Spirometry Reference Equations from the HCHS/SOL (Hispanic Community Health Study/Study of Latinos)

Lisa LaVange^{1*}, Sonia M. Davis¹, John Hankinson², Paul Enright^{3†}, Rebecca Wilson¹, R. Graham Barr⁴, Thomas K. Aldrich^{5†}, Ravi Kalhan⁶, Hector Lemus⁷, Ai Ni^{1§}, Lewis J. Smith⁶, and Gregory A. Talavera⁷

¹Collaborative Studies Coordinating Center, Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; ²Hankinson Consulting Inc., Athens, Georgia; ³University of Arizona, Tucson, Arizona; ⁴Columbia University Medical Center, New York, New York; ⁵Albert Einstein College of Medicine, Bronx, New York; ⁶Northwestern University, Chicago, Illinois; and ⁷San Diego State University, San Diego, California

ORCID ID: 0000-0003-2443-0876 (R.K.).

Abstract

Rationale: Accurate reference values for spirometry are important because the results are used for diagnosing common chronic lung diseases such as asthma and chronic obstructive pulmonary disease, estimating physiologic impairment, and predicting all-cause mortality. Reference equations have been established for Mexican Americans but not for others with Hispanic/Latino backgrounds.

Objectives: To develop spirometry reference equations for adult Hispanic/Latino background groups in the United States.

Methods: The HCHS/SOL (Hispanic Community Health Study/Study of Latinos) recruited a population-based probability sample of 16,415 Hispanics/Latinos aged 18–74 years living in the Bronx, Chicago, Miami, and San Diego. Participants self-identified as being of Puerto Rican, Cuban, Dominican, Mexican, or Central or South American background. Spirometry was performed using standardized methods with central quality control monitoring. Spirometric measures from a subset of 6,425 never-smoking

participants without respiratory symptoms or disease were modeled as a function of sex, age, height, and Hispanic/Latino background to produce background-specific reference equations for the predicted value and lower limit of normal.

Measurements and Main Results: Dominican and Puerto Rican Americans had substantially lower predicted and lower limit of normal values for FVC and FEV₁ than those in other Hispanic/Latino background groups and also than Mexican American values from NHANES III (Third National Health and Nutrition Examination Survey).

Conclusions: For patients of Dominican and Puerto Rican background who present with pulmonary symptoms in clinical practice, use of background-specific spirometry reference equations may provide more appropriate predicted and lower limit of normal values, enabling more accurate diagnoses of abnormality and physiologic impairment.

Keywords: spirometry; reference standards; Hispanic/Latino

(Received in original form October 3, 2016; accepted in final form June 13, 2017)

*Current address for L.L. is U.S. Food and Drug Administration, Silver Spring, Maryland.

†T.K.A. died prior to manuscript acceptance.

‡P.E. is currently retired.

§Current address for A.N. is Memorial Sloan Kettering Cancer Center, New York, New York.

The HCHS/SOL (Hispanic Community Health Study/Study of Latinos) was performed as a collaborative study supported by contracts from the NHLBI to the University of North Carolina (N01-HC65233), the University of Miami (N01-HC65234), Albert Einstein College of Medicine (N01-HC65235), Northwestern University (N01-HC65236), and San Diego State University (N01-HC65237). The following institutes, centers, or offices contribute to the HCHS/SOL through a transfer of funds to the NHLBI: National Center on Minority Health and Health Disparities, the National Institute on Deafness and Other Communication Disorders, the National Institute of Dental and Craniofacial Research, the National Institute of Diabetes and Digestive and Kidney Diseases, the National Institute of Neurological Disorders and Stroke, and the Office of Dietary Supplements. Investigators from the NHLBI of the National Institutes of Health were involved in the design and conduct of the study; in the collection, analysis, and interpretation of the data; and in the review and approval of the manuscript.

Am J Respir Crit Care Med Vol 196, Iss 8, pp 993–1003, Oct 15, 2017

Copyright © 2017 by the American Thoracic Society

Originally Published in Press as DOI: 10.1164/rccm.201610-1987OC on June 14, 2017

Internet address: www.atsjournals.org

At a Glance Commentary

Scientific Knowledge on the

Subject: Asthma and chronic obstructive pulmonary disease have been reported to be more prevalent among Hispanics/Latinos of Puerto Rican and Dominican background living in the United States. Diagnosis of these common chronic lung diseases is typically based in part on comparisons of lung volume measurements derived from spirometric tests to the predicted values and the lower limit of normal values derived from data on adults of Mexican background in NHANES III (Third National Health and Nutrition Examination Survey).

What This Study Adds to the

Field: Nonsmoking men and women of Dominican and Puerto Rican background without respiratory symptoms or disease had substantially worse FVC and FEV₁ than those in other Hispanic/Latino background groups, and reference equations fit for these two groups produced correspondingly lower predicted values and lower limits of normal. These findings suggest that the use of Hispanic/Latino background-specific spirometry reference equations may provide more accurate diagnoses of physiologic lung impairment for U.S. adults of Hispanic/Latino ethnicity.

Spirometry is the most clinically important pulmonary function test for medical diagnoses, and the FVC and FEV₁ are key spirometric parameters. A low FVC is associated with many lung conditions that restrict patients' functionality and quality of life. A low FEV₁/FVC indicates airway obstruction, which occurs with the most

common chronic lung diseases of adults, such as chronic obstructive pulmonary disease (COPD) and poorly controlled asthma. In most large epidemiologic cohort studies, FVC and FEV₁ have been shown to be excellent independent predictors of all-cause and cardiovascular mortality (1–3).

Spirometry reference equations are used to determine the percent predicted values and lower limits of the normal range (LLN) for lung function parameters on the basis of sex, age, and height. Equations from NHANES III (Third National Health and Nutrition Examination Study) (4), which included large numbers of healthy white individuals, Mexican Americans, and African Americans, are widely used throughout North America, as recommended by the 2005 pulmonary function testing guidelines (5). Spirometry reference equations for Hispanics/Latinos from four South American cities and Mexico City were subsequently published by the PLATINO (Latin American Research Project on Lung Obstruction) investigators (6). Further, the Global Lung Initiative (GLI) compiled spirometric results from around the world (including NHANES III) and published spirometry reference equations and LLN values for people aged 4–94 years in several ethnic and racial groups (7). The Multi-Ethnic Study of Atherosclerosis Lung Study researchers applied the NHANES III equations for white individuals, African Americans, and Mexican Americans to healthy never-smokers, and they noted no differences between observed and predicted values for these three racial/ethnic groups (8). Differences were noted for Hispanics/Latinos of non-Mexican background, but the sample size did not support development of reference equations.

Hispanic/Latino adults living in the United States include persons of varied backgrounds, generational statuses, and

countries of origin. Genetic admixtures indicate that continental ancestry proportions differ among Hispanic/Latino background groups (9), which could translate into variation in trunk/leg ratio. Extrinsic factors, such as airborne exposure, nutrition, and socioeconomic status (SES) during childhood, may also vary. These factors have been shown in previous studies to affect lung function in adulthood (10). The HCHS/SOL (Hispanic Community Health Study/Study of Latinos) was designed to address differences in baseline risk factors for and incidence of cardiovascular and pulmonary diseases among subpopulations defined by Hispanic/Latino background/heritage, namely Cuban, Mexican, Puerto Rican, Dominican, and Central and South American. The HCHS/SOL cohort sample selection, designed to yield subpopulation-specific sample sizes sufficient to support comparative analyses, as well as the high-quality spirometry testing procedures used, including identical instruments at all field centers and a central quality assurance program similar to that used in NHANES III, make HCHS/SOL uniquely positioned for an examination of spirometry reference values among U.S. Hispanic/Latino subpopulations.

Methods

Study Design

HCHS/SOL is a community-based probability sample cohort study designed to examine risk factors of chronic disease among Hispanics/Latinos living in the United States (11). The HCHS/SOL researchers recruited 16,415 self-identified Hispanics/Latinos aged 18–74 years residing in four U.S. communities (Bronx, NY; Chicago, IL; Miami, FL; and San Diego, CA) between March 2008 and June 2011. Households were selected using a stratified two-stage area probability sample (12, 13).

Author Contributions: L.L. was principal investigator (PI) of the HCHS/SOL (Hispanic Community Health Study/Study of Latinos) Coordinating Center until 2011 and contributed to protocol development, study design, cohort sample selection, data collection and management, statistical analysis, and manuscript writing; P.E. was PI of the HCHS/SOL Spirometry Reading Center and later a coinvestigator, developed the manual of operations for spirometry data collection and quality control oversight, conducted spirometry training and site visits, and contributed to manuscript writing; J.H. was coinvestigator of the HCHS/SOL Spirometry Reading Center, was responsible for spirometry data quality control and analysis, and contributed to manuscript writing; S.M.D., A.N., and R.W. were members of the HCHS/SOL Coordinating Center, conducted the statistical analyses, and contributed to manuscript writing; R.G.B. was PI of the HCHS/SOL Spirometry Reading Center after P.E. and contributed to spirometry data collection and quality control oversight; G.A.T. was PI of the San Diego Field Center; and T.K.A., H.L., R.K., and L.J.S. contributed to spirometry data collection. All authors provided critical revisions.

Correspondence and requests for reprints should be addressed to Sonia M. Davis, Dr.P.H., Collaborative Studies Coordinating Center, Department of Biostatistics, University of North Carolina at Chapel Hill, 137 East Franklin Street, Suite 203, Chapel Hill, NC 27599-8030. E-mail: sonia.davis@unc.edu

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Assessments

Information covering demographics, country of origin, years of residence in the United States, and medical history was obtained by questionnaire. Hispanic heritage was ascertained through self-identification. Height and weight were measured following a standard protocol that included central training and certification. A modified American Thoracic Society Division of Lung Disease respiratory questionnaire (the ATS-DLD-78) (14) was used to ascertain smoking status as well as respiratory diagnoses and symptoms (see online supplement). Spirometry was conducted in accordance with American Thoracic Society (ATS) and European Respiratory Society guidelines (15) using a dry rolling seal spirometer with automated quality checks (Occupational Marketing, Inc., Houston, TX) following the NHANES protocol (4) with minor modification for use of a three-curve rather than a five-curve acceptability minimum. The prevalence of COPD and asthma was described previously by Barr and colleagues (16).

Analysis Population

Reference equations are based on a healthy subset of participants without respiratory symptoms or disease, with no significant smoking history, and with good quality prebronchodilator (pre-BD) spirometry (quality grade C or better [17]). Exclusions, consistent with previous studies as being associated with significantly reduced lung function (18), included being a current smoker or having a smoking history greater than 10 pack-years, COPD, chronic bronchitis, current asthma, current wheeze, shortness of breath, positive BD response, post-BD FEV₁/FVC less than 0.70, weight greater than 136 kg, body mass index (BMI) greater than 35 kg/m², serious medical conditions, and current pregnancy. We also excluded those whose Hispanic background was not reported.

Statistical Analyses

Survey data analysis methods accounted for cluster sampling of households within census block groups and participants within households, weighting due to oversampling of older ages, and post-stratification adjustment for nonresponse (12). Estimates are applicable to the target population of healthy Hispanic/Latino adults living in the four communities.

A survey data linear regression model was fit for each parameter (FEV₁, FVC, and FEV₁/FVC) and each sex. Age, height, age², height², Hispanic/Latino background, and interactions of linear and quadratic functions of age and height with background were considered as predictors. Optimal models were selected on the basis of type III sums of squares and R² values (see online supplement). Reference equations were computed using the final models. The LLN was calculated as the predicted value minus $1.64 \times (\text{MSE}^{1/2})$, where 1.64 is the 95th percentile of the standard normal distribution and MSE is the model mean squared error (4, 7). The 95% confidence interval (CI) was calculated as $\text{LLN} \pm 1.96 \times \text{Var}(\text{LLN})$, where Var (LLN) is the variance of the predicted value plus $(1.64^2 \times \text{MSE}/2n)$, accounting for the variance of both the regression coefficients and the MSE.

Reference equations for each sex and background were compared graphically with the NHANES III equations for Mexican Americans and the GLI white equations. Backgrounds were compared for the predicted value at the average age and

height relative to $P < 0.002$, based on a Bonferroni adjustment for 21 pairwise comparisons among the seven backgrounds. Additional models compared consistency of the equations between field centers for background subpopulations with at least 200 participants within two or more field centers.

Results

The HCHS/SOL investigators recruited 16,415 participants, among whom 38% were of Mexican heritage, 16% were of Puerto Rican heritage, 20% were of Cuban heritage, 7% were of Central American heritage, 10% were of Dominican heritage, 5% were of South American heritage, and 4% were of other or mixed heritage. Pre-BD spirometry was obtained for 94% of the cohort and post-BD spirometry for 76% of those with obstructive lung function based on the pre-BD results ($n = 1,073$). Nine percent of participants had low-quality or missing spirometry and were therefore excluded from the analysis population. Tables 1 and E1 provide all exclusions, which resulted in

Table 1. Number and Percent of Hispanic Community Health Study/Study of Latinos Participants ($n = 16,415$) Excluded (Sequentially and Overall) from the Healthy Subset, by Reason for Exclusion

Exclusions	Total Number Excluded Sequentially*	Number Included	%	Total Number Excluded Overall†	%
Before exclusions	0	16,415	100.00		
Missing or unusable spirometry‡	1,451	14,964	91.2	1,451	8.8
Hispanic background missing	31	14,933	91.0	87	0.5
Current smoker§	2,935	11,998	73.1	3,259	19.9
Current COPD or chronic bronchitis	1,638	10,360	63.1	3,438	20.9
Current asthma or wheeze	1,017	9,343	56.9	3,039	18.5
Weight >136 kg or BMI >35 kg/m ²	1,397	7,946	48.4	2,889	17.6
Ever smoker with > 10 pack-years§	672	7,274	44.3	2,987	18.2
Shortness of breath§	518	6,756	41.2	2,646	16.1
Serious medical condition§	331	6,425	39.1	1,499	9.1
Age >74 yr	0	6,425	39.1	9	0.1
Total	9,990	6,425	39.1	9,990	60.9

Definition of abbreviations: BMI = body mass index; COPD = chronic obstructive pulmonary disease. *Exclusion criteria were implemented in the order listed; thus, participants were excluded only once despite possibly meeting multiple exclusion criteria.

†The total number of participants meeting a particular exclusion criterion regardless of order (i.e., multiple exclusions per participant are accounted for).

‡Spirometry quality grades D, F, or missing (prebronchodilator).

§Self-report.

||Self-report or spirometry.

Table 2. Estimated Population Characteristics from the Hispanic Community Health Study/Study of Latinos Healthy Subset (n = 6,425)

Variables	Women (n = 3,945)		Men (n = 2,480)	
	Mean	95% CI	Mean	95% CI
FEV ₁ , L	2.72	(2.70–2.74)	3.79	(3.76–3.82)
FVC, L	3.27	(3.25–3.29)	4.61	(4.58–4.65)
FEV ₁ /FVC, %	82.97	(82.76–83.17)	82.29	(82.05–82.54)
Age, yr	39.50	(38.87–40.12)	37.09	(36.54–37.64)
Height, cm	156.88	(156.62–157.15)	170.45	(170.07–170.83)
Weight, kg	67.19	(66.74–67.64)	80.17	(79.55–80.78)
BMI, kg/m ²	27.28	(27.11–27.45)	27.55	(27.37–27.73)

Definition of abbreviations: BMI = body mass index; CI = confidence interval.

Note: Estimates of population characteristics for the target population of healthy Hispanics/Latinos aged 18–74 years residing in the Hispanic Community Health Study/Study of Latinos geographic areas were computed via survey data analysis methods.

a healthy subset of 6,425 participants for reference equation estimation. Exclusions by field center and by asthma and COPD data source are provided in Tables E2A–E2D, and spirometry grades (A–D) for FEV₁ and FVC are provided in Tables E3A and E3B. Of note is a somewhat higher percentage excluded for missing or low-quality spirometry in the Bronx field center (16%) than in the other centers (6 to 8%). Initial problems were identified by the spirometry reading center's data review, and retraining was provided to improve the quality of spirometry data collection across all centers. Most participants with missing or low-grade spirometry also had another reason for exclusion; overall, only 1% of participants (n = 171) were excluded from the healthy subset due solely to the absence of usable spirometry, with the Bronx field center having the highest rate of 1.7%.

Population estimates of demographic characteristics for the healthy subset are described in Table 2. The average ages were 40 years for women and 37 years for men. Average BMI values for women and men were 27.3 kg/m² and 27.6 kg/m², respectively. The average age in this healthy subset was somewhat younger than in the full population, and the spirometry results, weight, and BMI reflected better health

status. For example, mean FEV₁ values were 2.72 (SE, 0.01) for women and 3.79 (SE, 0.02) for men in the healthy subset, compared with 2.58 (SE, 0.01) and 3.59 (SE, 0.02), respectively, for the full population (Table E4). Population estimates by sex and Hispanic/Latino background are provided in Tables E5A and E5B.

Model selection procedures resulted in the same model structure for FEV₁, FVC, and FEV₁/FVC (Table 3). Separate intercepts were fit for each Hispanic/Latino background, resulting in parallel curves for the different background groups with respect to age and height. Interactions between background and age or height did not substantially improve model fit and were nominally nonsignificant ($P > 0.05$). Linear effects for height provided adequate fit. A quadratic term for height was identified only for FVC among women ($P = 0.048$), but the increase in R^2 was minimal. A quadratic effect for age was needed to capture the variation across the 18–74 years age range ($P < 0.05$ in each model, except FEV₁ among men). For consistency, age² was maintained in all models, and height² was excluded. The final model explained 58% of the variation in FEV₁ for women ($R^2 = 0.58$) and 54% for men; for FVC, it explained 50% for women

and 48% for men; and for FEV₁/FVC, it explained 31% for women and 20% for men (Table 3).

The final model parameterization has a simpler structure than the NHANES III model parameterizations (4), which included quadratic terms in height for some of the models. Applying the NHANES III model parameterizations to the HCHS/SOL healthy subset yielded essentially the same R^2 values as our chosen models for FEV₁ and FVC, but a lower R^2 value (indicating a poorer fit) for FEV₁/FVC ($R^2 = 0.29$ and 0.18 for women and men, respectively).

The estimated regression coefficients for each of the six models are provided in Table 4. Predicted values and LLNs can be computed from these coefficients according to the following formulas:

$$\text{Predicted Value} = \beta_{0i} \times I(\text{Background} = i) + \beta_1(\text{Age} - 45) + \beta_2(\text{Age} - 45)^2 + \beta_3(\text{Height} - 160) \quad (1)$$

$$\text{LLN} = \beta_{0i} \times I(\text{Background} = i) + \beta_1(\text{Age} - 45) + \beta_2(\text{Age} - 45)^2 + \beta_3(\text{Height} - 160) - (1.64\sqrt{\text{MSE}}) \quad (2)$$

Table 5 provides the estimated LLNs and corresponding 95% CIs for all three parameters for age 45 years and average height. Estimated reference equations and lower bounds of 95% CIs (the LLNs) are depicted graphically for FEV₁ by sex and Hispanic/Latino background in Figures 1–4. Figures 1 and 2 show predicted FEV₁ versus age at average height, and Figures 3 and 4 show predicted FEV₁ versus height at age 45 years, for each background group and for men and women separately. The estimated reference equations for FVC and

Table 3. Model Parameters and R^2 Values for Each Spirometry Parameter and Sex

Model		R^2 Women	R^2 Men
FEV ₁ , L	$\beta_{0i} \times I(\text{Ethnic} = i) + \beta_1(\text{Age} - 45) + \beta_2(\text{Age} - 45)^2 + \beta_3(\text{Height} - 160)$	0.58	0.54
FVC, L	$\beta_{0i} \times I(\text{Ethnic} = i) + \beta_1(\text{Age} - 45) + \beta_2(\text{Age} - 45)^2 + \beta_3(\text{Height} - 160)$	0.50	0.48
FEV ₁ /FVC, %	$\beta_{0i} \times I(\text{Ethnic} = i) + \beta_1(\text{Age} - 45) + \beta_2(\text{Age} - 45)^2 + \beta_3(\text{Height} - 160)$	0.31	0.20

Note: In all models, age was centered at 45 years, and height was centered at 160 cm.

Table 4. Model Coefficients for Each Spirometry Parameter, Stratified by Sex

	Intercept β_0	(Age – 45) β_1	(Age – 45) ² β_2	(Height – 160) (cm) β_3	Model $\sqrt{\text{MSE}}$
FEV ₁ , L					
Women					
Central American	2.730	−0.0214	−0.0001	0.0289	0.3290
Cuban	2.701				
Dominican	2.586				
Mexican	2.763				
Puerto Rican	2.600				
South American	2.781				
Other or mixed	2.695				
Men					
Central American	3.160	−0.0236	−0.00001	0.0422	0.4440
Cuban	3.169				
Dominican	2.957				
Mexican	3.245				
Puerto Rican	3.006				
South American	3.277				
Other or mixed	3.202				
FVC, L					
Women					
Central American	3.360	−0.0190	−0.0003	0.0382	0.3953
Cuban	3.326				
Dominican	3.137				
Mexican	3.428				
Puerto Rican	3.212				
South American	3.450				
Other or mixed	3.298				
Men					
Central American	3.912	−0.0210	−0.0002	0.0552	0.5438
Cuban	3.889				
Dominican	3.585				
Mexican	4.042				
Puerto Rican	3.713				
South American	4.093				
Other or mixed	3.865				
FEV ₁ /FVC ratio, %					
Women					
Central American	81.344	−0.1858	0.0024	−0.0868	4.3781
Cuban	81.252				
Dominican	82.560				
Mexican	80.637				
Puerto Rican	81.092				
South American	80.680				
Other or mixed	81.963				
Men					
Central American	81.203	−0.1476	0.0025	−0.0714	4.9174
Cuban	81.697				
Dominican	82.637				
Mexican	80.620				
Puerto Rican	81.386				
South American	80.417				
Other or mixed	82.841				

Definition of abbreviation: MSE = mean squared error.

Background-specific coefficients β_i were added to the model intercept β_0 to obtain the intercept for each background.

The intercept is the value at 45 years of age and 160 cm height. The model standard error is the square root of the mean square error. Model coefficients are applied to obtain the predicted value (PV) and lower limit of normal (LLN). For example, FEV₁/FVC ratio (%) for a 65-year-old woman of Central American descent and height of 156 cm is PV = 81.3444 – (0.1858)(20) + (0.0024)(20²) – 0.0868(156 – 160) = 78.9, LLN = 78.9411 – (1.64)(4.3781) = 71.8.

FEV₁/FVC are similarly depicted in Figures E1A–E1D and E2A–E2D. The NHANES III reference equations and confidence bounds for Mexican Americans aged 20 years and older (4), as well as the GLI reference

equations for white individuals (7), are superimposed on each graph for comparison purposes. Both FEV₁ and FVC declined with age for male and female Hispanic/Latinos. There is a curvilinear

shape of the equations with respect to age, with little mean change from ages 18 to 30 years, followed by a linear decline. Healthy men and women with Dominican and Puerto Rican backgrounds had numerically

Table 5. Lower Limits of Normal and 95% Confidence Intervals for Spirometry Parameters for Men and Women Aged 45 Years and of Average Height

	Women		Men	
	LLN (95% CI)	Ratio Relative to Mexicans	LLN (95% CI)	Ratio Relative to Mexicans
FEV ₁ , L				
Central American	2.08 (2.03–2.12)	0.98	2.85 (2.78–2.93)	0.97
Cuban	2.05 (2.01–2.08)	0.97	2.86 (2.80–2.92)	0.97
Dominican	1.93 (1.87–1.99)	0.92	2.65 (2.58–2.72)	0.90
Mexican	2.11 (2.08–2.14)	—	2.94 (2.88–2.99)	—
Puerto Rican	1.95 (1.86–2.03)	0.92	2.70 (2.62–2.78)	0.92
South American	2.13 (2.07–2.18)	1.01	2.97 (2.89–3.05)	1.01
Other or mixed	2.04 (1.91–2.18)	0.97	2.89 (2.78–3.01)	0.99
FVC, L				
Central American	2.56 (2.51–2.61)	0.97	3.57 (3.48–3.66)	0.96
Cuban	2.53 (2.48–2.57)	0.96	3.55 (3.48–3.62)	0.96
Dominican	2.34 (2.28–2.40)	0.89	3.24 (3.16–3.33)	0.88
Mexican	2.63 (2.59–2.66)	—	3.70 (3.64–3.76)	—
Puerto Rican	2.41 (2.31–2.51)	0.92	3.37 (3.28–3.47)	0.91
South American	2.65 (2.58–2.72)	1.01	3.75 (3.65–3.85)	1.01
Other or mixed	2.50 (2.33–2.67)	0.95	3.52 (3.37–3.68)	0.95
FEV ₁ /FVC ratio, %				
Central American	74.51 (74.00–75.02)	1.01	72.42 (71.61–73.24)	1.01
Cuban	74.42 (73.92–74.92)	1.01	72.92 (72.34–73.50)	1.01
Dominican	75.73 (75.05–76.41)	1.03	73.86 (73.00–74.72)	1.03
Mexican	73.80 (73.45–74.15)	—	71.84 (71.25–72.43)	—
Puerto Rican	74.26 (73.47–75.05)	1.01	72.61 (71.70–73.52)	1.01
South American	73.85 (73.20–74.50)	1.00	71.64 (70.69–72.59)	1.00
Other or mixed	75.13 (73.79–76.47)	1.02	74.06 (72.86–75.27)	1.03

Definition of abbreviations: CI = confidence interval; LLN = lower limit of normal; MSE = mean squared error.

LLN was constructed using age = 45, height = 170 cm for men and 156 cm for women.

LLN = $(\beta_0 + \beta_1 \text{Age} + \beta_2 \text{Age}^2 + \beta_3 \text{Height}) - (1.64\sqrt{\text{MSE}})$. 95% CI = $\text{LLN} \pm 1.96\sqrt{\text{Var}(\text{LLN})}$, where

$\text{Var}(\text{LLN}) = \text{Var}(\beta_0 + \beta_1 \text{Age} + \beta_2 \text{Age}^2 + \beta_3 \text{Height}) + (1.64^2 \text{MSE}/2n)$, derived from the variance/covariance matrix of the regression coefficients and the variance of model MSE.

lower predicted values for FEV₁ and FVC than the predicted values from the NHANES III Mexican American and the GLI white reference equations, and also than other Hispanic/Latino groups from HCHS/SOL. All other background groups had numerically similar predicted values to those from both NHANES and GLI.

Predicted mean values and 95% CIs by sex and Hispanic/Latino background at age 45 years and average height are depicted for each parameter in Figures 5–7. Predicted values for the Dominican and Puerto Rican groups were statistically significantly lower than for many other Hispanic/Latino groups ($P < 0.002$ with adjustment for multiple comparisons) and demonstrated a trend ($P < 0.05$) compared with all other backgrounds individually (except the other/mixed group for women). Predicted values from the NHANES and GLI FEV₁ and FVC equations are within the 95% CIs for Mexican and South American backgrounds, but they are larger than the upper 95% CIs for Dominicans, Puerto Ricans, and Cubans.

Figures E3A–E3C show the predicted values at age 45 years and average height for background groups within each field center with a sample size of at least 200. Although only one field center (Bronx, NY) had a sufficient number of individuals of Dominican background for examination, Puerto Rican background was compared between the Chicago and Bronx centers. Average FEV₁ for men was found to be lower (worse) in Chicago than in the Bronx ($P = 0.026$). This difference persisted after further adjustment for BMI, education level, years in the United States, and age at immigration. There were no other differences between sites within background groups.

FEV₁/FVC declined with age. For men, the predicted and LLN values at ages above 50 years were numerically higher than with the NHANES III and GLI reference equations (4, 7). The curvilinear shape of the equation with respect to age for both sexes appears to be driven by a few outliers in the oldest age group with short FVC maneuvers. Dominican background

had numerically higher predicted FEV₁/FVC than all other Hispanic/Latino backgrounds other than the mixed or other category, with statistically significant differences ($P < 0.002$) versus some backgrounds.

Discussion

The study design and sample selection of the HCHS/SOL, the largest study conducted to date of U.S. Hispanics/Latinos, provide a unique opportunity to estimate reference equations for spirometry parameters based on healthy Hispanics with diverse backgrounds. Reference equations currently in use, such as the NHANES III or GLI equations, rely predominantly on data from Mexican Americans with suggested extrapolation to other Hispanic groups. The differences we observed among Hispanic/Latino background groups in predicted FEV₁, FVC, and FEV₁/FVC are clinically important and suggest that use of the HCHS/SOL reference equations may be

Figure 1. FEV₁ predicted (pred) value and lower limit of normal (LLN) versus age for women comparing reference equations from the HCHS/SOL (Hispanic Community Health Study/Study of Latinos) with the NHANES III (Third National Health and Nutrition Examination Study) for Mexican Americans and the Global Lung Initiative (GLI) for white individuals.

more appropriate, especially for adult patients with Dominican and Puerto Rican backgrounds who are seen in clinical practice with pulmonary symptoms. The use of these equations should provide a more accurate assessment of lung function abnormality and physiologic impairment and avoid the potential for overdiagnosis of restriction based on percent predicted FVC and, potentially, misclassification of COPD and asthma based on the LLN for the FEV₁/FVC ratio.

The LLN of the FEV₁/FVC ratio is clinically important because it is used to define airway obstruction for chronic lung diseases that commonly occur in adults, such as asthma and COPD. The lower LLNs from the GLI study are probably due to the inclusion of relatively large numbers of men and women over age 65 years, who exhibit larger variability in lung function.

The HCHS/SOL reference equations are based on model parameterizations that are straightforward to apply in a clinical setting and involve only sex, age, height, and Hispanic/Latino background obtained from self-identification as model covariates. There are many other factors that have been found to explain variability in lung function

among subpopulations (e.g., sitting height, waist circumference) that were not examined in the present study. Our analysis objective for HCHS/SOL was not to determine the best set of model predictors from all available measurements made on the study cohort, but rather to determine the best set among those measurements readily available during a routine clinical examination and traditionally used to determine normative values for patients seen in clinical practice.

Ethnic or racial differences in lung function have long been recognized, with studies confirming that for a given sex, age, and height, healthy African Americans and Mexican Americans have 12 to 15% and 5 to 10%, respectively, lower FEV₁ and FVC (but similar ratios) compared with white individuals (4). Consistent with the HCHS/SOL results, Hispanics of non-Mexican background in the Multi-Ethnic Study of Atherosclerosis Lung Study, most of whom were of Puerto Rican and Dominican background, had lower FEV₁ and FVC than Mexican Americans (8). The 2005 ATS and European Respiratory Society spirometry guidelines recommended the use of race- and

ethnicity-specific reference equations developed from NHANES III throughout North America (5). A large worldwide spirometry study (PURE [Prospective Urban Rural Epidemiology Study]) recently confirmed racial and ethnic differences of more than 25% in FEV₁ and FVC (19). An analysis of adults self-identifying as African American from five large cohort studies found that differences in FVC were directly proportional to the percentage of African ancestry determined by DNA testing (20), suggesting that differences in FVC may be due to the substantially higher proportion of African ancestry in HCHS/SOL Dominican background and, to a lesser degree, Puerto Rican background (9). We investigated this in additional analyses within the subset of the healthy sample with genetic data ($n = 4,037$ [63%]). After adjustment for percent African genetic ancestry, there were no differences between backgrounds for any parameter for women. Differences remained, although they were somewhat attenuated, in Puerto Rican men for FEV₁ (vs. all backgrounds except Central American) and FVC (vs. Mexican and South American). The fact that Puerto

Figure 2. FEV₁ predicted (pred) value and lower limit of normal (LLN) versus age for men comparing reference equations from the HCHS/SOL (Hispanic Community Health Study/Study of Latinos) with NHANES III (Third National Health and Nutrition Examination Study) for Mexican Americans and the Global Lung Initiative (GLI) for white individuals.

Figure 3. FEV₁ predicted (pred) value and lower limit of normal (LLN) versus height for women comparing reference equations from the HCHS/SOL (Hispanic Community Health Study/Study of Latinos) with NHANES III (Third National Health and Nutrition Examination Study) for Mexican Americans and the Global Lung Initiative (GLI) for white individuals.

Figure 4. FEV₁ predicted (pred) value and lower limit of normal (LLN) versus height for men comparing reference equations from the HCHS/SOL (Hispanic Community Health Study/Study of Latinos) with NHANES III (Third National Health and Nutrition Examination Study) for Mexican Americans and the Global Lung Initiative (GLI) for white individuals.

Rican differences were not fully explained by African genetic ancestry, whereas those between Dominicans and other groups were, may be reflective of the ancestral patterns; for example, average percent African genetic ancestry among Puerto Ricans was approximately half that of Dominicans (9).

One important difference in the methodology used to estimate the

HCHS/SOL and NHANES III reference equations is the use of sampling weights, which impact the point estimates, and other sample design parameters, such as clustering and stratification, which impact the estimates of variance. By applying survey regression methods to fit the prediction models, we are able to draw inference to the target population of all healthy Hispanics/Latinos residing in the four

community areas targeted for recruitment and not just to the study sample of participants. The resulting R^2 values, which indicate a smaller percentage of variation explained by our equations than that reported for the NHANES III reference equations for Mexican Americans, reflect this ability to make broader inference to the target population. For example, the R^2 values for healthy Mexican Americans in the NHANES III equations for FEV₁, FVC, and FEV₁/FVC were 73%, 71%, and 34% for women and 85%, 86%, and 27% for men (4), respectively, all greater than the corresponding R^2 values for the HCHS/SOL equations. Because the same spirometry protocol was used for both studies, the smaller R^2 values are likely due both to a more heterogeneous population being studied in HCHS/SOL and to HCHS/SOL's total variation estimate representing the target population (with sampling weights and design-based estimates) and not just the study sample, as in NHANES. A limitation is that results from the four urban communities might not be generalizable to the U.S. population.

Figure 5. Population predicted values and 95% confidence intervals of FEV₁ (liters) for men and women aged 45 years and of average height, by Hispanic background.

Figure 6. Population predicted values and 95% confidence intervals of FVC (liters) for men and women aged 45 years and of average height, by Hispanic background.

The problems inherent in applying spirometry reference values derived from healthy white individuals to all ethnic and racial groups are well recognized (21–23). Because nonwhite race/ethnicity is often associated with lower SES, relatively poor childhood nutrition, and exposure to higher levels of indoor and outdoor air pollution during childhood (24), some epidemiologists suggest that using race- or ethnicity-specific reference equations draws attention from these suboptimal conditions, making them seem normal (25). Indeed, mean FVC for a given sex, age, and height is higher for third-generation Japanese Americans than for first-generation Japanese Americans (26). Changes in SES over time within Japan during the 20th century (especially after World War II) are associated with changes in the sitting to standing height ratio, which explain differences in predicted FVC from one Japanese birth cohort to another (27).

These types of changes over time support the need for periodic updates to normalizing equations to ensure their usefulness in clinical practice.

On one hand, improvement in SES when Hispanic/Latino adults move from their native countries to the United States has probably increased lung volumes among their children growing up in the United States. On the other hand, differences in body build, exemplified by the lower average trunk/leg ratio among African Americans, and probably certain Hispanics/Latinos than non-Hispanic white individuals, likely explains at least some of the differences in percent predicted spirometric indices. Such differences should not be ignored in the evaluation of a specific patient, in the interest of counterbalancing hypothesized low SES. Furthermore, aside from leading to unnecessary treatment and treatment side effects, overdiagnosis of lung disease has the potential for needless

disqualification of Hispanics from certain occupations—notably firefighting. We are unaware of a study that has attempted to separate differences in percentage African ancestry from differences in childhood SES in their effect on the growth of lung volumes during childhood and adolescence (which is the source of the differences seen in adulthood).

Although an argument can be made against using race- and ethnicity-specific spirometry reference equations for epidemiologic purposes and for estimating the prevalence of lung disease in large population groups, we believe that they can be useful when spirometric results are used to assist clinicians who are detecting lung diseases in adults with respiratory symptoms. They are also useful when estimating the severity of physiologic impairment or disease control in patients who have been diagnosed with lung disease. For example, an adult cigarette smoker with dyspnea and a low post-BD FEV₁/FVC is often diagnosed as having COPD (28). In clinical practice, treatment decisions are often based in part on the percent predicted FEV₁. If such a patient is of Dominican or Puerto Rican background and the NHANES III reference equation for Mexican Americans is used, the percent predicted FEV₁ may be underestimated by as much as 10%, which could impact the course of treatment recommended.

We have presented reference equations and LLN values for spirometric parameters from a large, population-based study of U.S. Hispanics/Latinos. The findings suggest that applying predicted values and the LLN derived among Mexican Americans to Puerto Ricans and Dominicans results in substantial misclassification, which is likely to be clinically significant for some patients. Use of these background-specific equations in clinical practice has the potential to provide benefit through the use of more accurate normalizing values when diagnosing and treating lung disease compared with those based on equations for Mexican Americans alone. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors thank the staff and participants of HCHS/SOL for their important contributions. A complete list of staff and investigators was published previously (12) and is also available on the study website (www.csc.unc.edu/hchs/).

Figure 7. Population predicted values and 95% confidence intervals of FEV₁/FVC ratio (%) for men and women aged 45 years and average height, by Hispanic background.

References

- Chinn S, Gislason T, Aspelund T, Gudnason V. Optimum expression of adult lung function based on all-cause mortality: results from the Reykjavik study. *Respir Med* 2007;101:601–609.
- Mannino DM, Buist AS, Petty TL, Enright PL, Redd SC. Lung function and mortality in the United States: data from the First National Health and Nutrition Examination Survey follow up study. *Thorax* 2003;58:388–393.
- Schünemann HJ, Dorn J, Grant BJB, Winkelstein W Jr, Trevisan M. Pulmonary function is a long-term predictor of mortality in the general population: 29-year follow-up of the Buffalo Health Study. *Chest* 2000;118:656–664.
- 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;159:179–187.
- Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, Coates A, van der Grinten CP, Gustafsson P, Hankinson J, *et al.* Interpretative strategies for lung function tests. *Eur Respir J* 2005;26:948–968.
- Pérez-Padilla R, Valdivia G, Muñio A, López MV, Márquez MN, Montes de Oca M, Tálaro C, Lisboa C, Pertuzé J, B Jardim JR, *et al.* Spirometric reference values in 5 large Latin American cities for subjects aged 40 years or over [in Spanish]. *Arch Bronconeumol* 2006;42:317–325.
- Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, Enright PL, Hankinson JL, Ip MS, Zheng J, *et al.*; ERS Global Lung Function Initiative. Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012;40:1324–1343.
- Hankinson JL, Kawut SM, Shahar E, Smith LJ, Stukovsky KH, Barr RG. 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;137:138–145.
- Conomos MP, Laurie CA, Stilp AM, Gogarten SM, McHugh CP, Nelson SC, Sofer T, Fernández-Rhodes L, Justice AE, Graff M, *et al.* Genetic diversity and association studies in US Hispanic/Latino populations: applications in the Hispanic Community Health Study/Study of Latinos. *Am J Hum Genet* 2016;98:165–184.
- Enright PL, Kronmal RA, Higgins M, Schenker M, Haponik EF. Spirometry reference values for women and men 65 to 85 years of age: cardiovascular health study. *Am Rev Respir Dis* 1993;147:125–133.
- Sorlie PD, Avilés-Santa LM, Wassertheil-Smoller S, Kaplan RC, Daviglus ML, Giachello AL, Schneiderman N, Raji L, Talavera G, Allison M, *et al.* Design and implementation of the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol* 2010;20:629–641.
- Lavange LM, Kalsbeek WD, Sorlie PD, Avilés-Santa LM, Kaplan RC, Barnhart J, Liu K, Giachello A, Lee DJ, Ryan J, *et al.* Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol* 2010;20:642–649.
- Daviglus ML, Talavera GA, Avilés-Santa LM, Allison M, Cai J, Criqui MH, Gellman M, Giachello AL, Gouskova N, Kaplan RC, *et al.* Prevalence of major cardiovascular risk factors and cardiovascular diseases among Hispanic/Latino individuals of diverse backgrounds in the United States. *JAMA* 2012;308:1775–1784.
- Ferris BG. Epidemiology Standardization Project (American Thoracic Society). *Am Rev Respir Dis* 1978;118:1–120.
- Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P, van der Grinten CP, Gustafsson P, *et al.*; ATS/ERS Task Force. Standardisation of spirometry. *Eur Respir J* 2005;26:319–338.
- Barr RG, Avilés-Santa L, Davis SM, Aldrich TK, Gonzalez F II, Henderson AG, Kaplan RC, LaVange L, Liu K, Loredó JS, *et al.* Pulmonary disease and age at immigration among Hispanics: results from the Hispanic Community Health Study/Study of Latinos. *Am J Respir Crit Care Med* 2016;193:386–395.
- Hankinson JL, Eschenbacher B, Townsend M, Stocks J, Quanjer PH. Use of forced vital capacity and forced expiratory volume in 1 second quality criteria for determining a valid test. *Eur Respir J* 2015;45:1283–1292.
- Enright PL, Kaminsky DA. Strategies for screening for chronic obstructive pulmonary disease. *Respir Care* 2003;48:1194–1203.
- Duong M, Islam S, Rangarajan S, Teo K, O'Byrne PM, Schünemann HJ, Igumbor E, Chifamba J, Liu L, Li W, *et al.*; PURE-BREATH Study Investigators. Global differences in lung function by region (PURE): an international, community-based prospective study. *Lancet Respir Med* 2013;1:599–609.
- Kumar R, Seibold MA, Aldrich MC, Williams LK, Reiner AP, Colangelo L, Galanter J, Gignoux C, Hu D, Sen S, *et al.* Genetic ancestry in lung-function predictions. *N Engl J Med* 2010;363:321–330.
- Quanjer PH. Lung function, race and ethnicity: a conundrum. *Eur Respir J* 2013;41:1249–1251.
- Quanjer PH. Lung function, genetics and socioeconomic conditions. *Eur Respir J* 2015;45:1529–1533.
- Stanojevic S, Cole TJ, Stocks J, Quanjer PH. Disentangling the discordance between epidemiological associations and physiological mechanisms. *Thorax* 2014;69:869.
- Lawlor DA, Ebrahim S, Davey Smith G. Association between self-reported childhood socioeconomic position and adult lung function: findings from the British Women's Heart and Health Study. *Thorax* 2004;59:199–203.
- Burney PGJ, Hooper RL. The use of ethnically specific norms for ventilatory function in African-American and white populations. *Int J Epidemiol* 2012;41:782–790.
- Fulambarker A, Copur AS, Cohen ME, Patel M, Gill S, Schultz ST, Quanjer PH. Comparison of pulmonary function in immigrant vs US-born Asian Indians. *Chest* 2010;137:1398–1404.
- Quanjer PH, Kubota M, Kobayashi H, Omori H, Tatsumi K, Kanazawa M, Stanojevic S, Stocks J, Cole TJ. Secular changes in relative leg length confound height-based spirometric reference values. *Chest* 2015;147:792–797.
- Vestbo J, Hurd SS, Agustí AG, Jones PW, Vogelmeier C, Anzueto A, Barnes PJ, Fabbri LM, Martinez FJ, Nishimura M, *et al.* Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2013;187:347–365.