Implications of reference equations for interpretation of spirometry in three African countries: a cross-sectional study

Denise Banze[1,2,13], Claire J. Calderwood[3,4,13], Celina Nhamuave[2], Edson T. Marambire[1,4], Alfred Mfinanga[1,5], Leyla Larsson[6], Akanksha Mimi Malhotra[7], Lilian T. Minja[5], Olena Ivanova[6,8], Lindsay Zurba[9], Norbert Heinrich[6,8,10], Rashida A. Ferrand[3,4], Katherine Fielding[11], Katharina Kranzer[3,4,6,8], Andrea Rachow[6,8,12], John R. Hurst[7] and Celso Khosa[1,2] on behalf of the ERASE-TB consortium[14]

[1]CIH[LMU] Center for International Health, University Hospital, LMU Munich, Munich, Germany. [2]Instituto Nacional de Saúde, Marracuene, Mozambique. [3]Department of Clinical Research, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, UK. [4]Biomedical Research and Training Institute, Harare, Zimbabwe. [5]National Institute for Medical Research – Mbeya Medical Research Centre, Mbeya, Tanzania. [6]Institute of Infectious Diseases and Tropical Medicine, LMU University Hospital, LMU Munich, Munich, Germany. [7]UCL Respiratory, Division of Medicine, University College London, London, UK. [8]German Centre for Infection Research (DZIF) partner site Munich, Munich, Germany. [9]Education for Health Africa, Durban, South Africa. [10]Immunology, Infection and Pandemic Research, Fraunhofer Institute for Translational Medicine and Pharmacology, Munich, Germany. [11]Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene ans Tropical Medicine, London, UK. [12]Global Health Unit, Helmholtz Zentrum München, German Research Centre for Environmental Health (HMGU), Neuherberg, Germany. [13]These authors contributed equally. [14]For a list of the ERASE-TB consortium members and their affiliations see the supplementary material.

Corresponding authors: Denise Banze (denise.banze@ins.gov.mz)

Shareable abstract (@ERSpublications)
Choices about reference equations have important implications for clinical and public health decision-making in resource-constrained environments. There is a need for further work to ensure that GLI Global is globally representative. https://bit.ly/3Wpbxff

Cite this article as: Banze D, Calderwood CJ, Nhamuave C, *et al*. Implications of reference equations for interpretation of spirometry in three African countries: a cross-sectional study. *ERJ Open Res* 2025; 11: 00932-2024 [DOI: 10.1183/23120541.00932-2024].

This article has an editorial commentary: https://doi.org/10.1183/23120541.00186-2025

Received: 16 Sept 2024
Accepted: 7 Jan 2025

Abstract

Background The Global Lung Function Initiative (GLI) and American Thoracic Society recently endorsed a race-composite spirometry reference equation ("GLI Global"). Africa (outside North Africa) is not represented in the underlying dataset; GLI Global has not been evaluated in the region. We evaluated the fit and diagnostic implications of GLI and African (identified by scoping review) reference equations in three East/Southern African countries.

Methods Among healthy participants from a tuberculosis household contact cohort study in Mozambique, Tanzania and Zimbabwe (age ⩾10 years) with post-bronchodilator spirometry we calculated forced expiratory volume in 1 s (FEV_1), forced vital capacity (FVC) and FEV_1/FVC z-scores using different equations, the proportion of people with obstructive airways disease or preserved-ratio-impaired spirometry by different equations. We compared these measures across reference equations.

Results In total, 806 healthy people had good-quality post-bronchodilator spirometry. Across GLI equations, "African American" fitted best (mean±SD FEV_1 z-score −0.12±1.20, mean FVC z-score −0.35±1.19). Compared with "African American", GLI Global resulted in twice as many people being identified as having preserved-ratio impaired spirometry (22% *versus* 11%) with a similar proportion having obstruction (4.2% *versus* 3.8%). Reference equations developed in Africa conferred similar fit compared with the GLI African American equation.

Conclusions Reference equations have clinical and public health implications that demand careful consideration, particularly in resource-constrained environments. Use of GLI Global may result more people being identified as having lung function impairment. Further work that includes clinical outcomes is needed to ensure that GLI Global is globally representative. The key limitation of this work is the potential for people with undiagnosed respiratory disease to have been included in the analysis.

Introduction

Spirometry is the most widely used and validated test for diagnosis of chronic lung disease globally [1]. Since publication in 2012, multiple respiratory societies have endorsed the Global Lung Function Initiative (GLI) reference equations. Both GLI and the American Thoracic Society (ATS) have recently advocated a race-composite and/or race-independent approach to interpretation of lung function tests, representing a paradigm shift away from race-specific equations [2–4]. This recommendation is underscored by several key arguments: 1) it emphasises that race is a social construct, primarily influenced by external factors such as appearance and social background, with no inherent biological basis; 2) the use of race-specific normalisation in clinical practice may reinforce health disparities; 3) as much variation in lung function is observed within racial groups as between groups; 4) "normalising" for race fails to account for critical social, structural and environmental determinants of health, which disproportionately impact people of non-white ethnicities [5]; and finally, 5) relying on race-specific equations overlooks individuals with multi-ethnic backgrounds. Recent research from the United States supports the assertion that utilising race-specific equations may lead to underestimation of the true burden of disease among people of non-white ethnicities [6, 7].

At their inception, the GLI equations included a provisional "Other" category, representing a non-weighted average of the four included ethnic groups (white European (termed Caucasian), African American, North and South East Asian). This was proposed for use among individuals of mixed ethnicity, or those not otherwise represented [8]. As 77% of individuals in the underlying dataset were described as white Europeans, this group contributes most to the "Other" reference equations. GLI Global, the reference equation now recommended by ATS [4], was developed through re-analysis of the initial GLI dataset with equal weighting of each included ethnicity (it is thus a race-composite equation; the "race neutral" terminology reflects the perspective of the spirometry technician [9, 10]). Despite the term "Global", the underlying dataset is unchanged and remains heavily based on data from the United States. Importantly, the only data from Africa are 1143 individuals from North Africa (Algeria and Tunisia, accounting for 1.4% of the dataset), who were classified into the "Caucasian" equation. Consequently, the African American group exclusively contains data from Black Americans [8].

Differences in lung function at a population-level are complex and multifactorial: genetic ancestry, maternal health, intra-uterine and early childhood development, and ongoing respiratory exposures throughout the life course all play a role [11]. There is considerable variation in prevalence, and thus impact, of these determinants and their upstream contributors across world regions. Lung function is therefore determined by ubiquitous, complex and enduring structural and environmental factors. It is therefore, strictly taken, not possible to generate a reference equation that is not impacted by any factor that is detrimental to lung development and/or lung health. While a race-independent approach to the interpretation of spirometry of individuals in the same setting seems warranted [6], a one-size-fits-all reference to which much of world's population does not contribute is inadequate. There are important consequences of moving to a race-composite reference equation among people of non-white ethnicities. This includes the potential for over-diagnosis of lung function impairment, consequent over-treatment (where treatment is available) and increased burden on patients and health systems. These impacts are critically important in Africa where capacity for management of chronic lung diseases is extremely limited.

In this study, we evaluate the fit of different spirometry reference equations in a population of tuberculosis household contacts in three African countries, describe the clinical and public health implications of a move to GLI Global.

Methods

Ethical considerations and reporting guidance

Informed written consent was obtained from all participants 18 years and older. Individual assent and guardian consent was obtained from participants aged 10–17 years. Ethical approval was granted by all relevant institutions (additional details in online data supplement). This study is reported in accordance with Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (supplementary materials).

Study population and procedures

This cross-sectional analysis is part of a multicentre cohort study in Mozambique (Maputo), Tanzania (Mbeya) and Zimbabwe (Harare; ERASE-TB). Harare and Maputo, the respective countries' capital cities, are predominantly urban/peri-urban, whereas Mbeya is primarily rural, with most people engaged in small-scale farming. The methodology of the parent study has been described previously; additional details are provided in the supplementary materials [12]. In summary, adults with microbiologically confirmed

pulmonary tuberculosis were enrolled and asked about their household contacts. All people aged at least 10 years old and living in the same household as these people with tuberculosis (defined as sleeping for at least three nights per week in the same house) who were not themselves on tuberculosis treatment were invited to participate. All consenting individuals were assessed and followed up 6-monthly for a maximum of 24 months. Tuberculosis is strongly associated with poverty; therefore, these tuberculosis-exposed individuals are considered a healthy, but socioeconomically deprived population [13].

Pre- and post-bronchodilation spirometry was performed at baseline where possible, or otherwise at follow-up, according to ATS/European Respiratory Society (ERS) guidelines (Easy On-PC; supplementary materials) [14]. External quality assurance, including quality grading, followed ATS/ERS guidelines. We elected to use post-bronchodilator spirometry as this represented the best estimate of forced expiratory volume in 1 s (FEV_1). The sample size was determined by the primary objective of evaluating novel diagnostic tests for tuberculosis.

Participants were included if good-quality (A–C grade) post-bronchodilator spirometry was available. They were included in the "healthy" subpopulation used for primary analyses if they were nonsmokers or former smokers with less than one cumulative pack-year smoking history; did not report any lung disease (including no prior (self-reported) or current (screened for – details in supplementary materials) tuberculosis); and had a body mass index (BMI) in the normal range. Normal BMI was defined for adolescents (<19 years) as BMI-for-age z-score of between −2 and +2, calculated using sex-specific World Health Organization (WHO) reference equations, and for adults as BMI $\geqslant$18.5 $kg \cdot m^{-2}$ and <25.0 $kg \cdot m^{-2}$ [15].

Key variables of interest were FEV_1, forced vital capacity (FVC) and the ratio of FEV_1 to FVC (FEV_1/FVC), expressed as z-scores and as diagnostic categories. Additional potentially explanatory variables considered were age (expressed in years and categorised), sex, education, food insecurity, HIV status, diabetes status, BMI-for-age z-score and height-for-age z-score (both calculated among all participants using the 19-year-old normative values for all adults).

Scoping review

We identified reference equations developed in Africa through by scoping review (supplementary table S1). We searched Medline on 28 January 2024, with no date restrictions, using terms for spirometry/lung function, healthy and an expert-developed search for all countries in Africa. We considered articles where the primary objective was to develop a reference equation for spirometry interpretation and both male and female participants from a wide age range (*i.e.* not restricting to either older adults or children) were included. We used the authors' reported equations to calculate z-scores for our study population.

Statistical analysis

Data analysis was conducted in R. Primary analyses were conducted in the healthy subpopulation (as defined above). GLI z-scores were calculated using each of African American, Caucasian, Other and GLI Global reference equations and reference equations identified from the scoping review. We explored the distribution of FEV_1, FVC and FEV_1/FVC z-scores across equations through summary statistics (mean±SD). Since variation in FEV_1, FVC or FEV_1/FVC z-scores of up to 0.5 due to sampling variability been described [16], we considered a difference in means of more than 0.5 to be meaningful as previously recommended. We calculated the proportion of participants who had obstructive (FEV_1/FVC less than lower limit of normal (<LLN)) or preserved-ratio-impaired spirometry (PRISm; FVC<LLN with FEV_1/FVC$\geqslant$LLN) and displayed this graphically. The reference equation that resulted in z-scores with a mean closest to zero and SD closest to one was selected as the "best-fitting". Primarily considering this equation, we explored variation in FEV_1, FVC and FEV_1/FVC z-scores across site, age, sex, height- and BMI-for-age z-scores and measures of socioeconomic position. p-values were calculated to compare mean z-scores across groups using t-tests (binary variables) or ANOVA ($\geqslant$2 categories). For continuous variables we calculated Spearman rank correlation coefficients and associated p-values. We additionally explored the distribution of pre-bronchodilator FEV_1, FVC and FEV_1/FVC z-scores, across GLI reference equations. As a sensitivity analysis, analyses were repeated without BMI restriction.

Results

Study population

In total, 2109 people living in tuberculosis-affected households were enrolled (figure 1). Of those, participants were excluded due to not having a spirometry test done (n=293), not having post-bronchodilator spirometry (n=175) or having only low-quality spirometry (n=109). These participants were on average older and more likely to be female compared with those included in the analysis (supplementary table S2). A further 726 people were excluded from the healthy subpopulation; of these

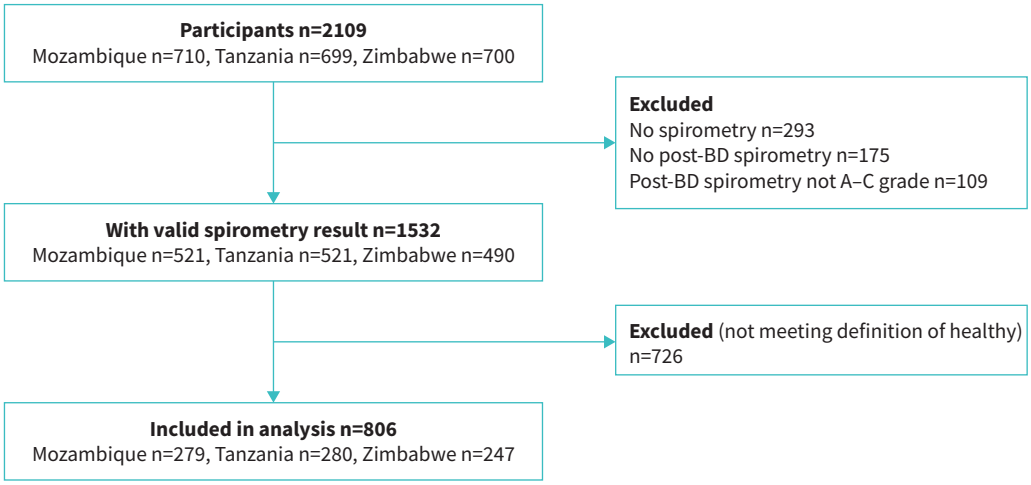

FIGURE 1 Participant flow diagram. Spirometry was performed and graded according to American Thoracic Society/European Respiratory Society standards. Of those excluded due to not meeting the definition of healthy, 518 of 726 were due to being under or overweight, defined as body mass index (BMI) for age z-score <−2 or >2 (if age <19 years) and BMI <18.5 $kg \cdot m^{-2}$ or ⩾25 $kg \cdot m^{-2}$ (if age ⩾19 years). BD: bronchodilator.

518 were excluded due to BMI outside the normal range. These people were on average older, more likely to be female or living with HIV compared with those included in the analysis (supplementary table S2). The remaining 806 healthy people for whom good-quality post-bronchodilator spirometry was available (247 in Harare, 280 in Mbeya and 279 in Maputo;) were included in the main analyses. Of these, 748 (93%) of spirometry traces were grade A. Among included participants, all were of black ethnicity, the median age was 18 years (interquartile range (IQR) 14–29 years), 57% were female and 9.8% were living with HIV (among which the median CD4 count was 574 (IQR 414–765); 78% were on ART; table 1).

TABLE 1 Characteristics of the study population (n=806)

Characteristic	Overall n=806	Female n=456	Male n=350
Age, years	18 (14–29)	19 (15–31)	17 (14–26)
Age category			
10–17 years	384 (48)	199 (44)	185 (53)
18–39 years	333 (41)	198 (43)	135 (39)
40+ years	89 (11)	59 (13)	30 (8.6)
Highest educational level[#]			
None/primary school	336 (43)	183 (42)	153 (46)
⩾ Secondary school	439 (57)	256 (58)	183 (54)
Insufficient food[¶]	189 (23)	112 (25)	77 (22)
HIV[+]	79 (9.8)	62 (14)	17 (4.9)
On ART	62 (78)	50 (81)	12 (71)
Diabetes[§]	29 (7.2)	18 (7.3)	11 (7.1)
Hypertension[§]	76 (19)	51 (22)	25 (16)
BMI-for-age z-score[ƒ]	−0.18 (−0.82–0.41)	0.05 (−0.62–0.61)	−0.48 (−0.99–0.11)
Height-for-age z-score[ƒ]	−0.76 (−1.40–−0.02)	−0.63 (−1.25–0.13)	−0.90 (−1.64– −0.21)

Data are presented as median (interquartile range) or n (%). ART: antiretroviral therapy; BMI: body mass index. [#]: 31 people were missing highest educational level. [¶]: Insufficient food from participant self-report ("was there any day in the past 6 months where you did not have enough food"). [+]: 3 people were missing HIV status, denominator for ART is the number of people with HIV. [§]: Diabetes and hypertension status are based on either a self-reported diagnosis, report of being on medication for these conditions, or (among participants ⩾18 years only) elevated glycated haemoglobin or blood pressure, respectively. [ƒ]: BMI and height-for-age z-scores were calculated using the World Health Organization reference population, with the highest category (19 years) used as the reference value for all adults.

Characteristics of study participants and households differed by site (supplementary tables S4 and S5). In Harare and Maputo 30% and 37% participants reported having insufficient food in the past 6 months, whereas in Mbeya only 4.3% participants did so. Participants in Harare and Maputo were on average taller than those in Mbeya (median height-for-age z-scores of −0.55, −0.36 and −1.21, respectively). In Mbeya, fewer people had completed secondary education (35% compared with 67–70%) and more households used solid fuel (biomass) stoves compared with the other two sites (31% compared with 8.6% in Maputo and 0% in Harare). Across all three sites, people were impoverished, with 92% households living on less than USD 1.90 per person per day (corresponding to the World Bank-defined international poverty line in 2021 [17]).

Results of scoping review

Scoping searches returned 72 studies, from which four (from Cameroon, Mozambique, Madagascar and Nigeria) were included [3–6]. Three of these studies only recruited adults (age ⩾18 years) with varying upper age limits and one included both children and adults (Cameroon). Three of the four papers used a linear model as the basis of prediction equations, concluding that more-complex models did not confer better fit compared with this approach in their data (supplementary table S1).

Lung function impairment as defined by different GLI reference equations

The proportion of people with spirometry-defined lung function impairment differed considerably depending on the reference equation applied (table 2 and figure 2b). Using the African American reference equation, one in seven people had any type of lung function impairment, but using the Caucasian reference equation, this was almost one in two. Slightly fewer people were classified as having lung function impairment when using GLI Global compared with Other (any impairment: African American 15%, GLI Global 26%, Other 33% and Caucasian 49%). Differences in proportion of people classified as having lung impairment across reference equations were driven by large differences in the proportion of people categorised as having PRISm (African American 11%, GLI Global 22%, Other 28% and Caucasian 45%; table 2), with little difference in classification for obstruction. Each set of z-scores resembled a normal distribution, with varying degrees of left-shift (supplementary figure S1). Across the GLI reference equations considered, the African American equation fitted best both overall and across subgroups of site, age category and sex (overall, mean±SD FEV_1 z-score −0.12±1.19, mean FVC z-score −0.35±1.15, mean FEV_1/FVC z-score 0.39±1.16; figure 2a and supplementary table S6). For all other GLI equations, mean FEV_1 and FVC z-scores were less than −0.5. FVC z-scores were proportionately lower than those for FEV_1, resulting in FEV_1/FVC ratios that were slightly elevated. Findings were similar when pre-bronchodilator spirometry values were used for interpretation and when analyses were repeated without BMI restriction (supplementary figure S2).

Using reference equations from Africa resulted in similar fit compared with the GLI African American equation, with equations from Cameroon fitting best (mean FEV_1 z-score 0.071.12 and FVC z-score 0.01 ±1.01; supplementary table S6 and supplementary figure S3). All subsequent analyses are conducted using the GLI African American equation.

TABLE 2 Proportion of people with spirometry-detected post-bronchodilator lung impairment by type and severity grading (n=806)

Characteristic	GLI reference equation			
	African American	Caucasian	Other/mixed	GLI Global
Overall	123 (15)	394 (49)	264 (33)	209 (26)
Obstruction[#]	31 (3.8)	31 (3.8)	40 (5.0)	34 (4.2)
Mild	17 (2.1)	18 (2.2)	23 (2.9)	21 (2.6)
Moderate–severe	14 (1.7)	13 (1.6)	17 (2.1)	13 (1.6)
PRISm[#]	92 (11)	363 (45)	224 (28)	175 (22)
Mild	67 (8.3)	213 (26)	141 (17)	130 (16)
Moderate–severe	24 (3.0)	148 (18)	83 (10)	43 (5.3)

Data are presented as n (%). GLI: Global Lung Function Initiative. [#]: Diagnostic and severity grading are according to the American Thoracic Society/European Respiratory Society criteria: obstruction indicated by ratio of forced expiratory volume in 1 s (FEV_1) to forced vital capacity (FVC) less than lower limit of normal (LLN); PRISm (preserved-ratio-impaired spirometry): FVC less than LLN with FEV_1/FVC greater than or equal to LLN; and severity grading of mild as −1.65 to −2.5 and moderate–severe as ⩽−2.51.

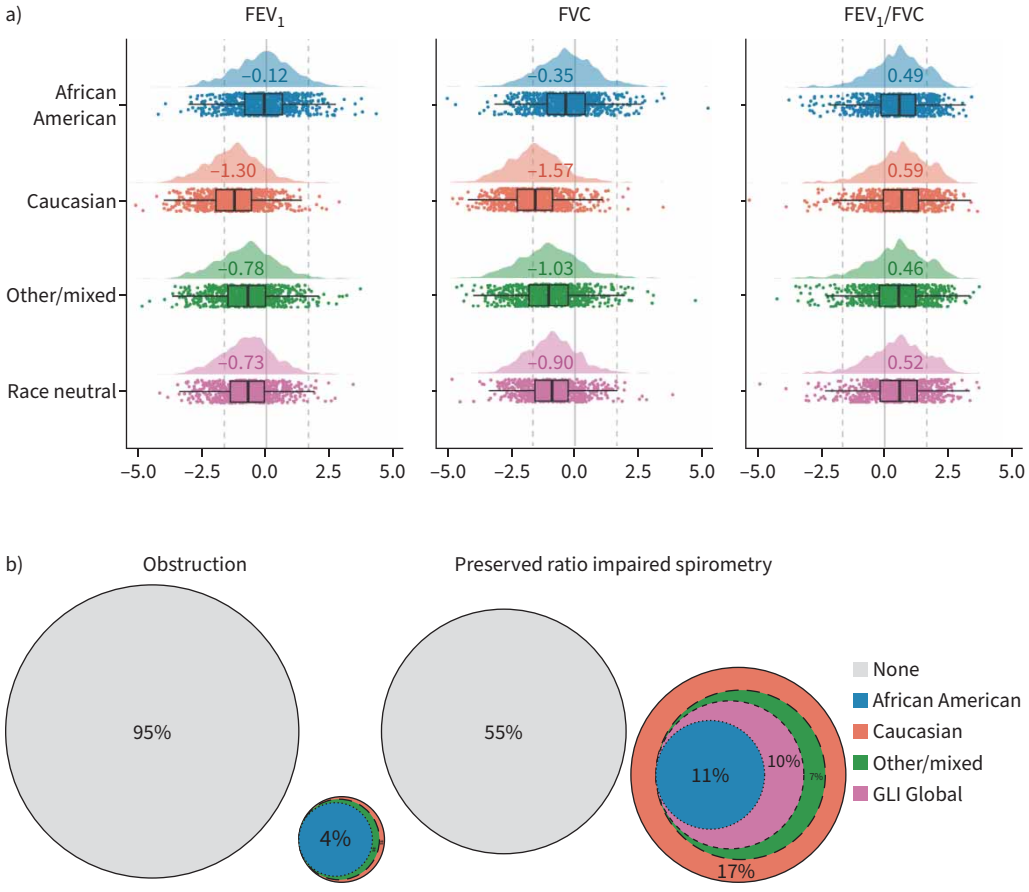

FIGURE 2 Fit of Global Lung Function Initiative (GLI) reference equations and impact on classification of obstructive lung disease and preserved-ratio-impaired spirometry among healthy members of tuberculosis-affected households (n=806). a) fit of various GLI reference equations to the healthy population in ERASE-TB across key indices (forced expiratory volume in 1 s (FEV_1), forced vital capacity (FVC) and ratio of FEV_1 to FVC). Annotations above box plots indicate mean z-score. b) implications of GLI reference equations for classification of obstruction and preserved-ratio-impaired spirometry, defined as FEV_1/FVC ratio less than the lower limit of normal (LLN) and FVC less than LLN with a normal FEV_1/FVC ratio, respectively. Different reference equations are indicated by colours (legend) and line type: solid: Caucasian/none; long-dash: other/mixed; mid-dash: GLI Global; fine dots: African American. For both panels, post-bronchodilator spirometry was used for interpretation. #: indicates <1%. GLI Global does not appear on the figure for obstructive lung disease as it is completely superimposed by the African American equation.

FEV_1 and FVC z-scores differed considerably across the three study sites (table 3 and supplementary figure S4), but not across age, sex or HIV status (as differences in z-scores were <0.5). Compared with Harare, participants in Maputo had, on average, lower FEV_1 and FVC z-scores (difference in means −0.48 and −0.30, respectively), and those from Mbeya tended to be higher (difference in means 0.71 and 0.84, respectively); whereas in the overall study population, people who were taller (greater height-for-age z-score) seemed to have lower FEV_1 and FVC z-scores (supplementary figure S5). We explored whether this association could be due to confounding by site: people in Mbeya (of the same sex and age) were shorter than those at the other two sites (leading to lower predicted values for FEV_1 and FVC). After stratification by site, no association between height-for-age and FEV_1 or FVC was seen. There was no evidence that people exposed to biomass fuels had lower lung function parameters compared with those using clean stoves (supplementary table S7) and there were no differences by BMI-for-age z-scores, or measures of socioeconomic position (not shown).

Discussion

Among healthy people living in tuberculosis-affected households in East and Southern Africa, the proportion with spirometry-defined lung function impairment differed considerably depending on the

TABLE 3 Summary measures of key spirometric indices across key subgroups, among healthy members of tuberculosis-affected households (n=806)

Stratum	Level	Absolute values		z-scores		
		Mean±SD	p-value	Mean±SD	Difference	p-value*
FEV_1						
Overall		2.63±0.76		−0.12±1.19		
Site	Harare	2.68±0.75	0.4	−0.20±1.16	Ref.	<0.001
	Maputo	2.57±0.68		−0.68±0.97	**−0.48**	
	Mbeya	2.66±0.84		0.51±1.11	**0.71**	
Sex	Female	2.42±0.54	<0.001	−0.20±1.19	Ref.	0.002
	Male	2.91±0.91		−0.02±1.17	0.18	
Age category	10–17 years	2.30±0.64	<0.001	−0.27±1.16	Ref.	<0.001
	18–39 years	3.07±0.71		0.04±1.22	0.31	
	40+ years	2.43±0.61		−0.12±1.09	0.15	
HIV	Negative	2.64±0.77	0.07	−0.13±1.18	Ref.	0.3
	Positive	2.54±0.68		−0.03±1.27	0.10	
FVC						
Overall		2.96±0.84		−0.35±1.15		
Site	Harare	2.96±0.82	0.8	−0.55±1.14	Ref.	<0.001
	Maputo	2.92±0.76		−0.81±0.98	−0.30	
	Mbeya	2.99±0.93	0.29±1.05		**0.84**	
Sex	Female	2.71±0.60	<0.001	−0.40±1.19	Ref.	0.05
	Male	3.28±0.98		−0.29±1.10	0.11	
Age category	10–17 years	2.55±0.68	<0.001	−0.42±1.12	Ref.	0.04
	18–39 years	3.42±0.79		−0.30±1.20	0.12	
	40+ years	2.99±0.73		−0.22±1.12	0.20	
HIV	Negative	2.96±0.84	>0.9	−0.36±1.15	Ref.	0.1
	Positive	2.96±0.79		−0.22±1.22	0.14	

FEV_1: forced expiratory volume in 1 s; FVC: forced vital capacity. z-scores are calculated using the Global Lung Function Initiative African American reference (Ref.) equation. Differences are absolute differences in z-score compared with the reference category, with differences >0.5 (our predefined meaningful difference) highlighted in bold. *: p-values are from t-tests (binary variables) and ANOVA (⩾2 categories), comparing means across subgroups.

reference equation applied. In this context, using GLI Global as the reference equation resulted in almost twice as many people being classified as having PRISm compared with the African American equation. The proportion of people classified as having PRISm was slightly lower when GLI Global was used, as compared with Other. A similar number had obstruction across different reference equations (that expected when using a LLN threshold in a healthy population, *i.e.* ∼5% [1]). GLI-derived FVC z-scores were on average lower than those for FEV_1, leading to mean FEV_1/FVC z-scores of greater than zero (but not meeting our predefined threshold for a meaningful difference), as has been noted across other African populations [18]. Use of local African reference equations resulted in FEV_1/FVC z-scores closer to zero.

There were important differences across sites, with participants in Mbeya having on average the highest z-scores and in Maputo, the lowest. Reasons for differences in distribution of lung function parameters between populations are multiple and may include differences in anthropometry (*e.g.* thoracic diameter and relative leg length), differential distribution of gene variants, and environmental exposures (particularly in early life) leading to impaired lung development. Secular changes such as improving socioeconomic conditions impact population-level anthropometry [19] and these effects could also be at play. Future studies that include the measurement of sitting height, which partially attenuates differences in lung function across ethnic groups [20], or other anthropometric measures (*e.g.* ulnar length and thoracic diameter) not included in this study could explore this further.

The fact that the use of GLI Global results in more people being classified as having abnormal lung function compared with ethnicity-specific equations has been described in high-income settings and is one of the reasons (*i.e.* to prevent missed diagnoses of respiratory disease) why its use has been advocated [4]. In the United States and United Kingdom, ethnicity-specific and universal reference equations have similar accuracy in predicting symptomatic, clinical and functional outcomes [21]. Longitudinal data from Africa

are limited. NHANES 1999 Caucasian equation-defined impaired spirometry in a population-based study in Malawi, East Africa [22] or GLI Other-defined impairment among people after tuberculosis treatment in Benin, West Africa correlated poorly with respiratory symptoms or functional ability [23], whereas using a locally derived equation resulted in a stronger association. This raises the question of whether lung impairment defined in an "ideal" reference population, applied in a disparate population, is clinically meaningful [22]. However, a multi-country study with data from diverse settings found that lower FEV_1 within the country-specific normal range was associated with higher mortality, cardiovascular events and respiratory hospitalisation [24]. In clinical decision-making it is important that individuals are not disadvantaged by acceptance of a lower "normal" lung function; therefore being denied treatments that may improve symptoms, quality of life or prevent death. Simultaneously, a higher threshold for "normal" lung function has important consequences, including resulting in people being deemed ineligible for certain employment [22]. These dual concerns highlight the need for spirometry (particularly when close to the LLN) to be interpreted in the context of an individual's clinical history, symptoms and other measures. In Africa, the clinical significance and treatment requirements of PRISm are unclear. Concurrently, access to pulmonologists and medications for respiratory disease are highly restricted both due to capacity and cost, with health systems that are ill-equipped to detect and manage the enormous burden of chronic respiratory disease [11, 25]. In a context where people with even severe respiratory disease cannot access care, over-diagnosis has critical implications for health systems and individuals. Greater advocacy, funding and research to drive policy and improve respiratory health in Africa are urgently needed. For now, the benefits and risks of moving to GLI Global, acknowledged by ATS, should be carefully considered in each individual context.

In addition to guiding clinical decision-making, GLI reference equations are used for population health and to attribute lung disease to specific exposures. In these applications, it is necessary to understand who is being compared with who. Several previous studies have evaluated the fit of GLI reference equations in healthy African populations and compared these to geographically specific equations [26–31]. In South Africa, GLI Other seemed best, whereas across several other East, southern and West African studies, GLI African American seemed so. The underlying reasons for these differences are multifactorial; however, it should be noted that an earlier study in South Africa, among a more socioeconomically deprived population, found African American to be the best-fitting equation [31]. Integration of existing data from Africa into GLI equations; together with new, population-representative data and prospective cohort studies, which include measurement of clinical outcomes such as symptom burden, quality of life and respiratory-related hospitalisations are urgently needed. In considering what reference equation to use, it is important to acknowledge the ubiquity of hazardous exposures (*e.g.* undernutrition, childhood infections, and indoor and outdoor air pollution) in many African contexts. In these settings, lung function measurements from healthy populations may not reflect "optimum" lung function, and use of an ethnicity-specific (or geographically specific) reference equation may "normalise poverty" [32]. However, this limitation also exists within GLI Global. If, as has been shown, people of non-white ethnicity in the GLI dataset are systematically disadvantaged, then such disadvantage persists in the resultant equations, regardless of how they are weighted. Considering the Caucasian reference equation as reflecting "ideal" lung function (*i.e.* being that of people with adequate means and living in a clean environment) may sometimes be appropriate but does not account for anthropometric or genetic differences [22]. Understanding the impact of ubiquitous exposures on lung function will becritical to drive policy and funding to address social and structural determinants of lung health in low-resource settings. This is of particular importance in the face of the existential threat of climate change which is resulting in food insecurity and worse air quality, with its greatest impacts in low- and middle-income countries.

Dichotomisation of lung function results as "normal" or "abnormal" using arbitrarily defined thresholds (*e.g.* LLN) exacerbates the impact of shifts in the underlying population distribution and therefore a nuanced approach is needed in interpreting borderline results. In exploring the associations between different exposures and lung function, use of continuous outcomes may be more appropriate. For public health research, we urge authors to include details of the reference equation (including which ethnic group, if any, was used) in reports of pulmonary function testing and consider applying different reference equations as part of their supplementary work to facilitate meaningful meta-analyses and cross-population comparisons.

The strengths of this study include data from a large multi-country cohort with standardised training, procedures and equipment. We used post-bronchodilator spirometry, performed according to ERS/ATS standards and including internal and external quality assurance of all flow-volume loops. Only high-quality recordings were included in the analysis. As a result, bias due to low-quality spirometry is unlikely, and differences described between sites cannot be attributed to technical differences.

While our overall sample size was large, within each site we had slightly fewer men meeting our criteria for "healthy" than is recommended for validation of GLI reference equations (recommendation of 150 *versus* Maputo 128, Mbeya 126 and Harare 96). We excluded people with a diagnosis of respiratory disease (including those with previous tuberculosis and smokers); however, people with undiagnosed respiratory disease may be included within the population that we defined as "healthy". This is in keeping with many other studies evaluating GLI reference equations. While our population (that of tuberculosis-affected households) is socioeconomically deprived and not population-representative, our data had a good fit to that derived from other, population-representative, studies from Africa. Finally, we do not have longitudinal outcome data or assessments of gas transfer and advanced imaging (*e.g.* computed tomography), which could have provided further insights into the underlying pathology of PRISm. These assessments are either extremely limited or not available in our setting.

In conclusion, among members of tuberculosis-affected households in East and Southern Africa, the GLI African American reference equation had the best fit and performed broadly similarly to African-derived reference equations. The key limitation of this work is the possibility that the analysis population included some people with undiagnosed respiratory disease. The choice of spirometry reference equation has critical implications for both clinical decision-making and public health. These need to be carefully considered, particularly in resource-constrained environments. Further work is needed to ensure that GLI Global is representative of a global population; to understand how GLI Global-defined lung function impairment correlates with clinical outcomes in diverse settings; and to understand the true impact of anthropometry, genes and the environment (including socioeconomic deprivation and early-life circumstances) on lung function and its interpretation.

Acknowledgements: We gratefully thank all those who participated in the study and the study personnel from the Biomedical Research and Training Institute and the Zvitambo Research Institute, Zimbabwe; the Instituto Nacional de Saúde, Mozambique; and the National Institute for Medical Research – Mbeya Medical Research Centre, Tanzania.

Data availability: Data and code to recreate the quantitative analyses presented in this manuscript will be made available *via* LSHTM Data Compass (https://doi.org/10.17037/DATA.00004267).

Provenance: Submitted article, peer reviewed.

Ethics statement: Ethical approval for the study was granted by the Medical Research Council in Zimbabwe (MRCZ/A/2618), the Mbeya Medical Research and Ethics Review Committee (SZEC-2439/R.A/V.1/101), the National Health Research Ethics Committee in Tanzania NIMR/HQ/R.8a/Vol.IX/3608), and Tanzanian Medicines and Medical Devices Authority (TMDA-WEB0021/CTR/0004/03), the National Bioethics Committee for Health in Mozambique (541/CNBS/21), and the ethical committees of London School of Hygiene and Tropical Medicine, UK (22522–2) and the medical faculty of the Ludwig Maximilian-Universität München, Germany (20–0771).

Author contributions: D. Banze, C.J. Calderwood, N. Heinrich, K. Kranzer, A. Rachow, J.R. Hurst and C. Khosa conceptualised and designed the study. D. Banze, C.J. Calderwood, C. Nhamuave, E.T. Marambire, A. Mfinanga, L.T. Minja and C. Khosa contributed to data acquisition. C.J. Calderwood and D. Banze undertook the analysis and interpretation of data, with input from R.A. Ferrand, K. Fielding, K. Kranzer, A. Rachow, J.R. Hurst and C. Khosa. D. Banze and C.J. Calderwood wrote the first draft, with input from A.M. Malhotra, J.R. Hurst and C. Khosa. All authors contributed to subsequent drafts and approved the final version for publication.

Conflict of interest: The authors declare no conflicts of interest.

Support statement: This project was supported by the grant (RIA2018D-2508 ERASE-TB) through the European and Developing Countries Clinical Trials Partnership (EDCTP), with assistance from the Department of Health and Social Care UK under the EDCTP2 programme supported by the European Union. The German Center for Infection Research (DZIF) (TTU 02.912_00) contributed further funding. C.J. Calderwood is funded by the Wellcome Trust (225466/Z/22/Z). Cepheid provided test kits at no cost to the consortium. The funders had no role in the study design, analysis, manuscript preparation or decision to publish. Funding information for this article has been deposited with the Open Funder Registry.

References

1 Stanojevic S, Kaminsky DA, Miller MR, *et al.* ERS/ATS technical standard on interpretive strategies for routine lung function tests. *Eur Respir J* 2022; 60: 2101499.

2 Bowerman C, Bhakta NR, Brazzale D, *et al.* A race-neutral approach to the interpretation of lung function measurements. *Am J Respir Crit Care Med* 2022; 207: 768–774.
3 Marciniuk DD, Becker EA, Kaminsky DA, *et al.* Effect of Race and Ethnicity on Pulmonary Function Testing Interpretation: An American College of Chest Physicians (CHEST), American Association for Respiratory Care (AARC), American Thoracic Society (ATS), and Canadian Thoracic Society (CTS) Evidence Review and Research Statement. *Chest* 2023; 164: 461–475.
4 Bhakta NR, Bime C, Kaminsky DA, *et al.* Race and ethnicity in pulmonary function test interpretation: an Official American Thoracic Society Statement. *Am J Respir Crit Care Med* 2023; 207: 978–995.
5 Louw SJ, Goldin JG, Joubert G. Spirometry of healthy adult South African men. Part I. Normative values. *S Afr Med J* 1996; 86: 814–819.
6 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: 819–829.
7 Elmaleh-Sachs A, Balte P, Oelsner EC, *et al.* Race/ethnicity, spirometry reference equations, and prediction of incident clinical events: the Multi-Ethnic Study of Atherosclerosis (MESA) Lung Study. *Am J Respir Crit Care Med* 2022; 205: 700–710.
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: 1324–1343.
9 Ubuane PO, Ajiboye OA, Akinola AO *et al.* Development of reference values and equations for the pulmonary function of Nigerian children aged 6–11 years measured with digital peak flow meter and its validation against local and the Global Lung Function Initiative (GLI) equations. *J Xiangya Med* 2024; 9: 1–21.
10 Mannino DM, Townsend MC. Spirometry in 2022: is a single set of prediction equations for all the best path forward? *Am J Respir Crit Care Med* 2023; 207: 659–661.
11 Meghji J, Mortimer K, Agusti A, *et al.* Improving lung health in low-income and middle-income countries: from challenges to solutions. *Lancet* 2021; 397: 928–940.
12 Marambire ET, Banze D, Mfinanga A, *et al.* Early risk assessment in paediatric and adult household contacts of confirmed tuberculosis cases by novel diagnostic tests (ERASE-TB): protocol for a prospective, non-interventional, longitudinal, multicountry cohort study. *BMJ Open* 2022; 12: e060985.
13 Calderwood CJ, Timire C, Mavodza C, *et al.* Beyond tuberculosis: a person-centred and rights-based approach to screening for household contacts. *Lancet Glob Health* 2024; 12: e509–e515.
14 Graham BL, Steenbruggen I, Barjaktarevic IZ, *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: E70–E88.
15 de Onis M, Onyango AW, Borghi E, *et al.* Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ* 2007; 85: 660–667.
16 Quanjer PH, Stocks J, Cole TJ, *et al.* Global Lungs Initiative. Influence of secular trends and sample size on reference equations for lung function tests. *Eur Respir J* 2011; 37: 658–664.
17 World Bank. Fact sheet: an adjustment to global poverty lines. Date last accessed: 2 Sep 2024. Date last updated: 14 Sep 2022. www.worldbank.org/en/news/factsheet/2022/05/02/fact-sheet-an-adjustment-to-global-poverty-lines
18 Pefura-Yone EW, Balkissou AD, Poka-Mayap V, *et al.* Spirometric reference equations for Cameroonians aged 4 to 89 years derived using lambda, mu, sigma (LMS) method. *BMC Pulm Med* 2021; 21: 344.
19 Quanjer PH, Kubota M, Kobayashi H, *et al.* Secular changes in relative leg length confound height-based spirometric reference values. *Chest* 2015; 147: 792–797.
20 Harik-Khan RI, Fleg JL, Muller DC, *et al.* The effect of anthropometric and socioeconomic factors on the racial difference in lung function. *Am J Respir Crit Care Med* 2001; 164: 1647–1654.
21 Diao James A, He Y, Khazanchi R, *et al. et al.* Implications of race adjustment in lung-function equations. *N Engl J Med* 2024; 390: 2083–2097.
22 Meghji J, Nadeau G, Davis KJ, *et al.* Noncommunicable lung disease in sub-Saharan Africa a community-based cross-sectional study of adults in urban Malawi. *Am J Respir Crit Care Med* 2016; 194: 67–76.
23 Fiogbe AA, Agodokpessi G, Tessier JF, *et al.* Prevalence of lung function impairment in cured pulmonary tuberculosis patients in Cotonou, Benin. *Int J Tuberc Lung Dis* 2019; 23: 195–202.
24 Duong M, Islam S, Rangarajan S, *et al.* Mortality and cardiovascular and respiratory morbidity in individuals with impaired FE (PURE): an international, community-based cohort study. *Lancet Glob Health* 2019; 7: e613–e623.
25 Plum C, Stolbrink M, Zurba L, *et al.* Availability of diagnostic services and essential medicines for non-communicable respiratory diseases in African countries. *Int J Tuberc Lung Dis* 2021; 25: 120–125.
26 Fawibe AE, Odeigah LO, Saka MJ. Reference equations for spirometric indices from a sample of the general adult population in Nigeria. *BMC Pulm Med* 2017; 17: 48.
27 Ivanova O, Khosa C, Bakuli A, *et al.* Lung function testing and prediction equations in adult population from Maputo, Mozambique. *Int J Environ Res Public Health* 2020; 17: 4535.

28 Arigliani M, Canciani MC, Mottini G, *et al.* Evaluation of the Global Lung Initiative 2012 reference values for spirometry in African children. *Am J Respir Crit Care Med* 2017; 195: 229–236.
29 Smith S-J, Gray DM, MacGinty RP, *et al.* Choosing the Better Global Lung Initiative 2012 equation in South African population groups. *Am J Respir Crit Care Med* 2020; 202: 1724–1727.
30 Madanhire T, Ferrand RA, Attia EF, *et al.* Validation of the global lung initiative 2012 multi-ethnic spirometric reference equations in healthy urban Zimbabwean 7–13 -year-old school children: a cross-sectional observational study. *BMC Pulm Med* 2020; 20: 56.
31 Masekela R, Hall GL, Stanojevic S, *et al.* An urgent need for African spirometry reference equations: the Paediatric and Adult African Spirometry study. *Int J Tuberc Lung Dis* 2019; 23: 952–958.
32 Masekela R, Mortimer KJ, Feary J, *et al.* Stop normalizing poverty: how can African children achieve their true lung health potential? *JPATS* 2023; 4: 3–7.