

1 **Word counts**

2 Text: 2,831

3 Abstract: 275

4 **Title:** Racial/ethnic-based spirometry reference equations: Are they accurate for admixed
5 populations?

6 **Running head:** Racial/Ethnic-Based Spirometry: Accuracy for Admixed

7

8 **Author list:** Jonathan Witonsky, MD; Jennifer R. Elhawary, MS; Celeste Eng, BS; José R.

9 Rodríguez-Santana, MD; Luisa N. Borrell[#], DDS, PhD; and Esteban G. Burchard[#], MD, MPH

10 [#]*Shared senior authors.*

11 **Affiliations:** From the Department of Pediatrics (Dr Witonsky), Department of Medicine (Ms

12 Elhawary and Ms Eng), and Department of Medicine and Department of Bioengineering and

13 Therapeutic Sciences (Dr Burchard), University of California, San Francisco, San Francisco,

14 CA; Centro de Neumología Pediátrica (Dr Rodríguez-Santana), San Juan, PR; and the

15 Department of Epidemiology and Biostatistics, Graduate School of Public Health and Health

16 Policy (Dr Borrell), City University of New York, New York, NY.

17 **Correspondence to:** Jonathan Witonsky, MD, Department of Pediatrics, University of

18 California, San Francisco, 550 16th Street, 4th Floor, San Francisco, CA, 94143; e-mail:

19 jonathan.witonsky@ucsf.edu

20 **Funding information:** This work was supported in part by the Sandler Family Foundation; the

21 American Asthma Foundation; the RWJF Amos Medical Faculty Development Program; Harry

22 Wm. and Diana V. Hind Distinguished Professor in Pharmaceutical Sciences II; the National

23 Institutes of Health, National Heart, Lung, and Blood Institute [R01HL117004, 1X01HL134589,

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

24 R01HL128439, R01HL135156, R01HL141992]; the National Institute of Environmental Health
25 Sciences [R01ES015794, R21ES24844]; the National Institute on Minority Health and Health
26 Disparities [R01MD010443, R56MD013312]; the National Human Genome Research Institute
27 [U01HG009080]; the National Institute of General Medical Sciences [T32GM007546], and the
28 UCSF Pediatric Allergy, Immunology, and Bone Marrow Transplantation Division. The content
29 of this publication is solely the responsibility of the authors and does not necessarily reflect the
30 views or policies of the Department of Health and Human Services, nor does mention of trade
31 names, commercial products, or organizations imply endorsement by the U.S. government.

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

32 **ABSTRACT**

33 **BACKGROUND:** Variation in genetic ancestry among admixed racial/ethnic groups may
34 influence the fit of guideline-recommended spirometry reference equations, which rely on self-
35 identified race/ethnicity.

36 **RESEARCH QUESTION:** What is the influence of genetic ancestry on the fit of the guideline-
37 recommended racial/ethnic-based spirometry reference equations in populations of genetically
38 admixed children?

39 **STUDY DESIGN AND METHODS:** Cross-sectional fit of guideline-recommended
40 racial/ethnic-based spirometry reference equations was evaluated in control subjects from case-
41 control studies of asthma. Anthropometry, blood samples, and spirometric measurements were
42 obtained in 599 healthy admixed children, aged 8 to 21-years. Genetic ancestry was estimated
43 using genome-wide genotype data. Equation fit was determined as a mean z-score between -0.5
44 and 0.5 and assessed in self-identified African American (N = 275) and Puerto Rican (N = 324)
45 children using the distribution to determine cut points of genetic ancestry.

46 **RESULTS:** For African American children, African American-derived equations fit for
47 predicting FEV₁ and FVC in those with an African ancestry above the median (81-100%),
48 whereas composite equations for “other/mixed” populations fit for predicting FEV₁ and FVC in
49 those with an African ancestry below the median (31-81%). Among Puerto Rican children,
50 White-derived equations fit for predicting FEV₁, and the composite equations fit for predicting
51 FVC for those with African ancestry above the median (21-88%). In contrast, in Puerto Rican
52 children with African ancestry below the median (6-21%), only equations derived in Whites
53 provide an adequate fit.

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

- 54 **INTERPRETATION:** Guideline-recommended spirometry reference equations yielded biased
55 estimates of lung function in admixed populations with high variation of African ancestry.
56 Spirometry is due for reference equations that incorporate genetic ancestry, either for more
57 precise application of the current equations or the derivation and utilization of new equations.

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

58 INTRODUCTION

59 There has been great debate over the use of racial/ethnic classification in medicine and
60 biomedical research.¹⁻³ Presently, pulmonary function testing is one of the few clinical
61 applications where self-reported race/ethnicity is used to define a “normal” range. Normative
62 equations of lung function have been developed by testing large populations categorized by self-
63 reported race/ethnicity.^{4,5} However, many populations have ancestral contributions from multiple
64 continents (i.e., genetically admixed) and, as a result, self-identified racial and ethnic categories
65 may not fully capture an individual’s genetic ancestry.⁶⁻⁹ Therefore, using self-reported
66 race/ethnicity may result in misclassifying individuals with respect to the normal range for
67 physiologic lung measures if the latter are more dependent on ancestry.¹⁰ This error could lead to
68 inaccuracies in evaluating individual pulmonary function.

69 The two most widely employed sets of spirometry reference equations used in pulmonary
70 function testing, and those recommended by American Thoracic Society and European
71 Respiratory Society (ATS/ERS) guidelines, are the third National Health and Nutrition
72 Examination Survey (NHANES III) equations, derived in White, African American, and
73 Mexican American adults and children aged 8-80 years, and the Global Lung Function Initiative
74 (GLI) equations, derived in White, African American, and North and South East Asian adults
75 and children aged 3-95 years.^{4,5,11} GLI also offers a composite equation—an average of the four
76 population-specific GLI equations—for use in multiracial or unrepresented populations. These
77 guideline-recommended spirometry reference equations adjust for clinically important
78 differences in normal lung function observed between racial/ethnic groups. Selecting an
79 inappropriate reference equation—or ignoring race/ethnicity by using a one-size-fits-all
80 approach—can cause unintended clinical consequences, including errors and delays in disease

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

81 detection and medical management, misclassification of disease severity, denial of disability
82 claims, and ineligibility for life-saving treatments such as transplants and other surgeries.^{12–20}
83 Several studies have demonstrated that the guideline-recommended lung function equations are
84 well-fitted to populations similar to those in which they were derived.^{16,21–25} However, to date,
85 none of these studies have evaluated equation fit after considering intra-population variation in
86 genetic ancestry, which we have demonstrated is associated with lung function in African
87 Americans and Latinos.^{26–28} Herein, we evaluate the guideline-recommended White, African
88 American, Mexican American, and composite reference equations for two independent
89 genetically admixed populations of African American and Puerto Rican children. Specifically,
90 we assessed guideline-recommended lung function equation fit in each population and
91 investigated their fit after classifying these populations using their genetic ancestry distribution.

92 **METHODS**

93 **Study Populations**

94 Spirometric predictions and genetic ancestry proportions were calculated using clinical and
95 genetic data from healthy control subjects among two case-control studies of asthma conducted
96 between 2006 and 2014: the Genes-environments & Admixture in Latino Americans (GALA II)
97 study and the Study of African Americans, Asthma, Genes, & Environments (SAGE).²⁹ Subjects
98 were aged 8 to 21-years old at the time of recruitment and were healthy with no reported history
99 of lung disease, asthma or allergies, use of medication for allergies, or any symptoms of
100 wheezing or shortness of breath during their lifetime. Parents and grandparents of study subjects
101 self-identified as Puerto Rican in GALA II or African American in SAGE. GALA II subjects
102 were recruited from five urban study centers throughout the U.S. (Chicago, Bronx, Houston, San

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

103 Francisco Bay Area, and Puerto Rico) and SAGE subjects were recruited from the San Francisco
104 Bay Area.

105 All participants provided written consent to being in the study. Consent was obtained from all
106 participants 18-years and older and parent/legal guardians of minor participants. The study
107 protocols for both GALA II and SAGE were approved by the University of California, San
108 Francisco (UCSF) Human Research Protection Program Institutional Review Board (IRB) and
109 all institutions participating in recruitment obtained the appropriate approvals from their IRBs for
110 recruitment related activity.

111 In total, 3,226 children without asthma were enrolled in the two studies (2,538 in GALA II and
112 688 in SAGE) with complete spirometry and genome-wide genetic data available for 1,049
113 children. After exclusion criteria (Figure 1) were applied, 599 children remained: 275 African
114 American children and 324 Puerto Rican children.

115 **Assessments**

116 Spirometric measurements were collected on all subjects with testing performed in accordance
117 with ATS recommendations.³⁰ Samples were genotyped with the Affymetrix Axiom® LAT1
118 array (World Array 4, Affymetrix, Santa Clara, CA), which includes 817,810 single nucleotide
119 polymorphisms (SNPs). This array was designed to capture genetic variation in African-descent
120 populations such as African Americans and Latinos.³¹ The genetic ancestry of each study
121 participant was determined using an unsupervised analysis in the ADMIXTURE software
122 package, as described elsewhere.²⁸

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

123 We examined the distribution of genetic ancestry in each population and used the median for
124 African ancestry to classify African American and Puerto Rican populations into two groups,
125 “*greater than median*” and “*less than or equal to median,*” based on African ancestry. Median
126 African ancestry cut points were 81.3% (range: 30.7-100%) for African American children and
127 21.3% (range: 6.4-87.5%) for Puerto Rican children. Reference equation fit was independently
128 assessed in the African American population (N = 275), the African American population with
129 African ancestry greater than the population’s median (N = 137), the African American
130 population with African ancestry less than or equal to the population median (N = 138), the
131 Puerto Rican population (N = 324), the Puerto Rican population with African ancestry greater
132 than the population’s median (N = 162), and the Puerto Rican population with African ancestry
133 less than or equal to the population median (N = 162).

134 **Statistical Analyses**

135 Demographic characteristics were compared within each subpopulation using the Student's t-test
136 for normally distributed continuous variables, the Wilcoxon Rank Sum Test for non-normally
137 distributed continuous variables, and chi-squared statistic for dichotomous variables.

138 Consistent with previous studies,^{32,33} cross-sectional fit of the equations was determined by
139 calculating FEV₁ and FVC z-scores. The z-score means for FEV₁ and FVC were calculated using
140 the NHANES III White, African American, and Mexican American equations and the GLI
141 White, African American, and composite equations. Because all study participants were healthy
142 with assumed normal lung function, an equation was deemed “best fit” when the mean z-score
143 was closest to zero (perfect fit being a mean z-score of zero), and an equation was deemed to
144 have “sufficient fit” if the mean z-score was between -0.5 and 0.5 using two one-sided t-tests

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

145 (TOST) for equivalence.^{21,34,35} Assessment of significance testing for TOST was performed
146 using 95% confidence intervals and *P*-values.³⁶ An equation was deemed appropriate if it
147 provided sufficient fit for both FEV₁ and FVC. A sensitivity analysis was performed assessing fit
148 using genetic ancestry tertile, quartile, and quintile distributions. To assess for misclassification,
149 the proportion of subjects with observed spirometry data below the lower limit of normal (LLN),
150 which corresponds to the 5th percentile (z-score = -1.65) of predicted values, were also evaluated
151 for FEV₁ and FVC. Data analyses were conducted using R.³³

152 **RESULTS**

153 Table 1 summarizes general characteristics of participants in each racial/ethnic group. Sex, age,
154 height, and weight distributions were comparable across each subpopulation. Lung function, as
155 measured by mean z-scores for FEV₁ and FVC, was higher for African American (zFEV₁: 0.41,
156 zFVC: 0.56) and Puerto Rican (zFEV₁: 0.69, zFVC: 0.55) subpopulations with African ancestry
157 less than or equal to the median than African American (zFEV₁: 0.29, zFVC: 0.43) and Puerto
158 Rican (zFEV₁: 0.43, zFVC: 0.23) subpopulations with African ancestry greater than the median.
159 These differences in lung function between subpopulations were statistically significant for the
160 Puerto Rican children but not for the African American children.

161 **Genetic Ancestry Proportions**

162 When comparing subpopulations according to the median distribution, the means (\pm standard
163 deviation, or SD) of African ancestry were 73.2% (\pm 9.6%) and 85.9% (\pm 3.2%) for African
164 American children below and above the median, respectively (Table 1). These values were
165 15.1% (\pm 3.3%) and 30.2% (\pm 9.8%) among Puerto Rican children. Figure 2 shows the

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

166 European, African, and Native American genetic admixture in the African American and Puerto
167 Rican populations.

168 **Fit to African American Children**

169 Table 2a provides equation fit to spirometric measurements in the study population and
170 subpopulations of African American children based on genetic ancestry distribution. The
171 NHANES III equations derived in African Americans and the GLI composite equations were
172 appropriate for African American children at the population level, providing sufficient fit for
173 predicting FEV₁ and FVC. The mean z-scores were 0.17 (95% CI: 0.06 to 0.28) for predicting
174 FEV₁ (2.5% below the LLN) and 0.34 (95% CI: 0.23 to 0.45) for predicting FVC (1.8% below
175 the LLN) using the NHANES III African American equations, whereas they were -0.32 (95% CI:
176 -0.43 to -0.20) for predicting FEV₁ (5.8% below the LLN) and -0.14 (95% CI: -0.27 to -0.02) for
177 predicting FVC (6.2% below the LLN) using GLI composite equations. When using the African
178 genetic ancestry median to classify African American children, only the GLI composite
179 equations were appropriate for predicting both FEV₁ and FVC with mean z-scores of -0.25 (95%
180 CI: -0.42 to -0.09) for FEV₁ (5.8% below the LLN) and -0.08 (95% CI: -0.25 to 0.10) for FVC
181 (5.8% below the LLN) for African American children classified below the median. In contrast,
182 for African American children with African ancestry above the median, the NHANES III
183 equations derived in African American children were appropriate for predicting FEV₁ (0.11;
184 95% CI: -0.04 to 0.26; 2.9% below the LLN) and FVC (0.27; 95% CI: 0.10 to 0.43; 2.9% below
185 the LLN). Figure 3a provides the fit for the NHANES III and GLI equations using other
186 distribution cut points of African ancestry in addition to the median. The figure underscores the
187 poor fit of the equations in African American children when using other African ancestry cut
188 points.

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

189 **Fit to Puerto Rican Children**

190 At the population level, the GLI equations derived in Whites were appropriate for Puerto Rican
191 children with mean z-scores of -0.06 (95% CI: -0.19 to 0.06) for predicting FEV₁ (6.8% below
192 the LLN) and -0.33 (95% CI: -0.46 to -0.21) for predicting FVC (11.1% below the LLN) (Table
193 2b). For Puerto Rican children with African ancestry below and above the median, the GLI and
194 NHANES III equations derived in Whites fit for predicting FEV₁. The White-derived equations
195 only fit for predicting FVC among Puerto Rican children with African ancestry below the
196 median. The GLI composite equation fit for predicting FVC in the subpopulation of Puerto Rican
197 children with African ancestry above the median. However, the GLI composite equations
198 underestimated FEV₁ predictions among those with African ancestry below the median (Figure
199 3b). In Puerto Rican children with African ancestry below the median, both GLI and NHANES
200 III equations derived in Whites were appropriate for predicting FEV₁ and FVC.

201 **DISCUSSION**

202 Pulmonary function testing is one of the few applications in clinical medicine where self-
203 identified race/ethnicity is used to predict normative values. Our findings confirm previous
204 observations that race/ethnicity-based equations can misestimate lung function after considering
205 intra-population variation in African ancestry.²⁶ The influence of genetic ancestry on guideline-
206 recommended equation fit is relevant to African American and Puerto Rican populations in
207 which there are especially wide distributions in the proportions of African ancestry.³⁷
208 Consequently, the guideline-recommended equations may exacerbate inequities in respiratory
209 health, which disproportionately affect the very populations for whom the race/ethnicity-based
210 spirometry equations are the least precise.³⁸ As a result, these populations are most prone to

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

211 underestimated lung function predictions. Misestimated lung function may result in delays of
212 disease detection or overestimated lung function predictions, which can result in unnecessary
213 treatment. This is of critical concern given that African American and Latino/Hispanic children
214 collectively make up approximately 40% of all children in the United States.³⁹

215 Previous studies have validated the use of the GLI equations derived in African American
216 populations for use in African and African American children.^{22,40} Our findings for African
217 American children in this study were consistent with those studies, although only for the total
218 population and ignoring African ancestry variation. After considering the variation in African
219 ancestry, the guideline-recommended equations derived in African American populations were
220 no longer well-fitted for predicting lung function in African American children with African
221 ancestry below the population's median. Only the GLI composite equation—an equation derived
222 as the average of all GLI data for use in multiracial or unrepresented populations—was
223 appropriately fitted for this subpopulation. This finding suggests that inappropriate spirometry
224 reference equations might be used for as many as half of all African American children, and
225 further illustrates how intra-population variation in African ancestry limits the clinical utility of a
226 single reference equation for an entire genetically admixed group.

227 In this study, the GLI equations derived in Whites were appropriate for the study population of
228 Puerto Rican children. This was an unexpected finding given that the largest study of reference
229 equation fit for Latino/Hispanic adults found the GLI White equations overestimated lung
230 function among adults who self-identified as Puerto Rican.¹⁸ The finding was also contrary to the
231 assumption that the GLI composite equation would be most appropriate for Puerto Ricans, a
232 population not represented by the four groups in which the GLI equations were derived.

233 Corroborating prior studies showing that African ancestry is inversely associated with lung

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

234 function in Puerto Rican individuals,^{27,28} the GLI White equations were not appropriate for
235 Puerto Rican children with African ancestry above the population's median. In this
236 subpopulation, no equation was appropriately fitted for predicting both FEV₁ and FVC,
237 suggesting that as many as half of all Puerto Rican children are being misclassified for these
238 measures. The lack of an appropriate equation for Puerto Rican children with more African
239 ancestry is especially problematic considering that asthma prevalence and mortality are highest
240 among Puerto Ricans compared to other racial/ethnic groups.^{41,42} Puerto Ricans have wide-
241 ranging proportions of genetic ancestry and self-identify across the spectrum of the U.S. Census
242 racial categories, making the population ideally suited to the ongoing study of race/ethnicity and
243 genetic ancestry as predictors of lung function.^{37,43,44}

244 There are some important limitations of this study. First, after classifying the populations by
245 genetic ancestry distribution, the sample sizes were below what is recommended for spirometry
246 equation validation studies.⁴⁵ Consequently, equations deemed a poor fit for a subpopulation
247 resulted because this study was underpowered to detect significance when using specified TOST
248 criterion. However, to address this limitation, a sensitivity analysis demonstrated how mean z-
249 scores tended to trend in the serially stratified subpopulations, revealing the continuous influence
250 of ancestry on equation fit. Second, the study did not include the ratio of FEV₁ to FVC
251 (FEV₁/FVC), which is the most commonly used outcome to assess airway obstruction. The
252 FEV₁/FVC ratio was excluded because prior reports suggest that racial/ethnic reductions in FEV₁
253 and FVC are proportional and, therefore, the ratio is independent of race/ethnicity.^{5,22,46,47} Third,
254 our findings are not generalizable to adults because the analysis was performed exclusively in
255 children aged 18-21. Lastly, our results were not replicated in other populations of African

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

256 American and Puerto Rican children, an essential next step in the evaluation of the guideline-
257 recommended equations.

258 Despite these limitations, this is the first study assessing the influence of genetic ancestry on the
259 applicability of guideline-recommended race/ethnicity-based spirometry reference equations for
260 admixed populations. This study shows that the use of guideline-recommended spirometry
261 reference equations in admixed populations is limited by intra-population variation in genetic
262 ancestry. In addition, this study contributes to the ongoing debate regarding the use of
263 race/ethnicity in clinical algorithms by demonstrating that race/ethnicity, alone, is insufficient for
264 estimating lung function in a large proportion of minority children.¹⁹ We demonstrate that the
265 use of genetic ancestry data instead of race/ethnicity in lung function equations could improve
266 lung function prediction equations for admixed populations. That said, we recognize that
267 incorporating race/ethnicity captures essential epidemiologic information that relates to lung
268 function and overall clinical outcomes not captured by genetic ancestry.¹⁹

269 **INTERPRETATION**

270 This study demonstrates that guideline-recommended spirometry reference equations, which rely
271 on self-identified race/ethnicity, inconsistently fit homogenous racial/ethnic groups after
272 classifying these populations by their genetic ancestry distribution. Recent scientific advances
273 allow for estimates of genetic ancestry to be easily and inexpensively measured as exemplified
274 by direct-to-consumer genetic ancestry testing. Unfortunately, current clinical practice has been
275 slow to adopt and incorporate modern advances from the Human Genome Project, which was
276 started over thirty years ago. Accordingly, spirometry is due for reference equations that
277 incorporate genetic ancestry, either for more precise application of the current equations or the

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

278 derivation and utilization of new equations. Meanwhile, the increasing availability of genetic
279 data in the clinical setting can be leveraged to guide selection of the most appropriate spirometry
280 reference equation. Our findings suggest that African American children with African ancestry
281 above 80% are best fit by equations derived in African American populations, whereas those
282 with African ancestry below 80% are best fit by the GLI composite equation. Puerto Rican
283 children with African ancestry below 20% are best fit by equations derived in White populations.
284 However, there is no appropriately fitted equation for Puerto Rican children with African
285 ancestry above 20%. As a result, extra caution must be taken when choosing a spirometry
286 reference equation and interpreting its results among racially/ethnically diverse populations.
287 Racially/ethnically diverse individuals might be best served if their spirometric measurements
288 are evaluated using multiple equations, with each result taken in the context of clinical
289 symptoms.

290 **Acknowledgments**

291 **Author contributions:** J. W. contributed to the conception and design of the work, analysis and
292 interpretation of the data, and drafting and revision of the manuscript. C. E. contributed to the
293 acquisition of data. J. R. E., J. R. R.-S., L. N. B., and E. G. B.: contributed to the analysis and
294 interpretation of the data, drafting and revision of the manuscript, and critical review of the
295 manuscript.

296

297 **Other contributions:** The authors thank the participants and their families for their contribution,
298 as well as the health care professionals and clinics for their support and participation in the
299 Genes-environments & Admixture in Latino Americans Study and the Study of African

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

300 Americans, Asthma, Genes & Environments. In particular, the authors thank the study
301 coordinator Sandra Salazar and the recruiters who obtained the data: Kelley Meade, MD, Adam
302 Davis, MPH, MA, Lisa Caine, Emerita Brigino-Buenaventura, MD, Duanny Alva, MD, Gaby
303 Ayala-Rodriguez, Ulysses Burley, Elizabeth Castellanos, Jaime Colon, Denise DeJesus, Iliana
304 Flexas, Blanca Lopez, Brenda Lopez, MD, Louis Martos, Vivian Medina, Juana Olivo, Mario
305 Peralta, Esther Pomares, MD, Jihan Quraishi, Johanna Rodriguez, Shahdad Saeedi, Dean Soto,
306 Ana Taveras, and Emmanuel Viera. The authors would also like to thank Andrew Zeiger, Oona
307 Risse-Adams, María G. Contreras, Fernando Picazo, MD, Martin Rofael, MD, Angel C. Y. Mak,
308 PhD, Donglei Hu, PhD, Scott Huntsman, MS, Eric Wohlford, MD, PhD, Sam S. Oh, PhD, MPH,
309 Maria Pino-Yanes, PhD, L. Keoki Williams, MD, MPH, Michael A. Lenoir, MD, Noah A.
310 Zaitlen, PhD, Elad Ziv, MD, Thomas Nuckton, MD, Christopher Gignoux, PhD, and Joshua
311 Galanter, MD for their advice on project conception and analysis.

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

312 **References**

- 313 1. Burchard EG, Ziv E, Coyle N, et al. The importance of race and ethnic background in
314 biomedical research and clinical practice. *N Engl J Med* 2003;348(12):1170–1175.
- 315 2. Phimister EG. Medicine and the racial divide. *N Engl J Med* 2003;348(12):1081–1082.
- 316 3. Cooper RS, Kaufman JS, Ward R. Race and genomics. *N Engl J Med* 2003;348(12):1166–
317 1170.
- 318 4. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric Reference Values from a Sample of
319 the General U.S. Population. *Am J Respir Crit Care Med* 1999;159(1):179–187.
- 320 5. Quanjer PH, Stanojevic S, Stocks J, Cole TJ. GLI-2012 All-Age Multi-Ethnic Reference
321 Values for Spirometry. :15.
- 322 6. The use of racial, ethnic, and ancestral categories in human genetics research. *Am J Hum*
323 *Genet* 2005;77(4):519–532.
- 324 7. Goldstein DB, Hirschhorn JN. In genetic control of disease, does “race” matter? *Nat Genet*
325 2004;36(12):1243–1244.
- 326 8. Yaeger R, Avila-Bront A, Abdul K, et al. Comparing genetic ancestry and self-described
327 race in african americans born in the United States and in Africa. *Cancer Epidemiol*
328 *Biomarkers Prev* 2008;17(6):1329–1338.
- 329 9. Yang JJ, Burchard EG, Choudhry S, et al. Differences in allergic sensitization by self-
330 reported race and genetic ancestry. *J Allergy Clin Immunol* 2008;122(4):820-827.e9.
- 331 10. Kumar et al. - 2010 - Genetic Ancestry in Lung-Function Predictions.pdf.
- 332 11. Graham BL, Steenbruggen I, Miller MR, et al. Standardization of Spirometry 2019 Update.
333 An Official American Thoracic Society and European Respiratory Society Technical
334 Statement. *Am J Respir Crit Care Med* 2019;200(8):e70–e88.
- 335 12. Quanjer PH, Stanojevic S, Enright PL, et al. MULTI-ETHNIC REFERENCE VALUES FOR
336 SPIROMETRY FOR THE 3-95 YEAR. 2012;39:11.
- 337 13. Myers JE. Differential ethnic standards for lung functions, or one standard for all? *S Afr Med*
338 *J* 1984;65(19):768–772.
- 339 14. Chien JW, Sullivan KM. Carbon monoxide diffusion capacity: how low can you go for
340 hematopoietic cell transplantation eligibility? *Biol Blood Marrow Transplant*
341 2009;15(4):447–453.
- 342 15. Fulambarker A, Copur AS, Javeri A, Jere S, Cohen ME. Reference Values for Pulmonary
343 Function in Asian Indians Living in the United States. *Chest* 2004;126(4):1225–1233.

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

- 344 16. Hankinson JL, Kawut SM, Shahar E, Smith LJ, Stukovsky KH, Barr RG. Performance of
345 American Thoracic Society-Recommended Spirometry Reference Values in a Multiethnic
346 Sample of Adults. *Chest* 2010;137(1):138–145.
- 347 17. Korotzer B, Ong S, Hansen JE. Ethnic Differences in Pulmonary Function in Healthy
348 Nonsmoking Asian-Americans and European-Americans. *Am J Respir Crit Care Med*
349 2000;161(4):1101–1108.
- 350 18. LaVange L, Davis SM, Hankinson J, et al. Spirometry Reference Equations from the
351 HCHS/SOL (Hispanic Community Health Study/Study of Latinos). *Am J Respir Crit Care*
352 *Med* 2017;196(8):993–1003.
- 353 19. Borrell LN, Elhawary JR, Fuentes-Afflick E, et al. Race and Genetic Ancestry in Medicine -
354 A Time for Reckoning with Racism. *N Engl J Med* 2021;
- 355 20. Quanjer PH, Pretto JJ, Brazzale DJ, Boros PW. Grading the severity of airways obstruction:
356 new wine in new bottles. *Eur Respir J* 2014;43(2):505–512.
- 357 21. Hall GL, Thompson BR, Stanojevic S, et al. The Global Lung Initiative 2012 reference
358 values reflect contemporary Australasian spirometry: Global Lung Initiative reference
359 values. *Respirology* 2012;17(7):1150–1151.
- 360 22. Bonner R, Lum S, Stocks J, Kirkby J, Wade A, Sonnappa S. Applicability of the Global
361 Lung Function Spirometry Equations in Contemporary Multiethnic Children. :2.
- 362 23. Brazzale DJ, Hall GL, Pretto JJ. Effects of adopting the new global lung function initiative
363 2012 reference equations on the interpretation of spirometry. *Respiration* 2013;86(3):183–
364 189.
- 365 24. Hulo S, Broucker V de, Giovannelli J, et al. Global Lung Function Initiative reference
366 equations better describe a middle-aged, healthy French population than the European
367 Community for Steel and Coal values. *Eur Respir J* 2016;48(6):1779–1781.
- 368 25. Belo J. Reference values for spirometry in elderly individuals: a cross-sectional study of
369 different reference equations. 2018;9.
- 370 26. Kumar R, Seibold MA, Aldrich MC, et al. Genetic Ancestry in Lung-Function Predictions. *N*
371 *Engl J Med* 2010;363(4):321–330.
- 372 27. Brehm JM, Acosta-Pérez E, Klei L, et al. African ancestry and lung function in Puerto Rican
373 children. *Journal of Allergy and Clinical Immunology* 2012;129(6):1484-1490.e6.
- 374 28. Pino-Yanes M, Thakur N, Gignoux CR, et al. Genetic ancestry influences asthma
375 susceptibility and lung function among Latinos. *Journal of Allergy and Clinical*
376 *Immunology* 2015;135(1):228–235.

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

- 377 29. Nishimura KK, Galanter JM, Roth LA, et al. Early-Life Air Pollution and Asthma Risk in
378 Minority Children. The GALA II and SAGE II Studies. *Am J Respir Crit Care Med*
379 2013;188(3):309–318.
- 380 30. Miller MR. Standardisation of spirometry. *European Respiratory Journal* 2005;26(2):319–
381 338.
- 382 31. Hoffmann TJ, Zhan Y, Kvale MN, et al. Design and coverage of high throughput genotyping
383 arrays optimized for individuals of East Asian, African American, and Latino race/ethnicity
384 using imputation and a novel hybrid SNP selection algorithm. *Genomics* 2011;98(6):422–
385 430.
- 386 32. Theodore Lytras. rspirometry: Implementation of Spirometry Equations [Internet]. 2020. Available
387 from: <https://github.com/thlytras/rspirometry>
- 388 33. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna,
389 Austria: R Foundation for Statistical Computing; 2020. Available from: [https://www.R-](https://www.R-project.org/)
390 [project.org/](https://www.R-project.org/)
- 391 34. Thompson BR, Stanojevic S, Abramson MJ, et al. The all-age spirometry reference ranges
392 reflect contemporary Australasian spirometry. *Respirology* 2011;16(6):912–917.
- 393 35. Ben Saad H, El Attar MN, Hadj Mabrouk K, et al. The recent multi-ethnic global lung
394 initiative 2012 (GLI2012) reference values don't reflect contemporary adult's North
395 African spirometry. *Respir Med* 2013;107(12):2000–2008.
- 396 36. Lakens D. Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-
397 Analyses. *Soc Psychol Personal Sci* 2017;8(4):355–362.
- 398 37. Bryc K, Durand EY, Macpherson JM, Reich D, Mountain JL. The genetic ancestry of
399 African Americans, Latinos, and European Americans across the United States. *Am J Hum*
400 *Genet* 2015;96(1):37–53.
- 401 38. Celedón JC, Roman J, Schraufnagel DE, Thomas A, Samet J. Respiratory health equality in
402 the United States. The American thoracic society perspective. *Ann Am Thorac Soc*
403 2014;11(4):473–479.
- 404 39. The Annie E. Casey Foundation. Child population by race in the United States [Internet].
405 Kids Count Data Center. Available from: [https://datacenter.kidscount.org/data/tables/103-](https://datacenter.kidscount.org/data/tables/103-child-population-by-race#detailed/1/any/false/1729,37,871,870,573,869,36,868,867,133/68,69,67,12,70,66,71,72/423,424)
406 [child-population-by-](https://datacenter.kidscount.org/data/tables/103-child-population-by-race#detailed/1/any/false/1729,37,871,870,573,869,36,868,867,133/68,69,67,12,70,66,71,72/423,424)
407 [race#detailed/1/any/false/1729,37,871,870,573,869,36,868,867,133/68,69,67,12,70,66,71,7](https://datacenter.kidscount.org/data/tables/103-child-population-by-race#detailed/1/any/false/1729,37,871,870,573,869,36,868,867,133/68,69,67,12,70,66,71,72/423,424)
408 [2/423,424](https://datacenter.kidscount.org/data/tables/103-child-population-by-race#detailed/1/any/false/1729,37,871,870,573,869,36,868,867,133/68,69,67,12,70,66,71,72/423,424)
- 409 40. Arigliani M, Canciani MC, Mottini G, et al. Evaluation of the Global Lung Initiative 2012
410 Reference Values for Spirometry in African Children. *Am J Respir Crit Care Med*
411 2017;195(2):229–236.

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

- 412 41. Centers for Disease Control and Prevention (CDC). Table 2-1 Lifetime Asthma Prevalence
413 Percents by Age, United States: National Health Interview Survey, 2018 [Internet]. 2019
414 [cited 2021 Feb 4]; Available from: <https://www.cdc.gov/asthma/nhis/2018/table2-1.htm>
- 415 42. Health, United States, 2010: With Special Feature on Death and Dying. Hyattsville (MD):
416 National Center for Health Statistics (US); 2011.
- 417 43. United States Census Bureau. Hispanic or Latino Origin by Race. Table ID: 803002
418 [Internet]. 2020; Available from:
419 <https://data.census.gov/cedsci/table?q=HISPANIC%20OR%20LATINO%20ORIGIN%20BY%20RACE&t=Hispanic%20or%20Latino&tid=ACSDT1Y2019.B03002&hidePreview=false>
420
421
- 422 44. United States Census Bureau. Quick Facts: Puerto Rico [Internet]. 2020; Available from:
423 <https://www.census.gov/quickfacts/PR>
- 424 45. Quanjer PH, Stocks J, Cole TJ, Hall GL, Stanojevic S. Influence of secular trends and
425 sample size on reference equations for lung function tests. *Eur Respir J* 2011;37(3):658–
426 664.
- 427 46. Stocks J, Sonnappa S, Lum S. Lung function testing in children: importance of race and
428 ethnic-specific reference equations. *Expert Review of Respiratory Medicine* 2014;8(5):527–
429 531.
- 430 47. Kirkby J, Bonner R, Lum S, et al. Interpretation of pediatric lung function: impact of
431 ethnicity. *Pediatr Pulmonol* 2013;48(1):20–26.
- 432

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

Table 1. Distribution of selected characteristics for participants in GALA II and SAGE based on race/ethnicity and genetic ancestry categories: 2006-2014.

	African American Children			Puerto Rican Children		
	All (N = 275)	African Ancestry (Median: 81.3%)		All (N = 324)	African Ancestry (Median: 21.3%)	
		≤Median (N = 138)	>Median (N = 137)		≤Median (N = 162)	>Median (N = 162)
Boys, %	41	43	39	45	43	48
Age, year	16.2 (3.8)	15.9 (3.9)	16.5 (3.7)	14.0 (3.0)	14.2 (3.2)	13.8 (2.7)
Height, cm	162.3 (13.6)	161.4 (13.6)	163.1 (13.6)	158.6 (11.6)	157.9 (12.7)	159.4 (10.3)
Weight, kg	67.3 (24.4)	66.1 (22.7)	68.4 (26.1)	57.8 (18.8)	56.6 (16.9)	59.0 (20.5)
Ancestry, %						
African	79.5 (9.6)	73.2 (9.6)	85.9 (3.2)*	22.7 (10.5)	15.1 (3.3)	30.2 (9.8)*
Native	—	—	—	10.2 (3.1)	10.6 (3.5)	9.8 (2.7)*
zFEV ₁	0.35 (0.95)	0.41 (0.97)	0.29 (0.94)	0.56 (1.21)	0.69 (1.14)	0.43 (1.28)*
zFVC	0.49 (0.99)	0.56 (0.97)	0.43 (1.01)	0.39 (1.27)	0.55 (1.15)	0.23 (1.36)*

Definition of abbreviations: zFEV₁ = z-score for FEV₁; zFVC = z-score for FVC.

Results are presented as mean (SD) unless otherwise specified. Spirometry z-scores based on Global Lung Function Initiative 2012 equations, using the African American-derived equation for African American children and the composite equation for Puerto Rican children.⁵

*Differences between subpopulations: $P < 0.05$ (Student's t-test and Wilcoxon Rank Sum Test).

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

Table 2a. Fit of Spirometry Equation Predictions to Spirometric Measurements in African American Children: 2006-2014

	African Ancestry		
	All (N = 275)	≤ Median 30.7 to 81.3% (N = 138)	> Median 81.4 to 100% (N = 137)
zFEV₁			
NHANES III, White	-0.95 (-1.07, -0.84) [24.0%]	-0.88 (-1.04, -0.73) [23.9%]	-1.03 (-1.19, -0.86) [24.1%]
NHANES III, African American	0.17 (0.06, 0.28) ^{*†} [2.5%]	0.24 (0.08, 0.39) [*] [2.2%]	0.11 (-0.04, 0.26) ^{*†} [2.9%]
NHANES III, Mexican American	-1.19 (-1.30, -1.07) [31.6%]	-1.12 (-1.27, -0.96) [28.3%]	-1.26 (-1.42, -1.09) [35.0%]
GLI, White	-0.88 (-0.99, -0.77) [18.2%]	-0.82 (-0.97, -0.67) [18.8%]	-0.94 (-1.09, -0.79) [17.5%]
GLI, African American	0.35 (0.24, 0.46) [*] [1.8%]	0.41 (0.25, 0.58) [1.4%]	0.29 (0.13, 0.44) [*] [2.2%]
GLI, Composite	-0.32 (-0.43, -0.20) ^{*†} [5.8%]	-0.25 (-0.42, -0.09) ^{*†} [5.8%]	-0.38 (-0.54, -0.22) [5.8%]
zFVC			
NHANES III, White	-0.85 (-0.97, -0.74) [19.3%]	-0.79 (-0.94, -0.63) [14.5%]	-0.92 (-1.09, -0.75) [24.1%]
NHANES III, African American	0.34 (0.23, 0.45) ^{*†} [1.8%]	0.41 (0.26, 0.56) [0.7%]	0.27 (0.10, 0.43) ^{*†} [2.9%]
NHANES III, Mexican American	-1.00 (-1.11, -0.88) [24.0%]	-0.93 (-1.09, -0.77) [21.0%]	-1.06 (-1.23, -0.90) [27.0%]
GLI, White	-0.79 (-0.91, -0.68) [17.8%]	-0.74 (-0.89, -0.58) [14.5%]	-0.85 (-1.01, -0.69) [21.2%]
GLI, African American	0.49 (0.37, 0.61) [2.2%]	0.56 (0.39, 0.72) [1.4%]	0.43 (0.26, 0.60) [2.9%]
GLI, Composite	-0.14 (-0.27, -0.02) ^{*†} [6.2%]	-0.08 (-0.25, 0.10) ^{*†} [5.8%]	-0.21 (-0.39, -0.02) [*] [6.6%]

Definition of abbreviations: NHANES III = Third National Health and Nutrition Examination Survey; GLI = Global Lung Function Initiative; zFEV₁ = z-score for FEV₁; zFVC = z-score for FVC.

Results are presented as mean (95% confidence interval) [percent below the fifth percentile].

^{*}Equation provides sufficient fit: mean z-score < |0.5| (*P* < 0.05) with 95% confidence intervals inside interval [-0.5, 0.5]).^{21,34,35}

[†]Equation appropriate for population/subpopulation (provides sufficient fit for both FEV₁ and FVC).

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

Table 2b. Fit of Spirometry Equation Predictions to Spirometric Measurements in Puerto Rican Children: 2006-2014

	All (N = 324)	African Ancestry	
		≤ Median 6.4 to 21.3% (N = 162)	> Median 21.4 to 87.5% (N = 162)
zFEV₁			
NHANES III, White	-0.16 (-0.28, -0.04) [*] [7.1%]	-0.05 (-0.21, 0.12) ^{*†} [4.9%]	-0.27 (-0.46, -0.09) [*] [9.3%]
NHANES III, African American	0.90 (0.78, 1.02) [0.3%]	1.02 (0.86, 1.18) [0.0%]	0.78 (0.60, 0.96) [0.6%]
NHANES III, Mexican American	-0.41 (-0.53, -0.28) [11.7%]	-0.28 (-0.45, -0.12) [*] [9.3%]	-0.53 (-0.72, -0.35) [14.2%]
GLI, White	-0.06 (-0.19, 0.06) ^{*†} [6.8%]	0.06 (-0.11, 0.22) ^{*†} [4.3%]	-0.18 (-0.36, 0.01) [*] [9.3%]
GLI, African American	1.23 (1.10, 1.36) [0.3%]	1.35 (1.18, 1.53) [0.0%]	1.10 (0.91, 1.30) [0.6%]
GLI, Composite	0.56 (0.43, 0.69) [1.5%]	0.69 (0.51, 0.86) [0.6%]	0.43 (0.24, 0.63) [2.5%]
zFVC			
NHANES III, White	-0.41 (-0.53, -0.30) [12.0%]	-0.28 (-0.43, -0.12) ^{*†} [7.4%]	-0.55 (-0.73, -0.37) [16.7%]
NHANES III, African American	0.76 (0.64, 0.87) [0.6%]	0.89 (0.74, 1.05) [0.0%]	0.62 (0.44, 0.80) [1.2%]
NHANES III, Mexican American	-0.56 (-0.68, -0.44) [15.7%]	-0.42 (-0.58, -0.27) [10.5%]	-0.70 (-0.88, -0.52) [21.0%]
GLI, White	-0.33 (-0.46, -0.21) ^{*†} [11.1%]	-0.19 (-0.35, -0.03) ^{*†} [6.8%]	-0.48 (-0.66, -0.29) [15.4%]
GLI, African American	1.00 (0.87, 1.13) [0.6%]	1.15 (0.98, 1.32) [0.0%]	0.86 (0.66, 1.06) [1.2%]
GLI, Composite	0.39 (0.25, 0.53) [4.3%]	0.55 (0.37, 0.73) [1.9%]	0.23 (0.02, 0.44) [*] [6.8%]

Definition of abbreviations: NHANES III = Third National Health and Nutrition Examination Survey; GLI = Global Lung Function Initiative; zFEV₁ = z-score for FEV₁; zFVC = z-score for FVC.

Results are presented as mean (95% confidence interval) [percent below the fifth percentile].

^{*}Equation provides sufficient fit: mean z-score < |0.5| (*P* < 0.05) with 95% confidence intervals inside interval [-0.5, 0.5]).^{21,34,35}

[†]Equation appropriate for population/subpopulation (provides sufficient fit for both FEV₁ and FVC).

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

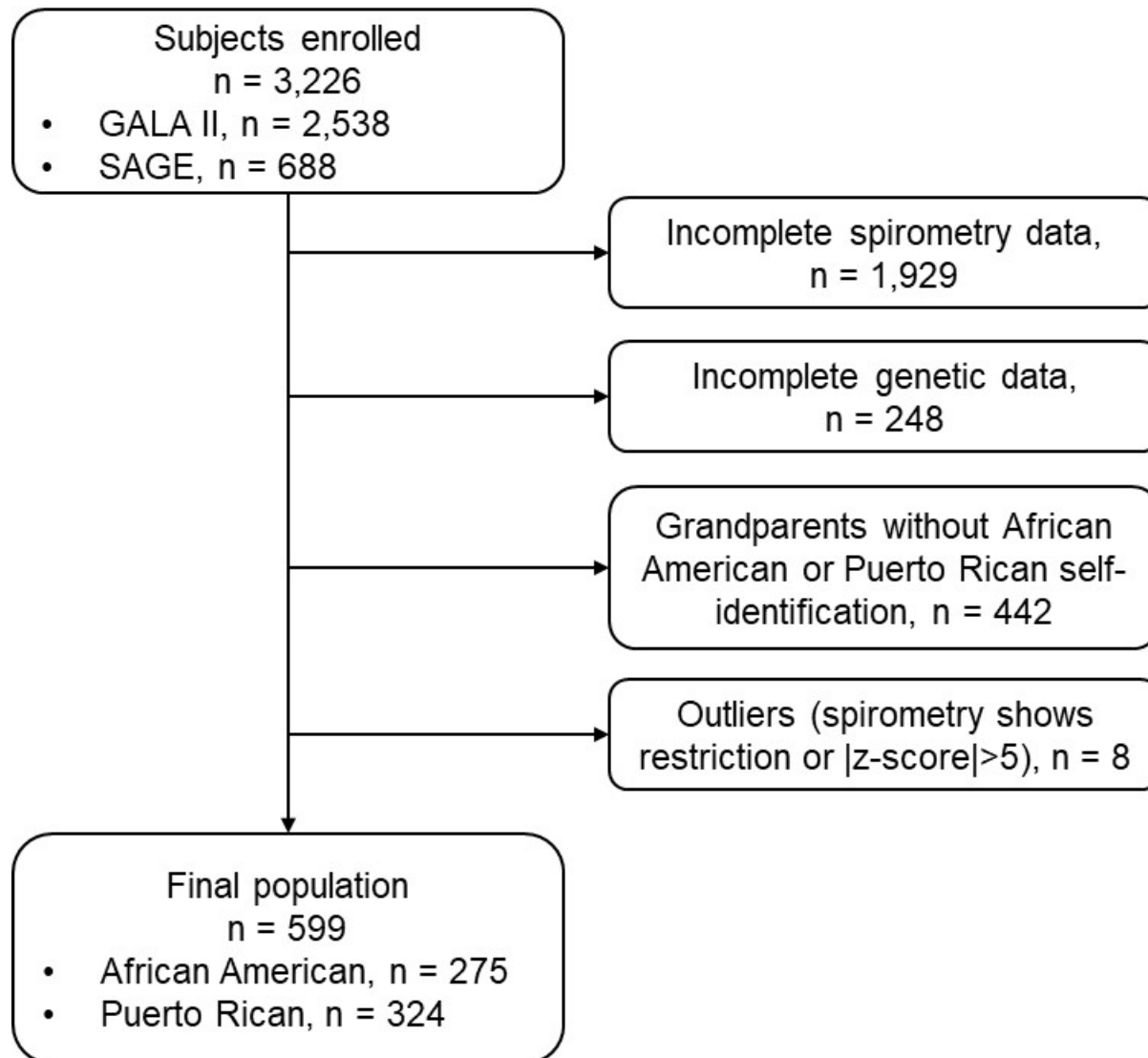


Figure 1. Study population. African American and Puerto Rican children (8-21 years old) recruited as part of parallel case-control studies of asthma from the Genes-environments & Admixture in Latino Americans (GALA II) study and the Study of African Americans, Asthma, Genes, & Environments (SAGE).

RACIAL/ETHNIC-BASED SPIROMETRY: ACCURACY FOR ADMIXED

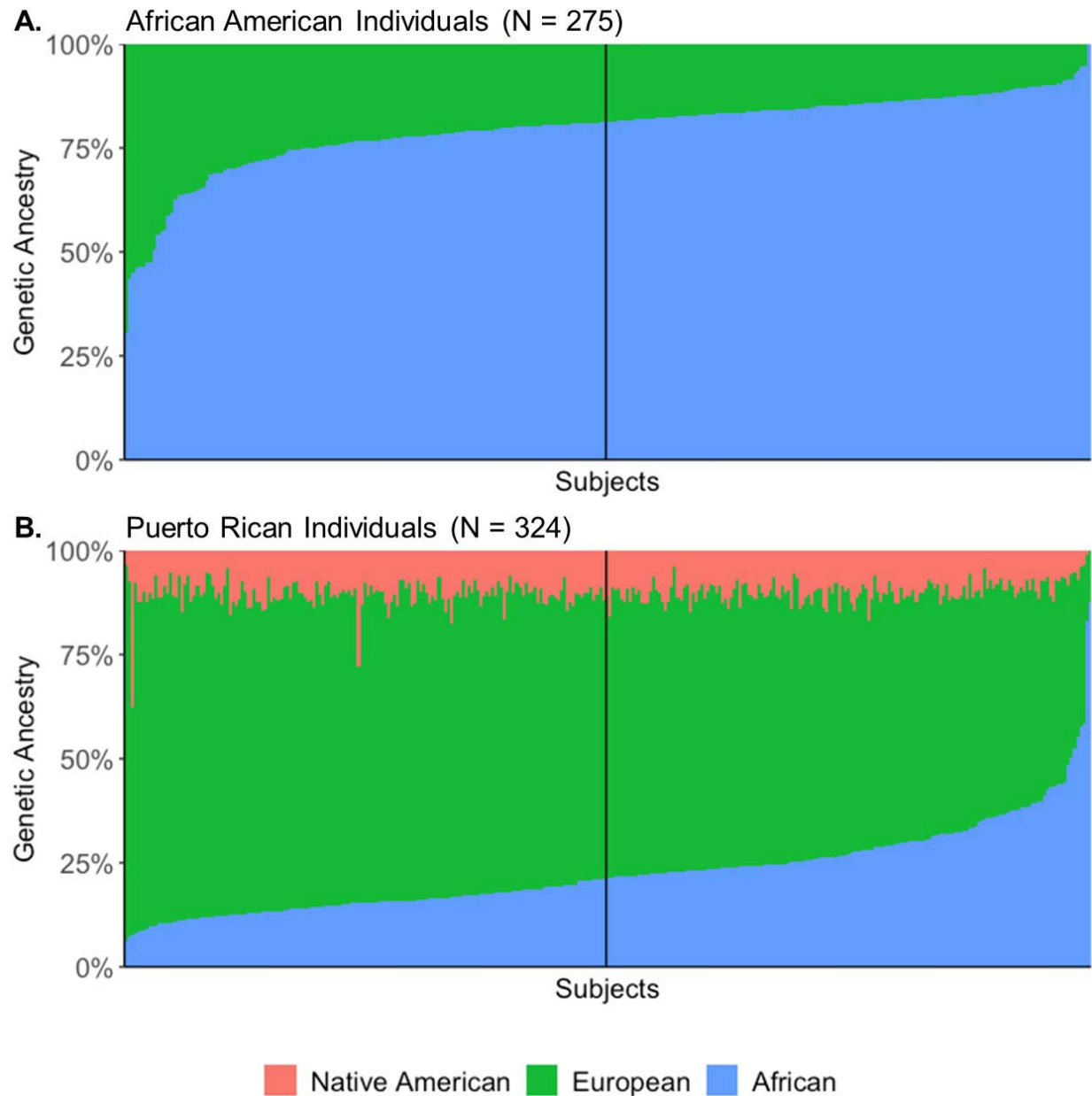


Figure 2. Genetic admixture in the (A) African American and (B) Puerto Rican children. The vertical black lines partition each population into paired subpopulations based on proportions of genetic ancestry. The African American and Puerto Rican subpopulations include individuals with African ancestry above and below the median.

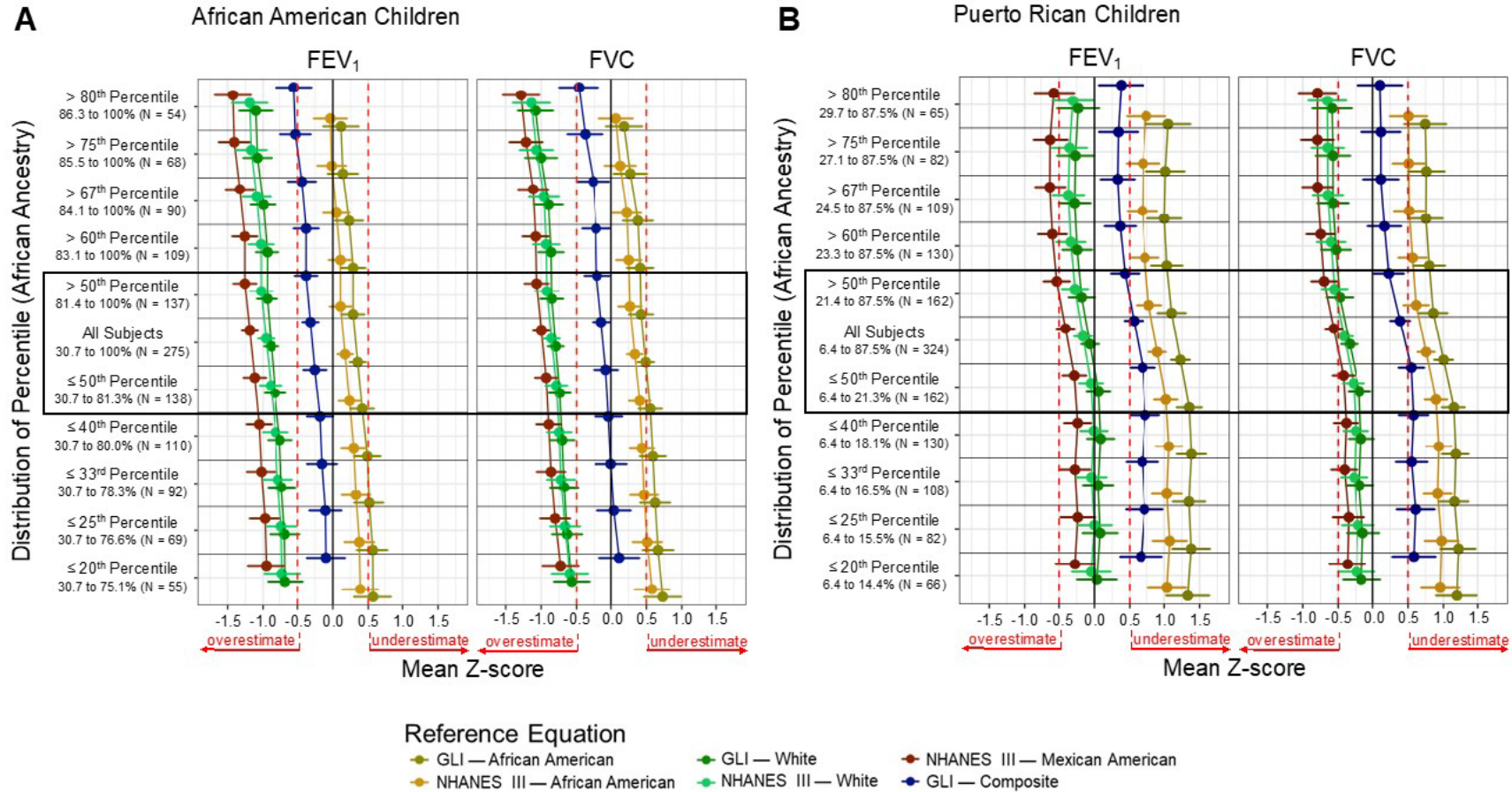


Figure 3. Spirometry reference equation fit for predicting FEV₁ and FVC in (A) African American and (B) Puerto Rican children with sensitivity analysis showing equation fit in subpopulations with increasingly disparate proportions of genetic ancestry, including ancestry tertiles, quartiles, and quintiles. Dot-and-whisker plots show mean z-score with 95% confidence intervals for each population/subpopulation. *Solid black vertical lines* indicate mean z-score of zero (perfect equation fit). *Dashed red vertical lines* indicate the interval [-0.5, 0.5] for sufficient equation fit. *Black boxes* outline the primary analysis of the study with sensitivity analysis above and below.