

Diagnostic Impact of a Race-Composite Pulmonary Function Test Results Interpretation Strategy



To the Editor:

It has been observed historically that lung function varies by race and ethnicity, even after accounting for differences in height, sex, and age. Until recently, guidelines have recommended using race-specific reference values to interpret pulmonary function test (PFT) results.¹ Many potential confounding factors contribute to population differences in lung function that transcend race, including socioeconomic status and environmental exposures.^{2,3} Furthermore, because race is a social construct, the use of race-based interpretation strategies runs the risk of perpetuating structural health care disparities.^{4,5} Although the limitations of using race-based reference equations increasingly are recognized, potential risks of not including race in interpretative strategies for PFT results also exist,⁶ and the best alternative approach currently is not clear.⁷ In an effort to balance these risks and benefits and to avoid perpetuating systemic

racism, the American Thoracic Society recently published guidelines recommending a transition to race-composite reference equation use in interpretation of PFT findings.

The Global Lung Initiative (GLI) established reference PFT equations from healthy participants of multiple races and ethnicities who did not use tobacco.⁸ One approach to performing race-neutral PFT results interpretation is to use a composite of reference values from all the included racial and ethnic groups, also known as GLI Other (GLI-O). The usefulness of GLI-O recently was examined in two articles. First, McCormack et al⁹ reanalyzed data from the Third National Health and Nutrition Examination Survey (NHANES III) survey and found that the use of FEV₁ z scores derived from GLI-O predicted mortality similarly to race-based approaches. Second, Baugh et al¹⁰ reanalyzed data from the Subpopulations and Intermediate Outcome Measures in COPD Study (SPIROMICS) and concluded that FEV₁ % predicted calculated using GLI-O more consistently identifies Black patients at high risk, including those with previously unidentified COPD. These studies demonstrated that low lung function (especially FEV₁) in Black patients should not be considered normal from an epidemiologic perspective.

Methods

This study was designed to help understand how the use of race-composite reference equations may impact the usefulness of spirometry as a diagnostic tool. We analyzed PFTs performed at the University of Rochester Medical Center between 1991 and 2020 using GLI race-specific reference equations (as recommended previously) and GLI-O reference equations. PFT records from White and Black patients aged 12 to 70 years were included in this analysis. In this study, race was determined by self-report documented at the

time PFTs originally were performed. Twelve thousand two hundred six PFT results were reanalyzed, 11,322 from White patients and 884 from Black patients. This study was approved by the University of Rochester Medical Center Institutional Review Board. We compared the number of White and Black patients with normal and abnormal z scores for FEV₁, FVC, and FEV₁ to FVC ratio (defined as $z \geq -1.645$ and $z < -1.645$, respectively). An obstructive defect was defined as FEV₁ to FVC ratio z score of < -1.645 , and a potential restrictive defect was defined as FVC z score of < -1.645 with FEV₁ to FVC ratio z score of ≥ -1.645 .

Results

A summary of demographic data for all included patients is provided in Table 1. No Black patients were found to have PFT findings consistent with obstruction only when using race-specific reference values. Additionally, only two Black patients were found to have PFT results suggestive of potential restriction when

using only race-specific reference values. As a result, although four groups of White patients are presented in Table 1, only three groups of Black patients are presented. Note that White patients whose FEV₁ results were abnormally low ($z < -1.645$) using both race-specific and GLI-O reference equations were older and taller in comparison with White patients with

TABLE 1] Patient Demographic Features

| Variable | Black Patients | | | | White Patients | | | | |
|-------------------------------|----------------|------------------|---------------------|--------------------|----------------|------------------|---------------------|-----------------------------|--------------------|
| | All | Agreed Normality | GLI-O Abnormal Only | Agreed Abnormality | All | Agreed Normality | GLI-O Abnormal Only | Race-Specific Abnormal Only | Agreed Abnormality |
| FEV ₁ | n = 884 | n = 444 | n = 108 | n = 331 | n = 11,322 | n = 6,561 | n = 145 | n = 589 | n = 4,027 |
| Age, y | 51.8 ± 18.1 | 50.7 ± 19.4 | 49.8 ± 19.1 | 53.9 ± 15.6 | 61.7 ± 17.0 | 60.0 ± 18.4 | 64.4 ± 9.8 | 62.6 ± 18.5 | 64.2 ± 14.0 |
| Sex, % male | 41.2 | 34.9 | 41.7 | 49.5 | 45.7 | 43.2 | 57.9 | 41.6 | 50.0 |
| Height, cm | 167.2 ± 9.6 | 166.5 ± 9.4 | 167.3 ± 10.3 | 168.3 ± 9.5 | 167.7 ± 11.7 | 167.6 ± 9.9 | 174.3 ± 9.2 | 166.2 ± 10.5 | 167.8 ± 14.2 |
| Absolute FEV ₁ , L | 2.1 ± 0.8 | 2.5 ± 0.7 | 2.1 ± 0.6 | 1.4 ± 0.5 | 2.2 ± 1.0 | 2.7 ± 0.9 | 2.4 ± 0.5 | 2.0 ± 0.6 | 1.4 ± 0.6 |
| Absolute FVC, L | 2.8 ± 0.9 | 3.2 ± 0.9 | 2.8 ± 0.8 | 2.3 ± 0.7 | 3.1 ± 1.1 | 3.6 ± 1.1 | 3.3 ± 0.8 | 2.8 ± 0.9 | 2.4 ± 0.8 |
| Obstruction | n = 884 | n = 611 | n = 17 | n = 256 | n = 11,322 | n = 7,691 | n = 390 | n = 2 | n = 3,239 |
| Age, y | 51.8 ± 18.1 | 52.9 ± 17.3 | 42.2 ± 23.5 | 49.8 ± 19.3 | 61.7 ± 17.0 | 61.0 ± 17.5 | 60.0 ± 19.5 | 49.0 ± 1.4 | 63.5 ± 15.3 |
| Sex, % male | 41.2 | 35.7 | 35.3 | 54.7 | 45.7 | 45.4 | 42.3 | 100.0 | 46.9 |
| Height, cm | 167.2 ± 9.6 | 166.8 ± 9.7 | 169.4 ± 6.5 | 168.1 ± 9.4 | 167.7 ± 11.7 | 167.7 ± 10.1 | 167.4 ± 10.0 | 169.9 ± 0.5 | 167.5 ± 14.9 |
| Absolute FEV ₁ , L | 2.1 ± 0.8 | 2.3 ± 0.8 | 2.2 ± 0.8 | 1.6 ± 0.7 | 2.2 ± 1.0 | 2.5 ± 0.9 | 2.2 ± 0.9 | 1.8 ± 0.3 | 1.5 ± 0.8 |
| Absolute FVC, L | 2.8 ± 0.9 | 2.9 ± 0.9 | 3.1 ± 1.0 | 2.7 ± 0.9 | 3.1 ± 1.1 | 3.3 ± 1.1 | 3.2 ± 1.2 | 2.7 ± 0.4 | 2.8 ± 1.1 |
| Potential restriction | n = 611 | n = 393 | n = 91 | n = 125 | n = 7,691 | n = 5,719 | n = 242 | n = 299 | n = 1,431 |
| Age, y | 52.9 ± 17.3 | 54.7 ± 15.7 | 51.3 ± 17.6 | 55.7 ± 13.4 | 61.0 ± 17.5 | 60.2 ± 18.1 | 66.4 ± 9.7 | 60.8 ± 17.2 | 63.4 ± 15.5 |
| Sex, % male | 35.7 | 35.4 | 34.1 | 38.4 | 45.4 | 42.0 | 55.4 | 50.5 | 56.5 |
| Height, cm | 166.8 ± 9.7 | 166.8 ± 9.5 | 166.7 ± 10.1 | 167.1 ± 10.2 | 167.7 ± 10.1 | 167.5 ± 10.0 | 170.2 ± 10.0 | 167.9 ± 10.0 | 168.4 ± 10.6 |
| Absolute FEV ₁ , L | 2.3 ± 0.8 | 2.6 ± 0.7 | 2.0 ± 0.6 | 1.6 ± 0.5 | 2.5 ± 0.9 | 2.8 ± 0.9 | 2.2 ± 0.6 | 2.1 ± 0.6 | 1.7 ± 0.6 |
| Absolute FVC, L | 2.9 ± 0.9 | 3.2 ± 0.9 | 2.5 ± 0.7 | 2.0 ± 0.6 | 3.3 ± 1.1 | 3.6 ± 1.0 | 2.9 ± 0.7 | 2.7 ± 0.7 | 2.2 ± 0.7 |

Data are presented as No. or mean ± SD, unless otherwise indicated. GLI-O = Global Lung Initiative Other.

consistently normal FEV₁ z scores using both race-specific and GLI-O reference equations. Demographic features among Black patients were comparable among the three FEV₁ classification groups, except that patients with abnormal FEV₁ z scores using either race-specific or GLI-O reference equations were more likely to be male.

Table 2 compares PFT values obtained using GLI race-specific references (in rows) vs race-composite GLI-O (in columns) for the PFT results of White and Black patients. The same data are shown in Figure 1 for ease of comparison. For FEV₁, 93.5% agreement was found between the two approaches in the PFT results of White patients, whereas the agreement was 87.6% in the PFT results of Black patients. The κ statistic 95% CI comparing FEV₁ z scores of White and Black patients did not overlap, indicating that the percent agreement using the two approaches was significantly higher for the PFT results of White patients than those of Black patients. Using GLI-O resulted in 5.2% of the FEV₁ values of White patients being reclassified as normal and 1.3% being reclassified as abnormally low. Using GLI-O resulted in 12.2% of FEV₁ results of Black patients being reclassified as abnormal and the FEV₁ results of only 1 patient as normal.

The percent agreement for FEV₁ to FVC ratio was better.⁸ For the PFT results of White and Black patients, the agreement between race-specific and GLI-O was 96.5% and 98.1%, respectively. Using race-specific reference equations, 28.6% of White patients' PFT results were found to be consistent with an obstructive pattern in comparison with 32.1% using GLI-O. For Black patients' PFT results, 29.0% were consistent with an obstructive pattern using race-specific reference equations compared with 30.9% using GLI-O.

We defined potential restrictive defect as a reduced FVC with preserved FEV₁ to FVC ratio. Using race-specific reference equations, 22.5% of White patients' PFT results and 20.8% of Black patients' PFT results were consistent with potential restriction. When using race-composite GLI-O, fewer PFT results (21.8%) of White patients but more PFT results (35.4%) of Black patients were classified as consistent with potential restrictive defects. This represents an absolute decrease of 0.7% in White patients' PFT results and an absolute increase of 14.6% in Black patients' PFT results. The percent agreement in diagnosing potential restriction was significantly lower for Black patients' PFT results (89.5%; $\kappa = 0.63$) compared with White patients' PFT results (93.0%; $\kappa = 0.80$).

Discussion

Adopting the race-composite GLI-O reference would result in reasonable agreement with race-specific reference equations for the diagnosis of obstructive ventilatory defects by spirometry. In contrast, about 15% more PFT results of Black patients would be reclassified as having potential restrictive defects. These results are not surprising because the FEV₁ to FVC ratio did not differ significantly by race or ethnicity, whereas Black patients have lower FVC on average than White patients in the race-specific GLI reference equations.⁸ Because health care outcomes for the reclassified groups currently are not known, the clinical implications of this reclassification remain to be determined. Extrapolating from recent studies,^{9,10} it seems premature to conclude that this group with new diagnoses with potential restriction should be dismissed as a new normal.

Differences in classification of PFT results among White patients were significantly less frequent when using GLI-O compared with race-specific reference equations. Of note, a small but significant number of White patients were found to have FEV₁ z scores of < -1.645 using GLI-O reference equations only. This finding was unexpected because in most cases, White patients with normal z scores calculated using race-specific reference equations also would be predicted to have normal z scores using GLI-O reference equations. Patients in this group were found to have borderline lung function with z score near the cutoff value of -1.645 . Demographic characteristics of patients in this group were not significantly different in comparison with demographics of the overall study population. Further research will be needed to explore characteristics of this group in greater detail. Overall, less than 5% of White patients' PFT results were classified differently with regard to obstruction or potential restriction. This is an important finding because the high agreement shown suggests low risk of harm related to underdiagnosis or overdiagnosis using different reference equations among White patients.

Our results confirm and extend the recent report from Kitazawa et al.¹¹ These investigators analyzed PFT results from 406 participants tested between 2018 and 2021. Similar to the findings of our study, the authors found an overall concordance rate of 87% when comparing interpretation of FEV₁ as normal or abnormal using race-specific Canadian, race-specific GLI, and race-composite GLI-O reference values. The most significant differences were noted comparing

TABLE 2] PFT Results Classification Using Race-Specific vs Race Neutral (GLI-O) Interpretation

| FEV ₁ | White Patients (n = 11,322) | | FEV ₁ | Black Patients (n = 884) | |
|---|---|---|---|---|---|
| | FEV ₁ GLI-O z < -1.645 (n = 4,172) | FEV ₁ GLI-O z ≥ -1.645 (n = 7,150) | | FEV ₁ GLI-O z < -1.645 (n = 439) | FEV ₁ GLI-O z ≥ -1.645 (n = 445) |
| FEV ₁ race-specific z < -1.645 (n = 4,616) | 4,027 (35.6) | 589 (5.2) | FEV ₁ race-specific z < -1.645 (n = 332) | 331 (37.4) | 1 (0.1) |
| FEV ₁ race-specific z ≥ -1.645 (n = 6,706) | 145 (1.3) | 6,561 (57.9) | FEV ₁ race-specific z ≥ -1.645 (n = 552) | 108 (12.2) | 444 (50.2) |
| Agreement | 93.5% | | Agreement | 87.6% | |
| κ (95% CI) | 0.86 (0.85-0.87) | | κ (95% CI) | 0.75 (0.71-0.80) | |
| Obstruction^a | White Patients (n = 11,322) | | Black patients (n = 884) | | |
| | GLI-O Obstructive (n = 3,629) | GLI-O Normal (n = 7,693) | | GLI-O Obstructive (n = 273) | GLI-O Normal (n = 611) |
| Race-specific obstructive (n = 3,241) | 3,239 (28.6) | 2 (0.0) | Race-specific obstructive (n = 256) | 256 (29.0) | 0 (0.0) |
| Race-specific normal (n = 8,081) | 390 (3.4) | 7,691 (67.9) | Race-specific normal (n = 628) | 17 (1.9) | 611 (69.1) |
| Agreement | 96.5% | | Agreement | 98.1% | |
| κ (95% CI) | 0.92 (0.91-0.93) | | κ (95% CI) | 0.95 (0.93-0.98) | |
| Potential restriction^b | White Patients (n = 7,691) | | Black patients (n = 611) | | |
| | GLI-O potential restriction (n=1,673) | GLI-O not suggestive of restriction (n = 6,018) | | GLI-O potential restriction (n = 216) | GLI-O not suggestive of restriction (n = 395) |
| Race-specific potential restriction (n = 1,730) | 1,431 (18.6) | 299 (3.9) | Race-specific Potential Restriction (n = 127) | 125 (20.5) | 2 (0.3) |
| Race-specific not suggestive of restriction (n = 5,961) | 242 (3.1) | 5,719 (74.4) | Race-specific not suggestive of restriction (n = 484) | 91 (14.9) | 393 (64.3) |
| Agreement | 93.0% | | Agreement | 89.5% | |
| κ (95% CI) | 0.80 (0.78-0.81) | | κ (95% CI) | 0.63 (0.57-0.70) | |

Data are presented as No. (%), unless otherwise indicated. GLI-O = Global Lung Initiative Other; PFT = pulmonary function test.

^aDefined as FEV₁ to FVC ratio z score < -1.645.

^bDefined as FVC z score < -1.645. All patients included here have normal FEV₁ to FVC ratio defined as FEV₁ to FVC ratio z score of ≥ -1.645. Patients with abnormal FEV₁ to FVC ratio were excluded in this group.

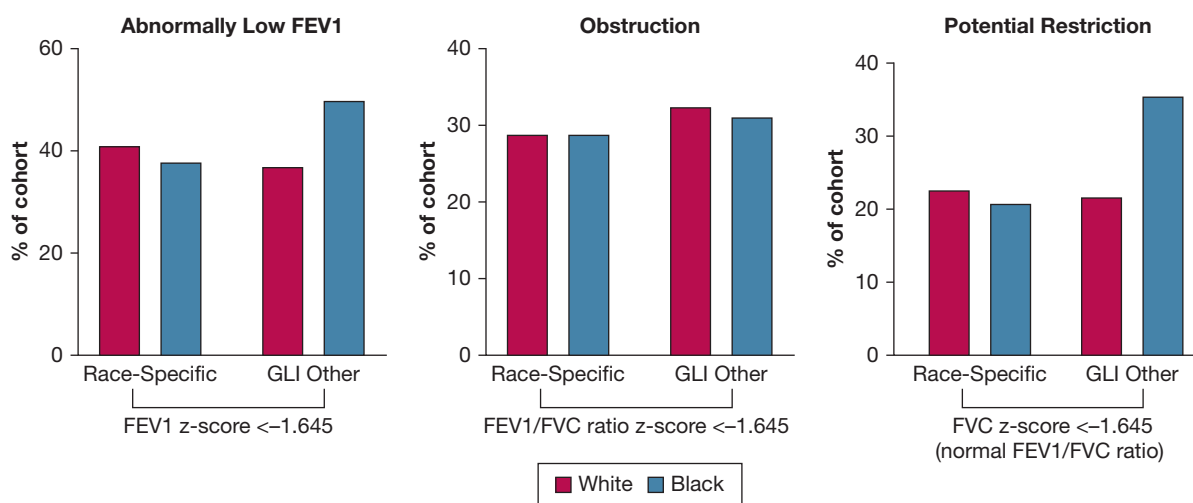


Figure 1 – A-C, Bar graphs showing observed change in pulmonary function test results classification with transition to race-neutral interpretation. Spirometry results were obtained using race-specific vs race-neutral reference equations and are shown for White patients and Black patients. A, Percentage of patients with an abnormally low FEV₁ (defined as $z < -1.645$). B, Percentage of patients with an obstructive defect (defined as FEV₁ to FVC ratio, $z < -1.645$). C, Percentage of patients with a potential restrictive defect (FVC z score < -1.645 with preserved FEV₁ to FVC ratio). The figure depicts absolute proportions obtained using the different approaches, so no error bars are warranted.

Canadian with race-specific GLI reference values among Black patients where FEV₁ was more likely to be considered normal using race-specific GLI reference values. The report from Kitazawa et al¹¹ primarily focuses on comparing race-specific GLI and GLI-O with Canadian reference equations. Our study increased generalizability to PFTs performed in the United States by focusing more on a comparison of currently used race-specific GLI reference equations with GLI-O reference equations.

Our study has both strengths and limitations. Strengths include a large number of tests from a clinical database of patients referred for PFTs. Additionally, this study used GLI-O reference equations that have been proposed as a race-composite alternative to race-specific reference equations. It seems likely that our results will reflect real-world experience in other centers in the United States, although the percent agreement will vary based on patient demographic features and prevalence of specific disease states. Limitations include that this was a single-center retrospective study incorporating a relatively small proportion of Black patients, the cohort was heterogeneous with respect to diagnoses, GLI-O is limited in its representation of all racial and ethnic groups, and clinical outcomes data currently are not available for the current patients. Additionally, recent publications have proposed use of GLI Global reference equations rather than GLI-O equations. Because the GLI Global reference equations were

released very recently, clear consensus on the best equation to use in race-neutral PFT results interpretation does not yet exist. However, reinterpretation of PFT results using GLI Global reference equations may be beneficial in future research. Finally, this study was limited to evaluation of the impact of change in reference equations on White and Black populations only because of limited data available from patients of other racial and ethnic backgrounds. To gain a more comprehensive understanding of the impact of transition to race-neutral reference equations, additional analyses incorporating patient data for individuals from a variety of racial and ethnic backgrounds and clarifying clinical characteristics of patients will be needed.

We hope the findings from our study will help clinicians understand the implications of using race-specific vs race-composite interpretation strategies for PFT results. Classification of PFT results comprises only one portion of clinical decision-making and importantly must be incorporated into a broader understanding of an individual patient's history, symptoms, and other diagnostic testing. Still, reclassification of PFT results is likely to impact clinical decision-making including recommendations for disability or activity and job restrictions, additional diagnostic testing, and medication prescription. More research focused on the impact of race-neutral approaches on the usefulness of PFT findings in the diagnosis and management of lung diseases is urgently

needed. This research will be crucial in identifying the best PFT interpretation strategy with the most important clinical goal of accurately identifying patients impacted by abnormal lung function.

Margaret J. Connolly, MD

Patrick A. Donohue, MD

Rohith Palli, PhD

Sandhya Khurana, MD

Xueya Cai, PhD

Steve N. Georas, MD

Rochester, NY

AFFILIATIONS: From the Division of Pulmonary and Critical Care Medicine (M. J. C., P. A. D., S. K., and S. N. G.), the Internal Medicine Residency Program (R. P.), and the Department of Biostatistics and Computational Biology (X. C.), University of Rochester School of Medicine and Dentistry.

Drs Khurana and Georas contributed equally to this manuscript.

This article was presented at the American Thoracic Society 2022 International Conference, San Francisco, California, May 18, 2022.

CORRESPONDENCE TO: Margaret J. Connolly, MD; email: maconnolly@som.umaryland.edu

Copyright © 2023 The Author(s). Published by Elsevier Inc under license from the American College of Chest Physicians. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

DOI: <https://doi.org/10.1016/j.chest.2023.06.011>

Funding/Support

The project described was supported by Multidisciplinary Training in Pulmonary Research [Award Number T32 HL066988], and the University of Rochester Medical Center Department of Medicine through a Pilot Award Program.

Financial/Nonfinancial Disclosures

None declared.

Acknowledgments

Other contributions: The authors thank Michael Clark, BS, Morgan Scientific, for his assistance with database management.

Role of sponsors: The sponsor had no role in the design of the study, the collection and analysis of the data, or the preparation of the manuscript.

References

1. Pellegrino R, Viegi G, Brusasco V, et al. Interpretative strategies for lung function tests. *Eur Respir J*. 2005;26(5):948-968.
2. Hegewald MJ, Crapo RO. Socioeconomic status and lung function. *Chest*. 2007;132(5):1608-1614.
3. Lum S, Bountziouka V, Sonnappa S, et al. Lung function in children in relation to ethnicity, physique and socioeconomic factors. *Eur Respir J*. 2015;46(6):1662-1671.
4. Braun L, Wolfgang M, Dickersin K. Defining race/ethnicity and explaining difference in research studies on lung function. *Eur Respir J*. 2013;41(6):1362-1370.
5. Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight—reconsidering the use of race correction in clinical algorithms. *N Engl J Med*. 2020;383(9):874-882.
6. Townsend MC, Cowl CT. US occupational historical perspective on race and lung function. *Am J Respir Crit Care Med*. 2022;206(6):789-790.
7. Bhakta NR, Kaminsky DA, Bime C, et al. Addressing race in pulmonary function testing by aligning intent and evidence with practice and perception. *Chest*. 2022;161(1):288-297.
8. Quanjer PH, Stanojevic S, Cole TJ, et al. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J*. 2012;40(6):1324-1343.
9. McCormack MC, Balasubramanian A, Matsui EC, et al. Race, lung function, and long-term mortality in the National Health and Nutrition Examination Survey III. *Am J Respir Crit Care Med*. 2022;205(6):723-724.
10. Baugh AD, Shiboski S, Hansel NN, et al. Reconsidering the utility of race-specific lung function prediction equations. *Am J Respir Crit Care Med*. 2022;205(7):819-829.
11. Kitazawa H, Jiang A, Nohra C, et al. Changes in interpretation of spirometry by implementing the GLI 2012 reference equations: impact on patients tested in a hospital-based PFT lab in a large metropolitan city. *BMJ Open Respir Res*. 2022;9(1):e001389.