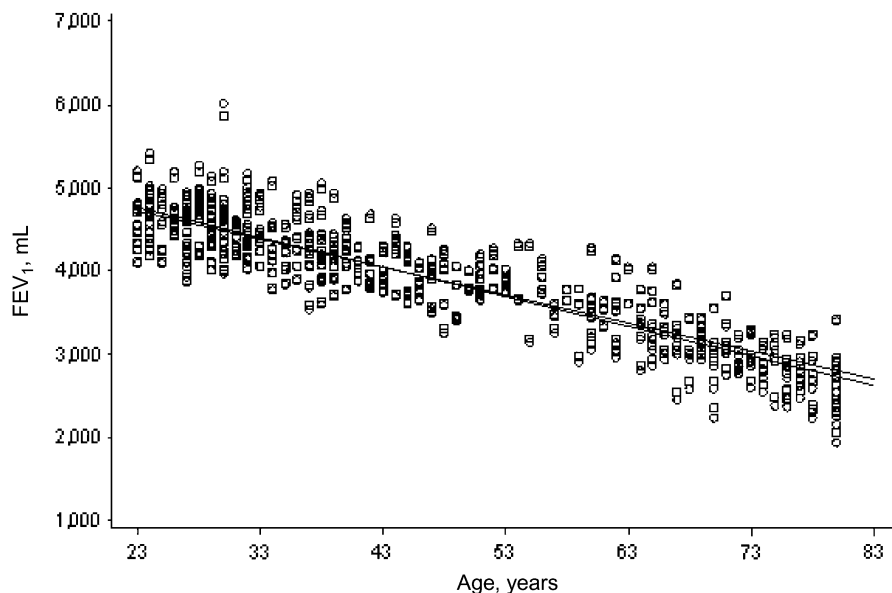
Figure 2. FEV$_1$ predicted from 1999 NHANES equation for white women (circles) and NHANES-derived equation for women, additive scale (squares). FEV$_1$, forced expiratory volume in 1 second; NHANES, National Health and Nutrition Examination Survey.

for FEV$_1$, FVC, or the FEV$_1$/FVC ratio. Both Chinese-American and non-Mexican Hispanic men and women, like African Americans, had lower FEV$_1$ and FVC values compared with whites of the same age and height, necessitating the inclusion of a main effect term for each of these race/ethnicities (Web Table 3). The final reference equations are presented in Web Table 3. Application of the equations derived from NHANES to the MESA cohort demonstrated

similar spirometry values compared with the observed MESA spirometry values.

DISCUSSION

We investigated the role of race/ethnicity in the derivation of reference equations in 2 large multiethnic cohorts. We found that effect modification by race/ethnicity explained

Figure 3. FEV$_1$ predicted from 1999 NHANES equation for white men (circles) and NHANES-derived equation for men, additive scale (squares). FEV$_1$, forced expiratory volume in 1 second; NHANES, National Health and Nutrition Examination Survey.

less than 1% of the variability in lung function, although the mean lung function for a given age and height was lower for participants of African origin compared with whites. Spirometry reference equations that consider race/ethnic groups pooled together are justified statistically and were found to improve precision compared with race/ethnic-specific equations.

When the currently recommended, race/ethnic-specific reference equations were developed more than 10 years ago, consideration was given to deriving a single reference equation for all race/ethnic groups (6). However, the authors ultimately decided to follow the 1991 ATS recommendations to select reference equations on the basis of the "ethnic origin" of the subjects being tested and to avoid potential extrapolation of data (16), given the relatively limited overlap in the range of height of Mexican Americans and whites in NHANES III.

The single, multiethnic equations presented here can address some, but not all, of the problems with race/ethnic-specific reference equations that we previously outlined. First, the multiethnic equations have improved precision because they are derived from the full cohort instead of from smaller samples for each race/ethnicity. Figures 2 and 3 show similar slopes and similar predicted values for the predicted FEV_1 by using the additive equation derived here compared with the 1999 NHANES III-derived equations. However, the mean width of the 95% confidence interval for the predicted FEV_1 among women using separate equations was approximately 100 mL for each race/ethnic group. Although a minimal, clinically important difference for the FEV_1 has not been firmly established, there has been some suggestion that a change in FEV_1 of 100 mL can be perceived by patients (17). Using the single reference equation decreases the width of the confidence interval for predicted FEV_1 in women by approximately one third.

Second, as race/ethnic-specific reference equations are problematic for patients of multiple race/ethnic backgrounds, multiethnic equations would be valuable for patients of mixed race/ethnicity. For some mixed-race/ethnicities, such as Mexican Americans, our findings suggest that mean lung volumes are the same among Mexican Americans and whites of the same age and height. However, main effect terms are required for African Americans, Chinese Americans, and non-Mexican Hispanics, reflecting mean differences in lung function compared with those of whites of the same age and height that are observed, but that are not fully accounted for in the literature (2–5).

The inclusion of race/ethnic terms implies that one would still need to identify an individual's race/ethnicity to determine the inclusion of these terms in the prediction equations. However, the size of the coefficient could be adjusted on the basis of ancestry, something that is not feasible for the race-specific equations (18–20). As ancestral informative markers become increasingly available in research and ultimately clinical settings, this coefficient could be individualized much more easily than could race/ethnic-specific equations.

Finally, despite the longstanding approach of separate reference equations by race/ethnicity, relatively few studies have examined this issue. Jacobs et al. (3) found significant

interactions for FEV_1 and FVC among whites and African Americans by using the interaction terms of race/ethnicity by sex by frame size. However, this 3-way interaction term is difficult to compare with the NHANES III reference equations, because the latter are stratified by gender due to effect modification by gender. A second small study of 80 Asian Americans and whites reported different relations of age and height to FVC (i.e., different slopes for the regression line) but not to FEV_1 (21). This finding was probably a false positive given that we were unable to replicate the findings in the MESA Lung Study, which included a much larger sample of Chinese Americans.

Sitting height has been suggested to account for a proportion of race/ethnic discrepancy. In the 1999 NHANES III reference equation derivations (6), sitting height explained some variability by race/ethnicity in a common equation that included all race/ethnicities but was less accurate than separate equations that did not include sitting height. The addition of sitting height did not improve separate race/ethnic equations that contained standing height. However, sitting height includes only one dimension of thoracic size, and the similar predicted FEV_1/FVC ratio among whites, African Americans, and Mexican Americans suggests that there are proportional differences related to frame size.

A strength of this investigation was the replication of the NHANES III findings in another cohort, the MESA Lung Study, which included non-Mexican Hispanic Americans and Chinese Americans. Both cohorts were multiethnic and used very similar spirometry methods. Predicted spirometry values in MESA calculated from the equations derived here produce values similar to observed MESA spirometry; similar results were previously obtained by applying the 1999 NHANES-derived equations to MESA (14).

We limited this analysis to adults who had likely achieved a plateau in lung function value with age and, hence, avoided the need for spline-based approaches, which are more important for pediatric samples (10). Direct inferences were therefore available from the linear models. However, our general findings are also likely to apply to spline-based approaches.

We acknowledge that there is some variation in predicted spirometry depending on which prediction equation is used. Comparing the 1999 NHANES-derived equations with equations derived here yields similar results for the equation derived on an additive scale, but a difference of greater than 10% on the log scale. This variation is slightly higher than the inherent variability of greater than 5% that can occur within an individual and a similar variability between spirometers. However, the primary intent of this paper was not to create new reference equations, but to fully investigate the issue of race/ethnicity in spirometry.

In summary, our findings show that there was no evidence of interaction by race/ethnicity on an additive or log scale in 2 large multiethnic cohorts. Less than 1% of the variability in lung function was explained by interaction terms for race/ethnicity across the 3 major US race/ethnic groups. Multiethnic rather than race/ethnic-specific spirometry reference equations are applicable for the US population.

ACKNOWLEDGMENTS

REFERENCES

1. Pellegrino R, Viegi G, Brusasco V, et al. Interpretative strategies for lung function tests. *Eur Respir J*. 2005;26(5):948–968.
2. Harik-Khan RI, Fleg JL, Muller DC, et al. The effect of anthropometric and socioeconomic factors on the racial difference in lung function. *Am J Respir Crit Care Med*. 2001; 164(9):1647–1654.
3. Jacobs DR Jr, Nelson ET, Dontas AS, et al. Are race and sex differences in lung function explained by frame size? The CARDIA Study. *Am Rev Respir Dis*. 1992;146(3):644–649.
4. Burchfiel CM, Marcus EB, Sharp DS, et al. Characteristics associated with rapid decline in forced expiratory volume. *Ann Epidemiol*. 1996;6(3):217–227.
5. Abramowitz S, Leiner GC, Lewis WA, et al. Vital capacity in the Negro. *Am Rev Respir Dis*. 1965;92:287–292.
6. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med*. 1999;159(1):179–187.
7. Jones NA, Smith AS. *The Two or More Races Population: 2000. Census 2000 Brief*. Washington, DC: US Census Bureau; 2001. (http://www.census.gov/prod/2001pubs/c2kbr01-6.pdf).
8. Bernstein R, Edwards T. An older and more diverse nation by midcentury. Washington, DC: US Census Bureau; 2008. (http://www.census.gov/newsroom/releases/archives/population/cb08-123.html).
9. National Center for Health Statistics. *NHANES III Reference Manuals and Reports*. Hyattsville, MD: Centers for Disease Control and Prevention; 1996.
10. Stanojevic S, Wade A, Stocks J, et al. Reference ranges for spirometry across all ages: a new approach. *Am J Respir Crit Care Med*. 2008;177(3):253–260.
11. Hankinson JL, Kinsley KB, Wagner GR. Comparison of spirometric reference values for Caucasian and African American blue-collar workers. *J Occup Environ Med*. 1996;38(2): 137–143.
12. Bild DE, Bluemke DA, Burke GL, et al. Multi-Ethnic Study of Atherosclerosis: objectives and design. *Am J Epidemiol*. 2002;156(9):871–881.
13. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J*. 2005;26(2):319–338.
14. Hankinson JL, Kawut SM, Shahar E, et al. 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*. 2010; 137(1):138–145.
15. Ferris BG. Epidemiology Standardization Project (American Thoracic Society). *Am Rev Respir Dis*. 1978;118(6 pt 2): 1–120.
16. Lung function testing: selection of reference values and interpretative strategies. American Thoracic Society. *Am Rev Respir Dis*. 1991;144(5):1202–1218.
17. Donohue JF. Minimal clinically important differences in COPD lung function. *COPD*. 2005;2(1):111–124.
18. Sans M. Admixture studies in Latin America: from the 20th to the 21st century. *Hum Biol*. 2000;72(1):155–177.
19. Salas A, Carracedo A, Richards M, et al. Charting the ancestry of African Americans. *Am J Hum Genet*. 2005; 77(4):676–680.
20. Parra EJ, Marcini A, Akey J, et al. Estimating African American admixture proportions by use of population-specific alleles. *Am J Hum Genet*. 1998;63(6):1839–1851.
21. Korotzer B, Ong S, Hansen JE. Ethnic differences in pulmonary function in healthy nonsmoking Asian-Americans and European-Americans. *Am J Respir Crit Care Med*. 2000; 161(4 pt 1):1101–1108.