

Original Manuscript

Race-adjusted Lung Function Increases Inequities in Diagnosis and Prognosis and Should Be Abandoned

Magnus Ekström, MD, PhD¹; David Mannino, MD^{2,3}

1. Lund University, Faculty of Medicine, Department of Clinical Sciences Lund, Respiratory Medicine and Allergology, Lund, Sweden
2. University of Kentucky College of Medicine, Department of Medicine, Lexington, Kentucky, USA
3. COPD Foundation, Washington, D.C., USA

Corresponding author: Magnus Ekström, Department of Medicine, Blekinge Hospital, SE-37185, Karlskrona, Sweden. Telephone: +46(0)455731000. Email: pmekstrom@gmail.com

Word count for body of text: 2,994 words

Word count for abstract: 287

Number of Tables: 2; **Figures:** 1; **References:** 28.

Word count for Abstract: 287

Abstract

Background: Lung function assessment is essential for respiratory medicine and health. Recommended international reference values differ by race, which is controversial. We evaluated the effect of adjusting lung function for race on prevalence of lung function impairment, breathlessness and mortality in the US population.

Methods: Population-based analysis of the National Health and Nutrition Examination Survey (NHANES) 2007–2012. Race was analyzed as black, white, or other. Lung function was assessed as forced expired volume in one second (FEV_1) and forced vital capacity (FVC). Predicted normal values were calculated for each person using the Global Lung Initiative (GLI)-2012 equations for 1) white; 2) black; and 3) other/mixed populations. Outcomes were compared for the different reference values in relation to: prevalence of lung function impairment ($<$ lower limit of normal [LLN]), moderate/severe impairment ($<50\%$ pred); self-reported exertional breathlessness; and mortality up to 31 December, 2015.

Findings: We studied 14,123 people (50% female); white ($n=5,928$), black ($n=3,130$), and other ($n=5,065$). Compared to those for white, black reference values identified markedly fewer cases of lung function impairment (FEV_1) both in black people (9.3% vs. 36.9%) and other non-white races (1.5% vs. 9.5%); and prevalence of moderate/severe impairment was approximately halved. Outcomes among those impaired differed by reference value used: white (best outcomes), other/mixed (intermediate), and black (worst outcomes). Black people with $FEV_1 \geq LLN_{black}$ but $< LLN_{white}$ had 48% increased rate of breathlessness and almost doubled mortality, compared to blacks $\geq LLN_{white}$. Lung function $\geq LLN_{white}$ identified people with good outcomes, similarly in black and white people. Findings were similar when analyzing FEV_1 or FVC.

Interpretation: Race adjustment of lung function should be abandoned. White reference values are most sensitive and specific to identify impairment, and could be applied across the population for improved assessment and health equity.

Funding: Swedish Research Council (Dnr: 2019-02081).

Research in context

Evidence before this study

We searched MEDLINE and Embase using search terms including “race”, “ethnicity”, “pulmonary function”, “spirometry”, and “prediction equations” from database inception and January 10, 2022, for papers published in English. A total 33 papers related to lung function and race were identified. Race-adjusted lung function reference values were recommended by major guidelines for use internationally. Race-specific references assume a 10-15% lower lung function, such as the forced expired volume in one second (FEV₁) and forced vital capacity (FVC), in black people and 4-6% lower in Asian people compared with in whites. Compared to not adjusting for race, race-adjusted lung function values have recently been questioned as they have been found to not improve prediction of outcomes in population-based studies or in people at risk of obstructive pulmonary disease. Concerns have been raised that, contrary to the intent, race-adjusted reference values may contribute to under diagnosis of disease in disadvantage minorities, with the largest differences reported in black (Afro-American) people, and may worsen race-related health inequalities. Data on the impact of race-adjusted lung function values across the ethnically diverse population are limited and data on how to decrease racial bias in lung function assessment are needed.

Added value of this study

We analyzed the impact of using different race-specific (GLI-2012) reference equations for FEV₁ and FVC across the US population in the National Health and Nutrition Examination Survey (NHANES) 2007-2012. Outcomes were prevalence of lung function impairment (value < lower limit of normal), breathlessness on exertion, and mortality up to December 31, 2015. Compared to using references for whites, black reference values were less likely to identify lung function impairment across all races but especially in blacks (9.3% vs. 36.9%); and those identified had lower lung function, more breathlessness, and worse prognosis. Black people with lung function normal by black standards but impaired by white standards had increased prevalence of breathlessness and mortality, compared to those normal also by white standards. Thus, race-adjusted reference values labeled black people as normal despite worse outcomes. White normal values identified people with similarly good lung function, and low rates of breathlessness and mortality across races groups.

Implications of all the available evidence

The findings from this study support that race-adjusted reference values markedly under diagnose lung function impairment, and related breathlessness, and mortality in underprivileged groups across the US population. Normal values for whites were most sensitive to identify lung function impairment related to worsening outcomes and people classified as having normal lung function with similar good outcomes irrespective of race group. These findings suggest that lung function should not be adjusted for race. When applied across the population, white reference values were most sensitive to identify smaller or earlier impairment and most specific to identify people with normal lung function with similarly good outcomes across race groups. Given the large impact shown, abandoning the use of race-adjusted lung function values is likely to contribute to improved health equity.

Introduction

Chronic respiratory disease is the third leading cause of death worldwide.¹ Pulmonary function testing using spirometry is key to the diagnosis of chronic respiratory disease, evaluation of breathlessness, whether people qualify for interventions such as lung transplant, or can be considered to be disabled.² Standards exist for both the performance³ and interpretation⁴ of spirometry. The American Thoracic Society (ATS) recommends that ‘laboratories must select appropriate reference values for the patients being tested’⁴ and goes on to recommend use of the Global Lung Initiative (GLI)-2012 prediction equations,⁵ which establishes race-specific reference values for whites, African Americans, North East Asians, and South East Asians. Currently, race-specific reference values for lung function are the recommended standard for use internationally.⁶⁻⁸

How do race-specific reference values for lung function work? The GLI-2012 prediction equations for normal lung function adjust for age, height, sex, and race.³ While historic prediction equations would apply an ‘adjustment factor’ of 0.88 (12% less) for black populations and 0.94 (6% less) for Asian populations,⁸ the GLI-2012 equations were developed without a fixed adjustment factor but rather using race-specific populations. However, even in the GLI-2012 equations, predicted lung function levels are 10-15% lower in African Americans and South East Asians relative to whites and North East Asians.²

Race adjustment is controversial. On one side of the argument is the thought that race is a surrogate measure that captures a number of factors predictive of poor health status and outcomes that are not really specific to a person’s racial make-up.⁹ The other side of the argument is that there are physiologic traits between populations that are based in genetics and captured, to some extent, by self-reported race.¹⁰ This is analogous to the differences between men and women in predicted lung function, even with identical heights.¹¹ In other area of medicine, race-specific normal values have recently been shown to discriminate and contribute to under-diagnosing and under-treatment in socioeconomically more vulnerable groups such as Afro-Americans and are currently revised not to adjust for race, such as normal values for renal function.¹²

We aimed to determine the effect of adjusting lung function for race on prevalence of lung function impairment, breathlessness and mortality in the US population.

Methods

Design and population

This was a population-based analysis of the National Health and Nutrition Examination Survey (NHANES) in the US from 2007 to 2012.¹³⁻¹⁵ We included all people aged ≥ 18 years with data on demographics and spirometry. Data were obtained on breathlessness on exertion (available for people aged ≥ 40 years), and mortality up to 31 December 2015. Participants provided written consent to participate in NHANES using a protocol approved by the National Center for Health Statistics Research Ethics Review Board.¹³⁻¹⁵ The present analysis used only de-identified NHANES data which are publicly available. The study is reported in accordance with STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) guidelines.¹⁶

Assessments

Data on age, sex, and race were from personal interviews. Race was categorized as Non-Hispanic whites (whites), Non-Hispanic black (blacks), and others (Hispanic, Asian, mixed race, etc.). Analysis focused on comparing black vs. white as these reference values differ the most.⁵ The category other was included to reflect the entire NHANES population.

Measured weight (kg), height (cm), and spirometry were obtained using mobile examination centres. Dynamic spirometry was performed in accordance with guidelines from the ATS and European Respiratory Society (ERS).¹⁷ Values were recorded as the highest obtained value (pre- or post-bronchodilator). For each participant, reference values for the forced expired volume in one second (FEV_1) and forced vital capacity (FVC) were calculated using the GLL-2012 equations for 1) white, 2) black, and 3) other/mixed populations.⁵ Thus, for each individual we calculated three predicted reference values for FEV_1 and FVC, respectively, for comparison.

Breathlessness on exertion was assessed using the question: 'Have you had shortness of breath either when hurrying on the level or walking up a slight hill?' (yes/no), corresponding to a breathlessness level of ≥ 1 point on the modified Medical Research Council (mMRC) scale.¹⁸ Breathlessness data were available in NHANES for people 40 years or older. Mortality was assessed using standardized NHANES procedures up to 31 December 2015.

Statistical analyses

The study population was weighted (using published NHANES weights for people undergoing examinations including spirometry), to represent the non-institutionalized US population during

the six year period. For all analyses, variance estimates were produced using Taylor Series Linearization methods,¹⁹ as recommended for NHANES.

Data were tabulated and compared between race groups using means (standard deviation [SD]) for normally distributed continuous variables, and frequency (percentage) for categorical variables. Lung function was evaluated as FEV₁ in the main analyses. Similar analyses of FVC are reported in the supplements.

Outcomes were compared between race groups (white, black, and other) in terms of: 1) predicted normal values using each race-specific prediction equation (white, black, or other/mixed); 2) prevalence of impaired lung function, defined as value < the lower limit of normal (LLN) using each race-specific prediction equation; and the prevalence of moderate to severe impairment, defined as < 50% of the predicted normal in accordance with GOLD (Global Initiative for Obstructive Lung Disease);⁶ 3) prevalence of breathlessness; and 4) mortality. As each prediction equation was applied to the same people, the effects of applying different race-specific reference values were independent of (adjusted for) participant characteristics by design.

Breathlessness and mortality were compared by race and lung function impairment (defined using different race-specific prediction equations) using five mutually exclusive categories: 'White Normal' (white race, value \geq LLN_{white}); 'White Abnormal' (white race, value < LLN_{white}); 'Black Normal' (white race, value \geq LLN_{white}); 'Black Abnormal (White Standard)' (black race, value < LLN_{white} but \geq LLN_{black}); or 'Black Abnormal (Black Standard)' (black race and value < LLN_{white} and < LLN_{black}). As normal values were higher for all persons using white than black prediction equations, all values < LLN_{black} were also < LLN_{white}. Probability of breathlessness was analyzed using logistic regression and was expressed as relative rate ratios (RRR). Mortality was analyzed using Cox proportional-hazards regression and expressed as hazard ratios (HR). Associations with breathlessness and mortality were also analyzed for lung function impairment using each race-specific prediction equation in the whole population, adjusting for age, sex and body mass index (BMI). All estimates were reported with 95% confidence intervals (CIs). Statistical analyses were performed with Stata version 16.0 (StataCorp LP; College Station, TX).

Results

A total 14,123 people (50% female) were studied with race self-reported as white (n=5,928), black (n=3,130), or other (n=5,065). Compared with the other groups, white people were slightly older and had somewhat higher absolute FEV₁ and FVC values, whereas the sex distribution and BMI was similar between race groups (Table 1). Breathlessness was more prevalent among black people, as compared with white and other (Table 1; p < 0.001 for both comparisons).

Prevalence of lung function impairment

The predicted normal FEV₁ was highest using the equation for whites, intermediate using that for other/mixed, and lowest when using the equation for black people (Table 2). This pattern was similar in all race groups. In black people the predicted normal FEV₁ dropped from 3.5 liter using the white equation to 3.0 liter using the black equation, a decrease by 14%.

The choice of race adjustment strongly influenced the prevalence of lung function impairment (Table 2). Compared with the equation for white people, black reference values identified markedly fewer cases of impaired lung function, both among whites (2.4% vs. 8.5%), in black people who had the largest absolute decrease (9.3% vs. 36.9%), as well as in other races (1.5% vs. 9.5%). Overall, compared with the reference values for blacks, white reference values identified about four times as many people as having impaired lung function, and identification of moderate to severe impairment (< 50% predicted) was approximately doubled (Table 2). All findings were similar when analyzing FVC instead of FEV₁ (supplemental Table S1).

Lung function and outcomes

People with impaired lung function (FEV₁<LLN) had increased rates of breathlessness and mortality across the whole population, but outcomes differed by the race-specific reference value used: black (worst outcomes), other/mixed (intermediate), and white (best outcomes). Breathlessness associations using each race-specific reference value were: for black (RRR 4.6; 95% CI, 3.2 – 6.6), other/mixed (RRR 3.4; 95% CI, 2.7 – 4.4), and white (RRR 2.8; 95% CI, 2.4 – 3.3). Mortality estimates were: for black (HR 3.5; 95% CI, 2.4 – 5.2), other/mixed (HR 2.8; 95% CI, 2.1 – 3.6), and white (HR 2.6; 95% CI, 2.1 – 3.4).

Outcomes by race (black or white) and lung function impairment defined using the different race-specific reference values are shown in Figure 1. People with normal lung function according to white standards (\geq LLN_{white}) had low rates of breathlessness and mortality which were similar in both white and black people. In contrast, black reference values introduced a race-related bias; black people with lung function normal by the black standard but impaired by the standard for whites (FEV₁ \geq LLN_{black} but < LLN_{white}) had significantly increased rates of breathlessness (RRR 1.48; 95% CI, 1.13–1.94) and mortality (HR 1.87; 95% CI, 1.42–2.46), compared to black people who were also normal by white standards (FEV₁ \geq LLN_{white}). Associations from Cox regression are shown in supplemental Table S2.

Findings were similar when analyzing FVC (Table S3 and Figure S1 in the supplement).

Discussion

The main findings are that, compared with using reference values (GLI-2012)⁵ for white people, race-specific equations for black or mixed/other populations markedly under diagnose lung function impairment, including moderate to severe disease, and misclassify black people as having normal lung function despite having substantially increased rates of breathlessness and mortality. In black people, race adjusted (black) reference values identified only ¼ of cases (9.3% vs. 36.9%) of impaired FEV₁ as compared with white reference values – and those identified by the black references had significantly worse lung function, more breathlessness and increased mortality. Black people with normal lung function according to black reference values but impaired according to those for whites had 48% increased rate of breathlessness and almost doubled mortality, as compared with black people with normal FEV₁ by white standards. This racial bias was largely avoided by applying the white reference values across the whole population; besides identifying the higher prevalence of lung function impairment with worse outcomes in black, people with normal lung function according to white reference values had similar good prognosis across all the race groups.

These findings have important implications for assessment of lung function and respiratory disease. Firstly, use of race-adjusted reference values, which are currently endorsed by major international guidelines,^{6, 7} should be abandoned. As we show, race-adjusted references (compared to white reference values) misclassify lung function as normal despite worse outcomes in as many as 28% of black people, corresponding to as many as 13.1 million people in the US alone.²⁰ A particularly alarming finding was that half of cases of moderate to severe lung function impairment were missed – which could lead to insufficient treatment or delayed interventions such as lung transplantation evaluation. Adjusting lung function values for race thus contribute to under diagnosis of disease and disability, failure to identify impaired lung function as contributing to breathlessness, potentially leading to insufficient or delayed treatment and compounded race-related health inequities. Second, to avoid this race-related bias, the present findings support that lung function should be assessed using a common prediction equation across the population. The choice of prediction equation may vary depending on aim of the assessment. In research and clinical medicine, the aims of spirometry are mainly to evaluate whether lung function is normal or impaired, whether breathlessness is related to impaired lung function, evaluate the severity of respiratory disease needing treatment or further evaluation, and to predict prognosis. With emphasis on sensitivity to detect and treat lung function impairment earlier, the present findings support the use of the GLI-2012 prediction equations for white people across the population. We found white references to be more sensitive than the GLI-2012 mixed/other reference values, which were

previously proposed for use across mixed populations.⁵ Importantly, while being more sensitive, white reference predictions were still specific to identify impairment associated with substantial morbidity in terms of increased breathlessness and mortality rates, which was seen across all the race groups. White reference values identified smaller or earlier lung function impairment, which should be appropriately evaluated, and may be more amenable to treatment.

The present findings extend previous reports that race-related differences in mortality were attenuated by applying using the same prediction equation (reference values for whites) across the population.²¹ The findings are in line with those that using race-adjusted reference values (compared to not adjusting for race) did not improve prediction of respiratory morbidity or mortality in a large cohort study,²² and predicted clinically important outcomes worse in black and white people at high risk of COPD (n=2,652).²³ Taken together, these previous data support our findings that using race-specific reference values, rather than counteract, in fact widen race-related inequities.

Lower lung function in black people is consistent with data that airflow obstruction and reduced lung function strongly associates with poverty at individual and community levels across multiple countries, independent of factors such as age, sex, and smoking and tuberculosis.²⁴ In the recent study by Baugh *et al.*, controlling for comorbid disease and measures of adversity weakened the association between race and FEV₁, suggesting that racial differences in lung function are at least partly reflect differential exposures.²³ It is increasingly acknowledged that race is, to a large part, a social construct.² Genetic and environmental factors inseparably interact in multiple and complex ways to influence all aspects of life, through prenatal and early life factors, circumstances throughout life, over the generations.² In the Eight America's project, Murray *et al.* described large disparities in mortality across race-county groupings and concluded that these differences could not be explained by race, income, or basic health-care access and utilization alone.²⁵ As pointed out,² the lower lung function in disadvantaged groups including Afro-Americans might, to an extent, reflect a higher accumulated exposure to adverse exposures and not disease. But as we show, race-adjusted lung function values may obscure the higher prevalence of impairment in these populations, misclassify people as healthy despite having worse outcomes, and contribute to under diagnosis of disease or presence of modifiable health exposures that could, when appreciated, be modified.² An example of the influence of environmental factors from another field is the generational change in health outcomes when comparing the population of southern Europe to northern Europe.²⁶ Even though these populations had similar race distributions, large differences in both adult height and childhood mortality seen in 1950 had largely disappeared by 1980.²⁶

The suggestion that race-adjustment should be abandoned in lung function assessment is consistent with similar developments to counteract racial bias in other medical areas, including tests in haematology, and references for kidney function.^{12, 27, 28}

Strengths of the present study include the use of a well characterized, large database representative for the racially diverse non-institutionalized US population. Race-specific prediction equations for normal lung function (FEV₁ and FVC) were evaluated using the international GLI-2012 reference values developed to be applicable globally, in accordance with guidelines.⁵⁻⁷ By comparing the predictions in the same population, the analyses were independent of differences in participant characteristics. Reference values were evaluated against clinically important outcomes in terms of prevalence of impairment, breathlessness and mortality.

A limitation of the present study is that data pertain to the US population, and studies in other settings are needed. However, we show how currently recommended race-specific prediction equations for lung function perform, introduce race-related bias, and now this bias may be attenuated or overcome by applying a similar reference across a large diverse population. The findings are likely to be valid and relevant for avoiding race-related bias in lung function assessment in many settings around the world.

In conclusion, reference values for lung function should not be adjusted for race, as it leads to substantial under diagnosis of impairment and misclassification of health status and outcomes in under privileged groups. Lung function should be evaluated using a common prediction equation across races, where reference values for whites are most sensitive to detect smaller and earlier impairment, which could promote improved management and health equity.

Acknowledgements

Declaration of interests: ME discloses personal fees from Astra Zeneca outside this study. DM is a former employee and current shareholder of GlaxoSmithKline and a consultant to AstraZeneca and the COPD Foundation, and is also an expert witness for the Schlesinger Law Firm.

Contributors statement: Both authors conceptualized the study design; ME performed the analyses, and wrote the first draft. Both authors verified the data and analyses, and were involved in the interpretation, revision and decision to submit the final manuscript.

Funding: ME was supported by an unrestricted grant from the Swedish Research Council (Dnr: 2019-02081).

Role of the funding sources: The funding sources had no influence of the planning, conduct or reporting of the study.

Data sharing statement: All data used are anonymized and publicly available, and scripts (for Stata) of all analyses will be made available from the authors upon reasonable request.

References

1. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015; **385**(9963): 117-71.
2. Bhakta NR, Kaminsky DA, Bime C, et al. Addressing Race in Pulmonary Function Testing by Aligning Intent and Evidence With Practice and Perception. *Chest* 2021.
3. Graham BL, Steenbruggen I, Miller MR, et al. Standardization of Spirometry 2019 Update. An Official American Thoracic Society and European Respiratory Society Technical Statement. *Am J Respir Crit Care Med* 2019; **200**(8): e70-e88.
4. Culver BH, Graham BL, Coates AL, et al. Recommendations for a Standardized Pulmonary Function Report. An Official American Thoracic Society Technical Statement. *Am J Respir Crit Care Med* 2017; **196**(11): 1463-72.
5. 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): 1324-43.
6. From the Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2021. Available from: <http://goldcopd.org/> [Accessed 29-07-2021].
7. Graham BL, Steenbruggen I, Miller MR, et al. Standardization of Spirometry 2019 Update. An Official American Thoracic Society and European Respiratory Society Technical Statement. *Am J Respir Crit Care Med* 2019; **200**(8): e70-e88.
8. Pellegrino R, Viegi G, Brusasco V, et al. Interpretative strategies for lung function tests. *Eur Respir J* 2005; **26**(5): 948-68.
9. Braun L. Race Correction and Spirometry: Why History Matters. *Chest* 2021; **159**(4): 1670-5.
10. Scanlon PD, Shriver MD. "Race correction" in pulmonary-function testing. *N Engl J Med* 2010; **363**(4): 385-6.
11. Miller MR. Black and white lungs: race and the history of spirometry. *Lancet* 2014; **2**: 2.
12. Diao JA, Wu GJ, Taylor HA, et al. Clinical Implications of Removing Race From Estimates of Kidney Function. *JAMA* 2021; **325**(2): 184-6.
13. National Center for Health Statistics. NHANES 2007-2008 Procedure Manuals. 2007. <https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/manuals.aspx?BeginYear=2007> (accessed January 9 2022).
14. National Center for Health Statistics. NHANES 2009-2010. 2009. <https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2009> (accessed January 10 2022).
15. National Center for Health Statistics. NHANES 2011-2012. 2011. <https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2011> (accessed January 10 2022).
16. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007; **335**(7624): 806-8.
17. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J* 2005; **26**(2): 319-38.
18. Bestall JC, Paul EA, Garrod R, Garnham R, Jones PW, Wedzicha JA. Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax* 1999; **54**(7): 581-6.
19. Encyclopedia of Survey Research Methods. 2008.

20. Pew Research Center. The Growing Diversity of Black America. 2021. <https://www.pewresearch.org/social-trends/2021/03/25/the-growing-diversity-of-black-america/> (accessed January 9 2022).
21. Gaffney AW, McCormick D, Woolhandler S, Christiani DC, Himmelstein DU. Prognostic implications of differences in forced vital capacity in black and white US adults: Findings from NHANES III with long-term mortality follow-up. *EClinicalMedicine* 2021; **39**.
22. Elmaleh-Sachs A, Balte P, Oelsner EC, Allen NB, Baugh A, Bertoni AG, Hankinson JL, Pankow J, Post WS, Schwartz JE, Smith BM, Watson K, Barr RG. Race/Ethnicity, Spirometry Reference Equations and Prediction of Incident Clinical Events: The Multi-Ethnic Study of Atherosclerosis (MESA) Lung Study. *AJRCCM* 2021. *In press*.
23. Baugh AD, Shiboski S, Hansel NN, et al. Reconsidering the Utility of Race-Specific Lung Function Prediction Equations. *Am J Respir Crit Care Med* 2021.
24. Townend J, Minelli C, Mortimer K, et al. The association between chronic airflow obstruction and poverty in 12 sites of the multinational BOLD study. *Eur Respir J* 2017; **49**(6).
25. Murray CJ, Kulkarni SC, Michaud C, et al. Eight Americas: investigating mortality disparities across races, counties, and race-counties in the United States. *PLoS Med* 2006; **3**(9): e260.
26. Bozzoli C, Deaton A, Quintana-Domeque C. Adult height and childhood disease. *Demography* 2009; **46**(4): 647-69.
27. Inker LA, Eneanya ND, Coresh J, et al. New Creatinine- and Cystatin C–Based Equations to Estimate GFR without Race. *N Engl J Med* 2021; **385**(19): 1737-49.
28. Weyand AC, McGann PT. Eliminating race-based reference ranges in haematology: a call to action. *The Lancet Haematology* 2021; **8**(6): e462-e6.

Tables

Table 1. Characteristics and lung function by race

Factor	White people	Black people	Other
N	5,928	3,130	5,065
Age, mean (SD)	45.7 (16.0)	42.1 (15.8)	39.5 (14.5)
Female, %	50.3%	53.7%	48.3%
Weight (kg), mean (SD)	83.0 (20.7)	87.9 (24.0)	76.7 (19.2)
Height (cm), mean (SD)	170.7 (9.8)	169.6 (9.5)	165.1 (9.6)
Body mass index (kg/m ²), mean (SD)	28.4 (6.5)	30.6 (8.1)	28.1 (6.2)
FEV ₁ (L), mean (SD)	3.3 (0.9)	2.9 (0.8)	3.2 (0.8)
FVC (L), mean (SD)	4.3 (1.1)	3.6 (1.0)	4.0 (1.0)
FEV ₁ /FVC, mean (SD)	0.77 (0.08)	0.80 (0.08)	0.81 (0.07)
Breathlessness prevalence*, %	29.3%	33.7%	24.4%
Deaths by December 31, 2015, %	3.0%	3.5%	2.2%

Characteristics are weighted to represent the mean US population in National Health and Nutrition Examination Survey (NHANES) 2007-2012. * Breathlessness prevalence is in people aged 40 years or older. Abbreviations: FEV₁ = forced expired volume in one second. FVC = forced vital capacity.

Table 2. Prevalence of impaired FEV₁ by race and race-specific reference value used

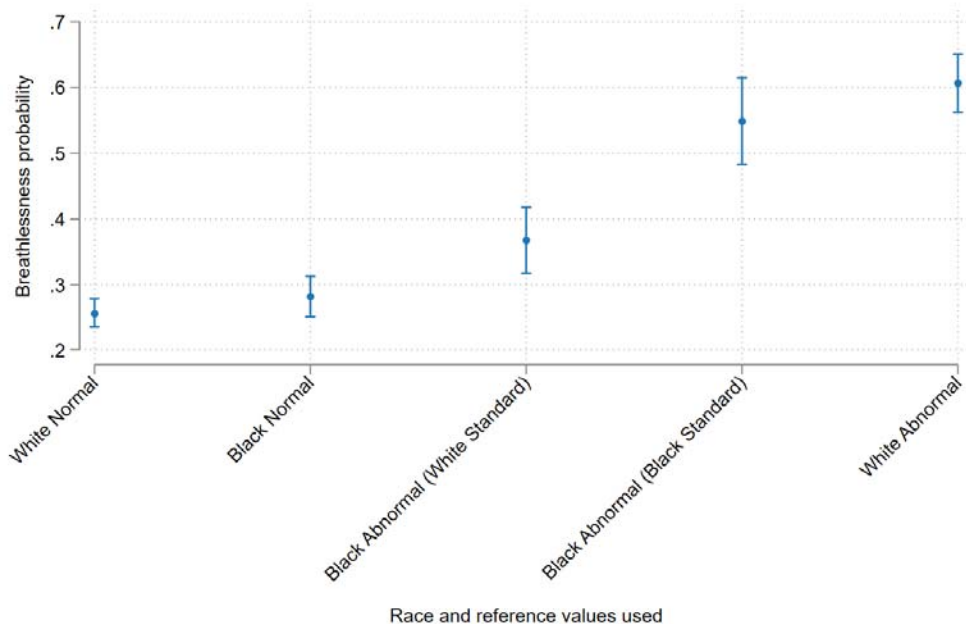
Factor	White people	Black people	Other
N	5,928	3,130	5,065
Predicted normal FEV₁, mean (SD)			
White reference values	3.5 (0.81)	3.5 (0.79)	3.4 (0.75)
Other/mixed reference values	3.2 (0.76)	3.2 (0.73)	3.1 (0.70)
Black reference values	3.0 (0.68)	3.0 (0.66)	2.9 (0.63)
Prevalence of impaired FEV₁ (<LLN) using, %			
White reference values	8.5%	36.9%	9.5%
Other/mixed reference values	5.1%	21.3%	4.4%
Black reference values	2.4%	9.3%	1.5%
Prevalence of moderate/severe FEV₁ impairment (<50%pred) using, %			
White reference values	0.8%	1.7%	0.5%
Other/mixed reference values	0.5%	1.1%	0.3%
Black reference values	0.4%	0.8%	0.2%

Reference values by GLI-2012.⁵ For abbreviations, see Table 1.

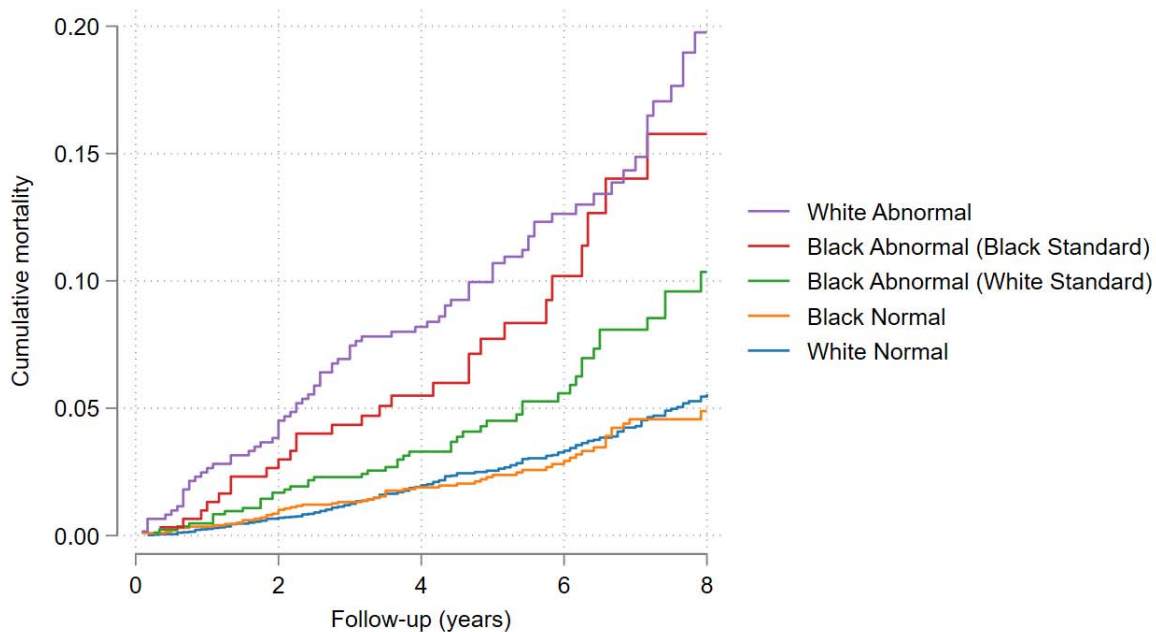
Figure legends

Figure 1. Outcomes by race and FEV₁ impairment defined using white and/or black normal values, in terms of a) breathlessness, and b) mortality. Breathlessness probability was analyzed using logistic regression, and mortality using Cox proportional hazards regression. Impaired lung function was defined as a forced expired volume in one second (FEV₁) < lower limit of normal (LLN) using GLLI-2012 predicted normal values for white and black people, respectively.⁵ Groups were categorized by race and FEV₁ impairment according to different race-specific prediction equations as: ‘White Normal’ (white race with FEV₁ ≥ LLN_{white}); ‘Black Normal’ (black race with FEV₁ ≥ LLN_{white}); ‘Black Abnormal (White Reference)’ (black race with FEV₁ < LLN_{white} but ≥ LLN_{black}); ‘Black Abnormal (Black Reference)’ (black race and FEV₁ < LLN_{white} and < LLN_{black}); and ‘White Abnormal’ (white race and FEV₁ < LLN_{white}). The main finding is that black people who were categorized as having a normal FEV₁ using LLN_{black} but not using LLN_{white} had increased breathlessness prevalence and mortality compared with people categorized as normal using reference values for white. Thus, black reference values misclassify black people as having normal lung function despite having worse outcomes. When defining normality using LLN_{white} for all, people with normal FEV₁ had similar breathlessness and mortality in both white and black people.

a) Breathlessness



b) Mortality



Number at risk				
611	588	514	302	118
305	297	205	102	32
836	822	588	312	112
1986	1969	1470	817	283
5309	5274	4729	2928	1051
				White Abnormal
				Black Abnormal (Black Standard)
				Black Abnormal (White Standard)
				Black Normal
				White Normal

Supplementary material

Table S1. Prevalence of impaired FVC by race and race-specific reference values used

Factor	White people	Black people	Other
N	5,928	3,130	5,065
Predicted normal FVC, mean (SD)			
White reference values	4.28 (0.98)	4.26 (0.94)	4.10 (0.89)
Other/mixed reference values	3.94 (0.90)	3.93 (0.87)	3.78 (0.82)
Black reference values	3.63 (0.81)	3.62 (0.78)	3.48 (0.74)
Prevalence of impaired FVC (<LLN) using, %			
White reference values	4.7%	32.5%	8.5%
Other/mixed reference values	2.1%	19.2%	3.5%
Black reference values	0.7%	7.0%	1.5%
Prevalence of moderate/severe FVC impairment (<50%pred) using, %			
White reference values	0.1%	0.7%	0.2%
Other/mixed reference values	0.1%	0.4%	0.1%
Black reference values	<1%	0.2%	<1%

Reference values by GLI-2012.⁵ Abbreviations: FVC = forced vital capacity; LLN = lower limit of normal; pred = predicted normal value.

Table S2. Associations with breathlessness and mortality by race and FEV₁ impairment

Group	Breathlessness RRR (95% CI)		Mortality Hazard ratio (95% CI)	
	Crude	Adjusted*	Crude	Adjusted*
White Normal	1 (ref)	1 (ref)	1 (ref)	1 (ref)
Black Normal	1.14 (0.95–1.37)	1.00 (0.84–1.20)	1.10 (0.83–1.47)	1.36 (1.01–1.83)
Black Abnormal (White Standard)	1.69 (1.36–2.08)	1.55 (1.24–1.94)	2.07 (1.51–2.83)	2.64 (1.88–3.70)
Black Abnormal (Black Standard)	3.52 (2.63–4.71)	2.91 (2.15–3.94)	3.46 (2.30–5.19)	3.27 (2.16–4.95)
White Abnormal	4.46 (3.64–5.46)	4.27 (3.45–5.30)	4.05 (2.84–5.78)	3.08 (2.16–4.38)

Breathlessness data were available and analyzed in people aged 40 years or older. Groups are categorized similar to in Figure 1. *Adjusted for age, sex, and body mass index. Abbreviations: CI = confidence interval; FEV₁ = forced expired volume in one second; RRR = relative rate ratio.

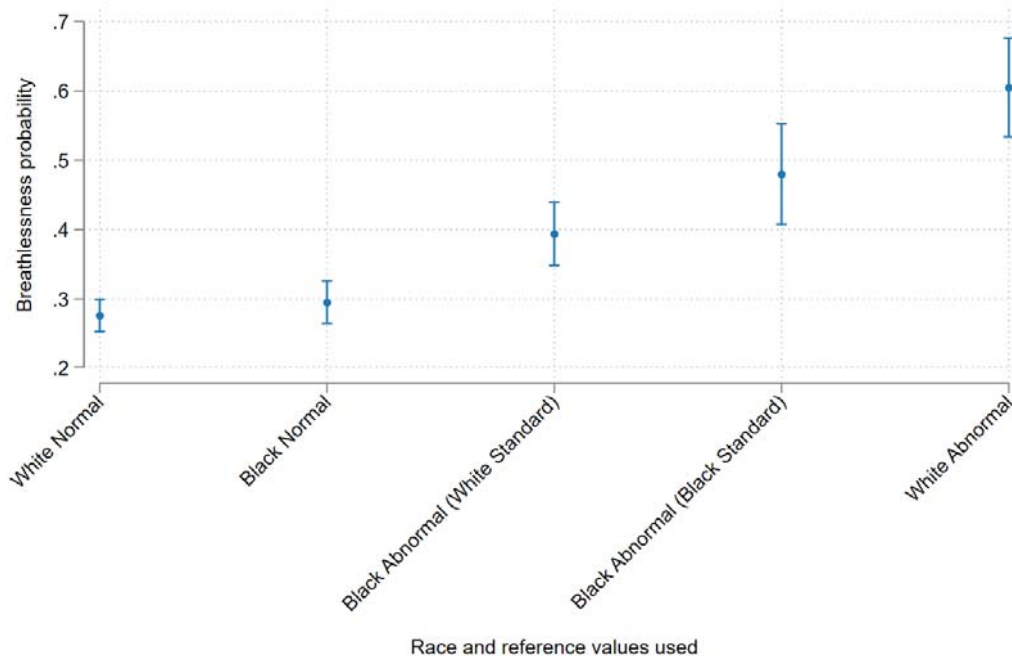
Table S3. Associations with breathlessness and mortality by race and FVC impairment

Group	Breathlessness RRR (95% CI)		Mortality Hazard ratio (95% CI)	
	Crude	Adjusted*	Crude	Adjusted*
White Normal	1 (ref)	1 (ref)	1 (ref)	1 (ref)
Black Normal	1.10 (0.93–1.30)	1.01 (0.85–1.19)	1.04 (0.78–1.40)	1.30 (0.97–1.74)
Black Abnormal (White Standard)	1.71 (1.37–2.13)	1.50 (1.20–1.87)	2.00 (1.39–2.89)	2.65 (1.79–3.91)
Black Abnormal (Black Standard)	2.42 (1.77–3.31)	1.76 (1.27–2.43)	3.48 (2.28–5.32)	3.63 (2.29–5.75)
White Abnormal	4.02 (2.94–5.49)	3.12 (2.18–4.44)	4.22 (2.80–6.36)	3.43 (2.36–4.99)

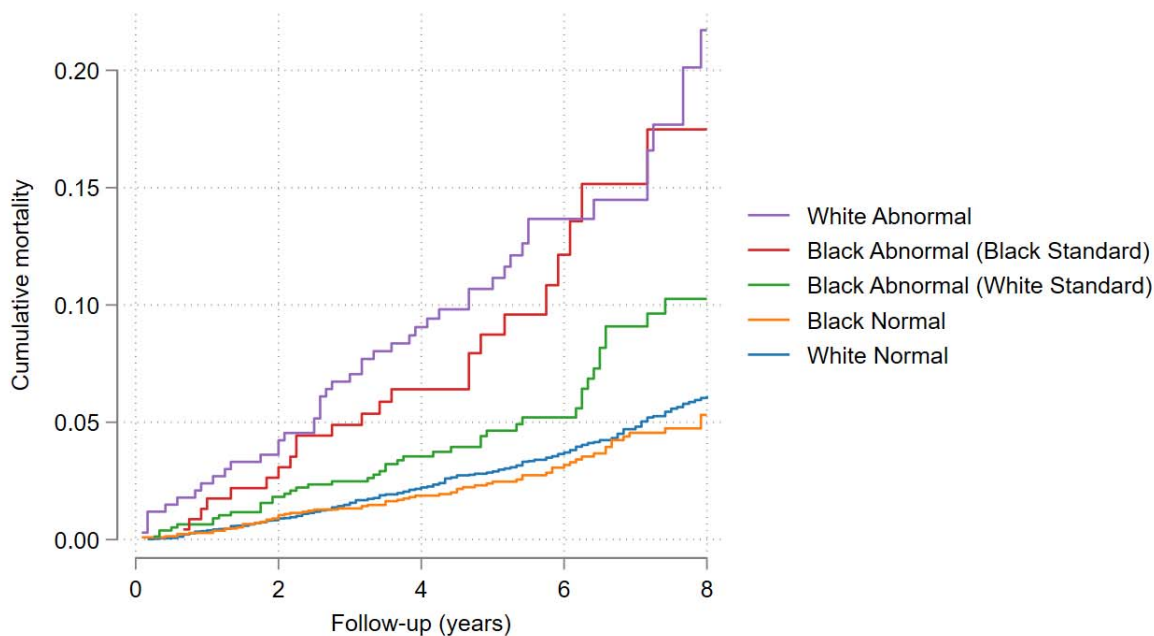
Breathlessness data were available and analyzed in people aged 40 years or older. Groups are categorized similar to in Figure S1. *Adjusted for age, sex, and body mass index. Abbreviations: CI = confidence interval; FVC = forced vital capacity; RRR = relative rate ratio.

Figure S1. Outcomes by race and FVC impairment defined using reference values for white and/or black people, in terms of a) breathlessness, and b) mortality. Breathlessness probability was analyzed using logistic regression, and mortality using Cox proportional hazards regression. Impaired lung function was defined as a forced vital capacity (FVC) < lower limit of normal (LLN) using GLI-2012 predicted normal values for white and black people, respectively.⁵ Groups were categorized by race and FVC impairment according to different race-specific prediction equations as: ‘White Normal’ (white race with $FVC \geq LLN_{white}$); ‘Black Normal’ (black race with $FVC \geq LLN_{white}$); ‘Black Abnormal (White Reference)’ (black race with $FVC < LLN_{white}$ but $\geq LLN_{black}$); ‘Black Abnormal (Black Reference)’ (black race and $FVC < LLN_{white}$ and $< LLN_{black}$); and ‘White Abnormal’ (white race and $FVC < LLN_{white}$). The main finding is that black people who were categorized as having a normal FVC using LLN_{black} but not using LLN_{white} had increased breathlessness prevalence and mortality compared with people categorized as normal using reference values for white. Thus, black reference values misclassify black people as having normal lung function despite having worse outcomes. When defining normality using LLN_{white} for all, people with normal FVC had similar breathlessness and mortality in both white and black people.

a) Breathlessness



b) Mortality



Number at risk				
337	325	282	153	60
230	224	152	73	21
774	760	547	275	91
2123	2104	1564	883	315
5583	5537	4961	3077	1109