BMJ Open Respiratory Research

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

Haruna Kitazawa ,[1,2] Annie Jiang ,[1] Cynthia Nohra ,[1] Honami Ota ,[1] Joyce K Y Wu ,[1,2] Clodagh M Ryan,[1,2,3] Chung-Wai Chow [1,2]

To cite: 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 Resp Res* 2022;**9**:e001389. doi:10.1136/bmjresp-2022-001389

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjresp-2022-001389).

Received 3 August 2022
Accepted 18 November 2022

[1]Department of Medicine, University Health Network, Toronto, Ontario, Canada
[2]Department of Medicine, Temerty Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada
[3]Sleep Research Laboratory, Toronto Rehabilitation Institute University Health Network, Toronto, Ontario, Canada

Correspondence to
Dr Chung-Wai Chow;
Chung-Wai.Chow@uhn.ca

ABSTRACT

Background The Global Lung Function Initiative (GLI-2012) focused on race/ethnicity as an important factor in determining reference values. This study evaluated the effects of changing from Canadian reference equations developed from an all-Caucasian cohort with European ancestry to the GLI-2012 on the interpretation of spirometry in a multiethnic population and aimed to identify the ethnic groups affected by discrepant interpretations.

Methods Clinically indicated spirometry in a multiethnic population (aged 20–80 years) collected from 2018 to 2021 was analysed. The predicted and lower limit of normal (LLN) values were calculated using three sets of reference equations: Canadian, GLI-race/ethnic-based (GLI-Race) and GLI-race/ethnic-neutral (GLI-Other). We compared the prevalence of concordance in the abnormal diagnoses (defined as <LLN) for forced vital capacity (FVC), forced expiratory volume in 1 s (FEV_1), and FEV_1/FVC among the three reference values, and evaluated whether race/ethnicity was associated with discordance.

Results Data from 406 participants were evaluated (non-Caucasian 43.6%). There was 85%–87% concordance for normal/abnormal FVC and FEV_1 interpretations among the Canadian, GLI-Race and GLI-Other reference equations. In all ethnic groups, application of the Canadian references for interpretation led to a higher prevalence of abnormal (<LLN) FVC and FEV_1compared with GLI-Race and GLI-Other. This trend was more prominent in Black, South-East Asian and Mixed/other ethnic groups when comparing the Canadian to the GLI-Race equations. In contrast, the discordance rates were similar among ethnic groups when compared with the GLI-Other reference equations. Interpretation of FEV_1/FVC had a high rate of agreement among all equations.

Conclusion Interpretation using Canadian reference equations was associated with a higher prevalence of restrictive physiology compared with the GLI-2012 equations, particularly if the GLI-Race were used. These observations were mostly found in non-white Caucasian groups, highlighting the need to choose reference equations that reflect closely the ethnic mix of the population being evaluated in order to optimise patient management.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Race/ethnicity is one of the important features in determination of normal reference values for spirometry. The Global Lung Function Initiative (GLI-2012) developed reference equations that account for differences in lung function between race/ethnic groups.

WHAT THIS STUDY ADDS

⇒ Use of the Canadian reference equations derived from an all-white Caucasian cohort with European ancestry results in overinterpretation of abnormal forced vital capacity and forced expiratory volume in 1 s (< lower limit of normal) in all ethnic groups compared with the GLI-2012. The magnitude of the discordance was more prominent in Black, South-East Asian and Mixed/other ethnic groups when compared with GLI-Race/ethnic-based reference equations.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our data revealed the extent by which discordance in interpretations occurs in each reference equation set, the Canadian, the GLI-race/ethnic-based and GLI-race/ethnic-neutral reference equations and found that race/ethnicity was significantly associated with these discrepancies. This study suggests that lung function laboratories should carefully evaluate the choice of reference equations for the interpretation of lung function tests to better reflect the ethnic mix of the patient population to provide optimal clinical care.

INTRODUCTION

Spirometry is the most commonly used pulmonary function test (PFT) and plays a central role in diagnosis and management of lung diseases. It is interpreted in the context of reference values derived from a healthy population that is reflective of the patient

being evaluated.[1 2] The use of inappropriate prediction equations can lead to misinterpretation to result in missed or misdiagnosis of restrictive and/or obstructive lung disease.[3 4] For these reasons, the Global Lung Function Initiative (GLI) developed new reference equations, the GLI-2012, to better reflect the patient populations with the intent of improving diagnostic acumen. GLI-2012 has created reference equations for all ages and multiethnic groups.[5] In addition to the four main race/ethnicity groups (Caucasian, Black, North-East (NE) Asian and South-East (SE) Asian), GLI-2012 also provides an 'Other' equation that corresponds to other groups and individuals of mixed ethnic origin, by averaging the four main groups.

Studies comparing GLI-2012 to other reference equations in various races and respiratory diseases[6–12] have shown that the GLI race/ethnic-based reference equations (GLI-Race) could fit the population in several validation samples. However, a recent publication found no evidence that interpretation using the GLI-Race reference equations improved the prediction of clinical events compared with the race/ethnic-neutral equations for Other/Mixed ethnicity (GLI-Other).[13] Moreover, Baugh *et al* reported that the % predicted values for forced expiratory volume in 1 s (FEV_1) and forced vital capacity (FVC) derived from GLI-other equations more accurately reflected clinically relevant outcomes than those derived from race-specific equations[14] in an American population of smokers with and at-risk for chronic obstructive pulmonary disease and that race-specific equations underestimated disease severity.[15]

The PFT laboratories in the University of Toronto affiliated academic hospitals, a major academic medical centre in Canada, use a set of reference equations that was developed from data collected in 627 all-white Caucasian Canadians of European descent.[16] Toronto is a large metropolitan city that has seen significant growth over the past several decades and is one of the most ethnically diverse populations in the North America. Data from 2016 Canadian Census show that 45% of Toronto respondents reported Asian (40%) or African (5%) ethnic origin.[17] It is questionable whether the Canadian reference equations[16] adequately represent the current population in the region. We postulate that the application of difference reference equations will change the interpretation of spirometry and alter the prevalence of abnormal findings. To our best knowledge, few studies have compared the Canadian[16] and the GLI-2012 reference equations.[18]

The primary aim of this study is to determine the effects of changing from the Canadian to the 2012 GLI-Race and GLI-Other reference equations on the interpretation of spirometry in a large group of patients recently evaluated with spirometry. The secondary aim is to identify the ethnic groups where discrepancies in the interpretation occur.

METHODS

Participants and classification of race/ethnicity

This is a retrospective analysis of pulmonary function data that were collected from 6 January 2018 to 8 December 2021 in ongoing research studies (REB number 19–5582 and 17–5652) where data regarding ethnicity of the study subjects were available. All participants signed informed consent. The current study included adults who were recruited as healthy control subjects and the patients followed by the General Internal Medicine or the Respirology Services and referred for PFTs for clinical assessment of respiratory symptoms. Ethnicity of participants was self-reported using categories as shown in online supplemental table S1. Only data from participants aged 20–80 years were included as the data for the Canadian reference equations only considered this age range.[16]

Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting or dissemination plans of this study.

Spirometry

Spirometry and full pulmonary function studies were performed using the MIR Minispir (MIR, Rome, Italy) or Bodybox (Medisoft, Sorinnes, Belgium) by qualified technologists in clinic or in the Toronto General Pulmonary Function Laboratory. All testing procedures followed American Thoracic Society and European Respiratory Society Guidelines.[19] Quality control audits of the data were conducted monthly. Only data that passed quality control were used for this study. For participants who had repeated testing done during the study period, only the first test was included in the analysis. Only prebronchodilator values were collected and analysed.

Predicted and the lower limit of normal values of lung function

Three sets of reference equations were used for derivation of normal reference values: (1) Canadian set[16] which does not take into account race/ethnicity, (2) GLI-race/ethnic-based spirometry reference equations (*GLI-Race*) and (3) GLI 2012-Other equation which is race/ethnic-neutral (*GLI-Other*).[5] For each reference set, % predicted and the lower limit of normal (LLN) values, corresponding to the lowest fifth percentile of predicted values, were calculated. For the GLI-Race calculations, we applied the GLI-2012 classifications according to the self-reported ethnicities[5 20] (online supplemental table S2). For respondents who self-identified as Chinese but where the geographic region of origin in China is unknown (n=21), we considered them as either NE Asian or SE Asian categories in GLI-Race equations (online supplemental table S2). The results analysed by applying NE Asian (model 1) are shown in the manuscript, and those

analysed by applying SE Asian (model 2) are found in the online supplemental materials.

Statistical analysis

Comparisons between race/ethnicity groups were conducted using one-way analysis of variance (ANOVA) or Kruskal-Wallis test for continuous variables and Pearson's chi-square test or Fisher's exact test for categorical variables. Kruskal-Wallis test was performed to compare LLN differences between the Canadian and GLI-2012 reference equations among ethnic groups, and Bonferroni correction was used when multiple comparisons were calculated. Abnormal FVC, FEV_1, and FEV_1/FVC were defined when the measured values were less than LLN. First, we compared interpretation based on the Canadian reference equations with the GLI-2012 (GLI-Race or GLI-Other) reference sets for all participants or each ethnicity group using rates of concordance and discordance. Concordance was defined as the same outcome, while discordance was defined as different outcomes when comparing the two equations. Second, univariable and multivariable logistic regression modelling were performed to identify the factors related to one of the discordant pairs in interpretations between the Canadian and GLI-2012 reference equation (abnormal (< LLN) in the Canadian reference equations and normal (≥ LLN) in GLI-2012). All statistical analyses were performed using R (V.4.1.1)/Rstudio (V.1.4.1717).

RESULTS

During the study period, 419 patients underwent spirometry and from whom race/ethnic data were collected. We excluded 13 patients from analysis as their age range fell outside the Canadian reference equations (n=6, <20 years old; n=7, >80 years old). The demographic data revealed that 43.6% of participants were non-Caucasian with the majority self-identifying as SE Asian or Other/Mixed ethnicity (table 1). Interpretation using the Canadian reference equations led to lower percent predicted values for both FVC (FVC % predicted) and FEV_1 (FEV_1 % predicted) compared with the GLI-Race or GLI-Other

Table 1 Participant demographics and spirometry*

	All	Caucasian	Black	NE Asian	SE Asian	Other/Mixed	P
Number of participants	406	229	20	21	92	44	
Male	174 (42.9)	85 (37.1)	7 (35.0)	11 (52.4)	46 (50.0)	25 (56.8)	0.044
Age, years	48.68 (35.55, 59.87)	47.35 (35.96, 57.11)	48.06 (35.29, 57.96)	45.94 (39.80, 57.26)	55.89 (46.05, 68.12)	37.89 (29.19, 56.84)	<0.001
Height, cm	168.23±9.74	170.51±9.72	170.68±6.22	165.43±9.46	162.58±8.74	168.36±8.37	<0.001
Weight, kg	79.11±23.27	84.66±24.81	89.03±22.76	69.42±15.71	65.82±15.55	77.83±18.40	<0.001
BMI, kg/m^2	27.77±7.15	29.02±7.85	30.78±8.71	25.21±4.50	24.75±4.66	27.34±5.48	<0.001
Smoking status							0.11
Current smoker	16 (4.0)	6 (2.7)	0 (0.0)	1 (4.8)	8 (8.8)	1 (2.3)	
Ex-smoker	107 (26.7)	69 (30.5)	4 (20.0)	3 (14.3)	24 (26.4)	7 (16.3)	
Never smoker	278 (69.3)	151 (66.8)	16 (80.0)	17 (81.0)	59 (64.8)	35 (81.4)	
Lung function							
FVC, L	3.63±1.06	3.92±1.02	3.31±1.04	3.36±1.08	3.14±0.95	3.46±1.01	<0.001
FEV_1, L	2.88±0.87	3.11±0.83	2.68±0.83	2.71±0.80	2.42±0.82	2.81±0.82	<0.001
FEV_1/FVC, %	79.23±7.62	79.48±6.37	81.27±6.82	81.51±4.73	76.56±10.45	81.49±6.69	0.001
FVC % predicted as per							
Canadian equation	87.69±16.48	92.26±15.76	76.92±15.06	83.06±17.75	84.13±14.77	78.45±15.61	<0.001
GLI-Race	94.26±16.69	94.15±15.66	92.34±18.26	87.98±18.34	99.12±17.27	88.56±16.78	0.003
GLI-Other	97.61±17.79	102.30±17.03	85.47±16.78	92.43±19.36	94.08±16.17	88.56±16.78	<0.001
FEV_1 % predicted as per							
Canadian equation	85.44±16.17	89.93±15.48	76.39±14.27	83.00±15.81	80.27±15.95	78.13±13.60	<0.001
GLI-Race	91.86±16.78	92.99±16.17	92.04±16.88	88.02±16.79	92.05±18.70	87.26±15.23	0.242
GLI-Other	94.86±18.05	99.82±17.36	84.84±15.53	92.12±17.73	88.95±17.82	87.26±15.23	<0.001

Data are presented as n (%), mean±SD, or median (IQR).
Information on weight, BMI, and smoking status was missing in 2, 2 and 5, respectively.
*We analysed the data by classifying Asian participants as NE Asian (model 1) if they could not be classified clearly as either NE Asian or SE Asian.
BMI, body mass index; FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; GLI, Global Lung Function Initiative; NE, North East; SE, South East.

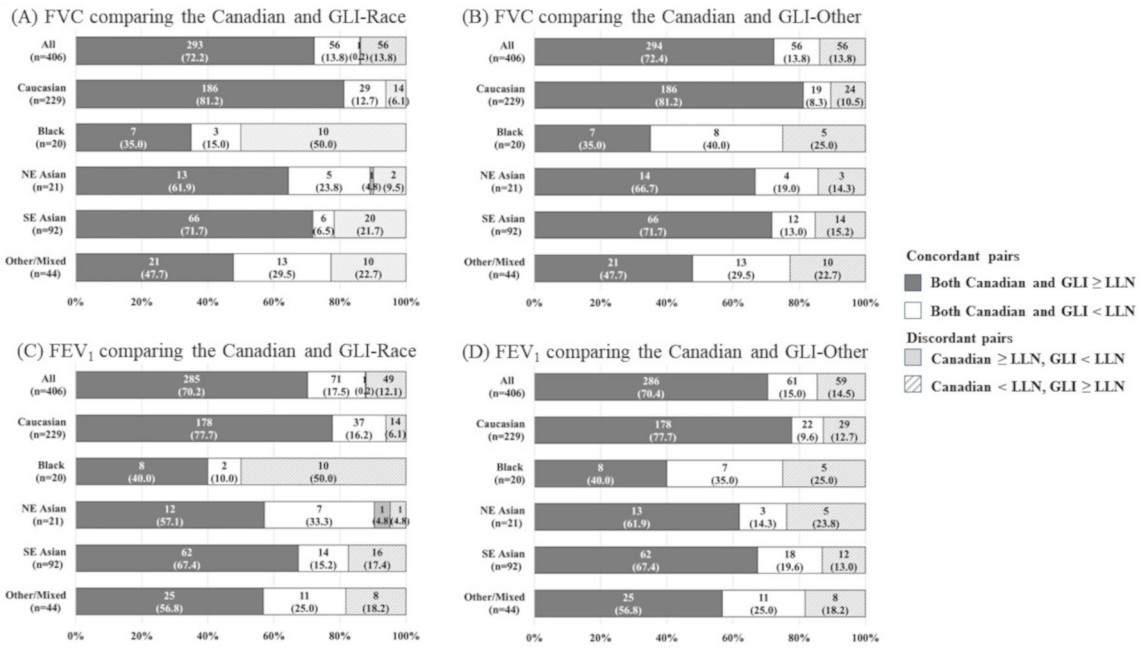

Figure 1 Stacked barchart comparing the concordant and discordant pairs for abnormal diagnosis (<LLN) when using Canadian reference equation compared with GLI-race/ethnic-based (GLI-Race) or GLI-race/ethnic-neutral (GLI-Other) equation. (A and B) FVC; (C and D) FEV_1. (A) and (C) compare the Canadian reference equations and GLI-Race, and (B) and (D) compare the Canadian reference equations and GLI-Other. Data are presented as n (%). Percentages may not total 100 due to rounding. We analysed the data by classifying Asian participants as NE Asian (model 1) if they could not be classified clearly as either NE Asian or SE Asian. FVC, forced vital capacity, FEV1, forced expiratory volume in 1 s, GLI, Global Lung Function Initiative, NE, North East, SE, South East, LLN, lower limit of normal.

reference equations in all participants and in all ethnic subgroups (p<0.001 respectively, paired t-test).

The concordance rates in the interpretation of abnormal FVC (defined as <LLN) between the Canadian, GLI-Race or GLI-Other reference equations are shown in figure 1A,B and table 2. Concordance was observed in 86% of participants. However, 56 participants (13.8%) whose FVC was considered abnormal according to the Canadian reference equations were interpreted to be normal by the GLI-Race or GLI-Other. This discordance rate was particularly pronounced for Black (50%), Other/Mixed (22.7%) and SE Asian (21.7%) ethnicity groups when the Canadian reference equations were compared with the GLI-Race (figure 1A and table 2). In contrast, when the interpretations based on the Canadian reference equations were compared with GLI-Other, the difference between the ethnic groups was reduced, with the exception of the Other/Mixed group which had a discordance rate of 22.7%. Similar observations were made in the diagnosis of abnormal FEV_1 (FEV_1<LLN), with the highest discordance rates in the Black, Other/Mixed and SE Asian groups (figure 1C,D and table 3). For the diagnosis of abnormal FEV_1/FVC (FEV_1/FVC<LLN), the concordance rate was more than 95% in all ethnicities among the interpretations according to the Canadian, GLI-Race and GLI-Other reference equations (online supplemental tables S3 and S4).

The source of the discrepancies in the interpretations is only due to difference in the values of the LLN among the three reference sets. The Canadian reference equations consistently predicted higher LLN values for FVC and FEV_1 than the GLI-Race and GLI-Other, except for NE Asian group (table 4, online supplemental figure S1). In addition, differences in the LLN for both FVC and FEV_1 were significantly different between ethnic groups when either GLI-Race or GLI-Other was applied (p<0.001, respectively). These differences were more pronounced when GLI-Race was applied to the non-Caucasian ethnic groups (Black, SE Asian, Other/Mixed) in comparison with Caucasian group. Although difference in the LLN value of FEV_1/FVC was also different between ethnic groups (p<0.001), the Canadian reference equation did not necessarily predict a higher LLN value for FEV_1/FVC.

Next, we conducted univariable and multivariable logistic regression analyses to identify the factors that contribute to the discordance in the interpretations of FVC and FEV_1 using the Canadian vs the two GLI-2012 reference sets. When GLI-Race was applied, sex, Black, SE Asian, Mixed/other ethnic groups were significantly associated with discrepancies in both FVC and FEV_1 interpretation (table 5). However, when GLI-Other was applied, only male sex was found to be significant factor in the discordance of the FVC interpretation, while age and weight were found to be factors significantly associated with discordance in the FEV_1 interpretation (table 6). We also repeated the same analyses by classifying Asian participants who could not be classified clearly as either NE Asian or SE Asian as SE Asian (model 2); the results

Table 2 Number of concordant and discordant interpretations for abnormal FVC (FVC<LLN) when using Canadian reference equation compared with GLI-Race or GLI-Other equation*

Ethnicity		GLI-Race				GLI-Other		
		FVC≥LLN	FVC<LLN	Total		FVC≥LLN	FVC<LLN	Total
All	Canadian				Canadian			
	FVC≥LLN	293 (72.2%)	1 (0.2%)	294	FVC≥LLN	294 (72.4%)	0 (0%)	294
	FVC<LLN	56 (13.8%)	56 (13.8%)	112	FVC<LLN	56 (13.8%)	56 (13.8%)	112
	Total	349	57	406	Total	350	56	406
Caucasian	Canadian				Canadian			
	FVC≥LLN	186 (81.2%)	0 (0%)	186	FVC≥LLN	186 (81.2%)	0 (0%)	186
	FVC<LLN	14 (6.1%)	29 (12.7%)	43	FVC<LLN	24 (10.5%)	19 (8.3%)	43
	Total	200	29	229	Total	210	19	229
Black	Canadian				Canadian			
	FVC≥LLN	7 (35.0%)	0 (0%)	7	FVC≥LLN	7 (35.0%)	0 (0%)	7
	FVC<LLN	10 (50.0%)	3 (15.0%)	13	FVC<LLN	5 (25.0%)	8 (40.0%)	13
	Total	17	3	20	Total	12	8	20
NE Asian	Canadian				Canadian			
	FVC≥LLN	13 (61.9%)	1 (4.8%)	14	FVC≥LLN	14 (66.7%)	0 (0%)	14
	FVC<LLN	2 (9.5%)	5 (23.8%)	7	FVC<LLN	3 (14.3%)	4 (19.0%)	7
	Total	15	6	21	Total	17	4	21
SE Asian	Canadian				Canadian			
	FVC≥LLN	66 (71.7%)	0 (0%)	66	FVC≥LLN	66 (71.7%)	0 (0%)	66
	FVC<LLN	20 (21.7%)	6 (6.5%)	26	FVC<LLN	14 (15.2%)	12 (13.0%)	26
	Total	86	6	92	Total	80	12	92
Other/Mixed	Canadian				Canadian			
	FVC≥LLN	21 (47.7%)	0 (0%)	21	FVC≥LLN	21 (47.7%)	0 (0%)	21
	FVC<LLN	10 (22.7%)	13 (29.5%)	23	FVC<LLN	10 (22.7%)	13 (29.5%)	23
	Total	31	13	44	Total	31	13	44

*We analysed the data by classifying Asian participants as NE Asian (model 1) if they could not be classified clearly as either NE Asian or SE Asian.
FVC, forced vital capacity; GLI, Global Lung Function Initiative; LLN, lower limit of normal; NE, North East; SE, South East.

were similar (online supplemental tables S5–S11). When the smoking history or types of spirometers were adjusted in the statistical models, the findings were similar (data not shown).

DISCUSSION

The present study evaluated the impact of different reference equation on interpretation of spirometry in a multi-ethnic cohort in a large Canadian city. It revealed that application of the all-Caucasian Canadian reference equations led to the over-interpretation of abnormal (<LLN) FVC and FEV_1 compared with GLI-Race and GLI-Other equations in all ethnic groups. The magnitude of the discordance was especially large in Black, Mixed/other and SE Asian population. The discordance was statistically significant even after adjusting for the key factors used for derivation of the reference values, that is, sex, age and height, when the Canadian reference equations were compared with GLI-Race reference equations. Although we observed discordance between ethnic groups when comparing the Canadian and the GLI-Other reference equations, these were not statistically significant; only male sex was found to be a significant factor in the discordance of FVC while age and weight were significant factors in the discordance of FEV_1. Unsurprisingly, FEV_1/FVC was highly consistent between the two equations.

Although some studies have shown disagreement,[21] GLI-Race reference equations generally fit for multiple race/ethnicities.[6–12] The current study revealed that the Canadian reference equations, compared with GLI-Race, led to higher rates of abnormal FVC and FEV_1, especially in non-Caucasian groups. Moreover, the LLN values of FVC and FEV_1 did not show perfect agreement even in the Caucasian group when comparing the Canadian and GLI-Race reference equations. For the Caucasian group, predicted LLN values of FVC and FEV_1 according to the Canadian reference equations were higher than those

Table 3 Number of concordant and discordant interpretations for abnormal FEV_1 (FEV_1<LLN) when using Canadian reference equation compared with GLI-Race or GLI-Other equation*

Ethnicity		GLI-Race				GLI-Other		
		FEV_1≥LLN	FEV_1<LLN	Total		FEV_1≥LLN	FEV_1<LLN	Total
All	Canadian				Canadian			
	FEV_1≥LLN	285 (70.2%)	1 (0.2%)	286	FEV_1≥LLN	286 (70.4%)	0 (0%)	286
	FEV_1<LLN	49 (12.1%)	71 (17.5%)	120	FEV_1<LLN	59 (14.5%)	61 (15.0%)	120
	Total	334	72	406	Total	345	61	406
Caucasian	Canadian				Canadian			
	FEV_1≥LLN	178 (77.7%)	0 (0%)	178	FEV_1≥LLN	178 (77.7%)	0 (0%)	178
	FEV_1<LLN	14 (6.1%)	37 (16.2%)	51	FEV_1<LLN	29 (12.7%)	22 (9.6%)	51
	Total	192	37	229	Total	207	22	229
Black	Canadian				Canadian			
	FEV_1≥LLN	8 (40.0%)	0 (0%)	8	FEV_1≥LLN	8 (40.0%)	0 (0%)	8
	FEV_1<LLN	10 (50.0%)	2 (10.0%)	12	FEV_1<LLN	5 (25.0%)	7 (35.0%)	12
	Total	18	2	20	Total	13	7	20
NE Asian	Canadian				Canadian			
	FEV_1≥LLN	12 (57.1%)	1 (4.8%)	13	FEV_1≥LLN	13 (61.9%)	0 (0%)	13
	FEV_1<LLN	1 (4.8%)	7 (33.3%)	8	FEV_1<LLN	5 (23.8%)	3 (14.3%)	8
	Total	13	8	21	Total	18	3	21
SE Asian	Canadian				Canadian			
	FEV_1≥LLN	62 (67.4%)	0 (0%)	62	FEV_1≥LLN	62 (67.4%)	0 (0%)	62
	FEV_1<LLN	16 (17.4%)	14 (15.2%)	30	FEV_1<LLN	12 (13.0%)	18 (19.6%)	30
	Total	78	14	92	Total	74	18	92
Other/Mixed	Canadian				Canadian			
	FEV_1≥LLN	25 (56.8%)	0 (0%)	25	FEV_1≥LLN	25 (56.8%)	0 (0%)	25
	FEV_1<LLN	8 (18.2%)	11 (25.0%)	19	FEV_1<LLN	8 (18.2%)	11 (25.0%)	19
	Total	33	11	44	Total	33	11	44

*We analysed the data by classifying Asian participants as NE Asian (model 1) if they could not be classified clearly as either NE Asian or SE Asian.
FEV_1, forced expiratory volume in 1 s; GLI, Global Lung Function Initiative; LLN, lower limit of normal; NE, North East; SE, South East.

derived from reference equations developed from data of Canadian Caucasian adults[22] and the third National Health and Nutrition Examination Survey (NHANES) III reference equations for Caucasian (non-Hispanic White) ethnic group[14] (online supplemental table S12). Thus, there is also the possibility of over-interpretation of abnormal lung function (<LLN) in Caucasian patients when the Canadian reference equations are applied. An Italian group compared the reference equations which were developed from several existing reference equations and concluded the necessity of applying reference equations derived from normal subjects who are as similar as possible to the study population being evaluated and using similar conditions of measurements.[23] These results emphasised the importance of using appropriate reference predictions that are most representative of the population in question.

The significance of considering race/ethnicity in lung function prediction equations is not limited to genetic/racial differences. Race and ethnicity are constructed by a complex combination of social, cultural and genetic factors.[24] It has been suggested that other factors, such as socioeconomic status and education, are associated with lung function.[25 26] For example, Asian-Indians born in USA have higher pulmonary function compared with immigrant Asian Indians, suggesting the effect of differing environmental conditions.[27] In other words, in multi-ethnic cities such as the one where this study was conducted or other large cosmopolitan cities around the world, the race/ethnicity categories may be ambiguous in many participants. A recent report of 567 Asian subjects living in the USA found that the GLI-Other reference equations adequately fitted spirometry data compared with the NHANES III and GLI-Race equations.[28] Comparison of the percent predicted lung function based on GLI-Other versus the GLI-Race equations in 3972 Black participants who participated in the NHANES III study showed that FEV_1 and FVC z-scores based on GLI-Other

Table 4 Differences in the calculated lower limit of normal (LLN) values for FVC, FEV_1 and FEV_1/FVC by race/ethnicities between the different reference equations*

	All	Caucasian	Black	NE Asian	SE Asian	Mixed/Other	P
n	406	229	20	21	92	44	
LLN FVC differences (L)							
Canadian equation—GLI-Race	0.29 (0.11, 0.49)	0.16 (0.08, 0.34)	0.66 (0.55, 0.95)†	−0.01 (−0.07, 0.07)†	0.49 (0.31, 0.71)†	0.47 (0.29, 0.62)†	<0.001
Canadian equation—GLI-Other	0.37 (0.25, 0.57)	0.38 (0.27, 0.60)	0.36 (0.29, 0.57)	0.35 (0.15, 0.51)	0.29 (0.15, 0.50)†	0.47 (0.29, 0.62)	<0.001
LLN FEV_1 differences (L)							
Canadian equation—GLI-Race	0.22 (0.15, 0.33)	0.20 (0.14, 0.23)	0.59 (0.54, 0.74)†	0.05 (−0.03, 0.10)†	0.33 (0.21, 0.46)†	0.35 (0.30, 0.46)†	<0.001
Canadian equation—GLI-Other	0.36 (0.27, 0.43)	0.38 (0.31, 0.44)	0.37 (0.33, 0.43)	0.26 (0.18, 0.37)†	0.23 (0.14, 0.35)†	0.35 (0.30, 0.46)	<0.001
LLN FEV_1/FVC difference (%)							
Canadian equation—GLI-Race	1.16 (−0.34, 1.90)	1.81 (1.31, 2.21)	1.07 (0.67, 1.40)†	−0.79 (−1.44,−0.11) †	−1.47 (−1.98,−0.80) †	0.10 (−0.24, 0.34)†	<0.001
Canadian equation—GLI-Other	0.13 (−0.35, 0.57)	0.03 (−0.50, 0.42)	0.12 (−0.16, 0.39)	0.28 (−0.51, 0.87)	0.57 (−0.07, 1.16)†	0.10 (−0.24, 0.34)	<0.001

Data are presented as median (IQR). P value is the result of overall by Kruskal-Wallis test.
*We analysed the data by classifying Asian participants as NE Asian (model 1) if they could not be classified clearly as either NE Asian or SE Asian.
†Bonferroni correction for multiple comparisons with p<0.05 compared with Caucasian group.
FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; GLI, Global Lung Function Initiative; LLN, lower limit of normal; NE, North East; SE, South East.

had more agreement between White and Black populations rather than GLI-Race equations.[29] Although the GLI-Race references led to lower FEV_1 and FVC in Black compared with the White populations, modelling of mortality risk was similar when GLI-Other was applied for lung function interpretation.[29] These studies have led some scholars to suggest that consideration of race/ethnicity may be counterproductive in the interpretation of PFT.[30] There is considerable debate as to whether race-specific equations or universal reference equations is superior.[31] Our data revealed the extent to which discrepancies occurs in each reference set according to race/ethnic groups by comparing the Canadian reference equations to both GLI-Race and GLI-Other of the GLI. Our findings suggest that the choice of reference equations should be carefully evaluated in different ethnic groups and considered when interpreting PFT.

In the current study, there was a difference in the interpretation of FEV_1 and FVC, but no discrepancy in interpretation of airway obstruction (FEV_1/FVC<LLN) between the reference equations. While some have argued that incorporation of FEV_1/FVC ratio in interpretations of PFTs may minimise the impact of race/ethnicity,[30] recent reports of clinical outcomes in preserved ratio impaired spirometry (PRISm), defined as FEV_1<80% predicted and FEV_1/FVC≥70%, suggest that the use of inappropriate reference equations could have clinical consequence. In two large population studies in the USA where the NHANES III reference equation was applied[31] and in a Belgian study where the GLI-Race equations were used, participants with PRISm have increased respiratory symptoms, mortality and faster FEV_1 decline.[32 33] Although some patients with PRISm transitioned to normal spirometry over time,[32 34] early identification of this group is important. The prevalence of PRISm decreases when post-bronchodilator data are assessed.[35] While our study evaluated pre-bronchodilator spirometry data, discrepancies in the interpretations of spirometry between the GLI and Canadian equations of restrictive patterns and PRISm would still be found.

There are several limitations to this study. First, GLI-2012 equations do not incorporate all race/ethnicities. For example, Black reference equations in GLI-2012 were generated by the data only from African Americans, and classification of non-African American Black individuals is inconsistent.[36 37] By classifying all Black participants in the GLI-2012 Black ethnic group, we are undoubtedly not accounting for the diversity of the people from the African continent and other locations, such as the Caribbean, the geographic origin of the many of the Black population in Toronto. For other races/ethnicities not covered by GLI-2012, reference equations for geographically and ethnically proximate groups were applied according to GLI-2012.[20] Second, race/ethnicity was self-reported, which may not be accurate enough for the clinical purposes as described in original GLI-2012 paper,[5] as inter-racial families are common in the greater Toronto area. Third, the number of participants included in some

Table 5 Univariable and multivariable analysis to identify the factors causing discordance in FVC and FEV_1 interpretation defined as abnormal (<LLN) in Canadian reference equations and normal (≥LLN) in GLI-Race*

		Univariable analysis			Multivariable analysis		
Variables		**OR**	**(95% CI)**	**P**	**OR**	**(95% CI)**	**P**
A) FVC							
Sex	Female	Ref			Ref		
	Male	3.32	(1.85 to 6.18)	<0.001	2.89	(1.24 to 6.92)	0.015
Age		0.997	(0.98 to 1.02)	0.78	0.99	(0.97 to 1.01)	0.46
Height		1.03	(0.998 to 1.06)	0.066	1.02	(0.97 to 1.07)	0.44
Weight		1.002	(0.99 to 1.01)	0.77	0.999	(0.98 to 1.02)	0.89
Race/ethnicity				<0.001			<0.001
	Caucasian	Ref			Ref		
	Black	15.36	(5.49 to 43.94)	<0.001	19.19	(6.48 to 59.10)	<0.001
	NEA	1.62	(0.24 to 6.38)	0.55	1.51	(0.22 to 6.53)	0.62
	SEA	4.27	(2.06 to 9.05)	<0.001	4.93	(2.04 to 12.42)	0.001
	Mixed/Other	4.52	(1.82 to 10.94)	0.001	4.06	(1.53 to 10.63)	0.004
B) FEV_1							
Sex	Female	Ref			Ref		
	Male	0.83	(0.44 to 1.51)	0.54	0.40	(0.16 to 0.95)	0.042
Age		0.996	(0.98 to 1.02)	0.73	1.003	(0.98 to 1.03)	0.80
Height		1.005	(0.97 to 1.04)	0.73	1.05	(0.998 to 1.11)	0.061
Weight		1.001	(0.99 to 1.01)	0.88	1.002	(0.99 to 1.02)	0.77
Race/ethnicity				<0.001			<0.001
	Caucasian	Ref			Ref		
	Black	15.36	(5.49 to 43.94)	<0.001	15.96	(5.62 to 46.52)	<0.001
	NEA	0.77	(0.041 to 4.14)	0.80	1.22	(0.064 to 7.14)	0.86
	SEA	3.23	(1.51 to 7.02)	0.003	5.85	(2.31 to 15.40)	<0.001
	Mixed/Other	3.41	(1.28 to 8.58)	0.010	5.04	(1.77 to 13.93)	0.002

*We analysed the data by classifying Asian participants as NE Asian (model 1) if they could not be classified clearly as either NE Asian or SE Asian.
FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; NEA, North East Asian; SEA, South East Asian.

ethnicity groups was not large although the ethnic mix is reflective of the population in the region. While our PFT Laboratory assesses 250 patients weekly, race/ethnicity data collection is not routine clinical practice. Only data from REB-approved studies where this information was collected were included in this paper. While we assessed mathematical considerations such as the difference in LLN values and the clinical concordance rate, larger-scale studies are required to validate our findings. Fourth, this study compared the Canadian and two GLI-2012 equations for interpretation of spirometry, rather than clinical outcomes or mortality risk. As the Canadian reference equations only considered the age range of 20–80 years,[16] we excluded 13 patients outside of this age range. As the GLI allowed for the modelling of complex nonlinear relationships over a wide age range, this omission will not have significant impact.

It should also be noted that the purpose of the current study is not to evaluate clinical outcomes associated with the labelling of the pulmonary function pattern but rather, to evaluate the discrepancies in the prevalence of pulmonary function abnormalities when different reference equations are applied to the same test data. The observed discrepancies in the interpretation of spirometry based on the choice of reference equations provide the rationale for ongoing discussion as the physiological pattern of PFT are key early factors that determine the clinical pathway of patients with respect to subsequent investigations, treatment and other therapeutic management. Thus, the use of lung function prediction equations should be carefully considered in each medical centre to ensure best practices in providing medical care to multiethnic populations.

CONCLUSION

By changing from the all Caucasian Canadian reference equations to the GLI-2012, the prevalence of

Table 6 Univariable and multivariable analysis to identify the factors causing discordance in FVC and FEV_1 interpretation defined as abnormal (<LLN) in Canadian reference equations and normal (≥ LLN) in GLI-Other*

		Univariable analysis			Multivariable analysis		
Variables		**OR**	**(95% CI)**	**P**	**OR**	**(95% CI)**	**P**
A) FVC							
Sex	Female	Ref			Ref		
	Male	4.03	(2.21 to 7.67)	<0.001	3.04	(1.35 to 7.01)	0.008
Age		0.99	(0.97 to 1.01)	0.42	0.99	(0.97 to 1.01)	0.35
Height		1.05	(1.02 to 1.08)	0.001	1.01	(0.97 to 1.06)	0.55
Weight		1.01	(1.002 to 1.02)	0.020	1.01	(0.996 to 1.03)	0.14
Race/ethnicity				0.13			0.15
	Caucasian	Ref			Ref		
	Black	2.85	(0.87 to 8.11)	0.062	3.15	(0.92 to 9.49)	0.050
	NEA	1.42	(0.32 to 4.61)	0.59	1.56	(0.33 to 5.62)	0.53
	SEA	1.53	(0.74 to 3.08)	0.24	2.06	(0.86 to 4.92)	0.10
	Mixed/Other	2.51	(1.07 to 5.60)	0.028	2.38	(0.93 to 5.84)	0.063
B) FEV_1							
Sex	Female	Ref			Ref		
	Male	0.90	(0.51 to 1.57)	0.72	0.77	(0.35 to 1.65)	0.50
Age		0.98	(0.96 to 0.99)	0.011	0.97	(0.95 to 0.995)	0.017
Height		1.01	(0.98 to 1.04)	0.47	0.999	(0.96 to 1.05)	0.98
Weight		1.01	(0.999 to 1.02)	0.063	1.01	(1.001 to 1.03)	0.028
Race/ethnicity				0.36			0.25
	Caucasian	Ref			Ref		
	Black	2.30	(0.71 to 6.45)	0.13	2.22	(0.67 to 6.33)	0.16
	NEA	2.16	(0.67 to 5.99)	0.16	3.04	(0.89 to 9.25)	0.059
	SEA	1.03	(0.49 to 2.08)	0.93	1.69	(0.72 to 3.87)	0.22
	Mixed/Other	1.53	(0.61 to 3.49)	0.33	1.67	(0.63 to 4.11)	0.28

*We analysed the data by classifying Asian participants as NE Asian (model 1) if they could not be classified clearly as either NE Asian or SE Asian.
FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; NEA, North East Asian; SEA, South East Asian.

restrictive lung physiological patterns as defined by FVC and FEV_1<LLN is expected to be decreased. The impact is particularly high in Black, SE Asian and Mixed/Other ethnic groups where the potential misclassification of restrictive defects would lead to unnecessary invasive tests such as CT imaging of the chest, lung biopsies and serological testing to arrive at a final diagnosis. Conversely, under-calling of abnormal spirometry as normal will result in missed diagnosis of lung disease. Our findings suggest that careful evaluation of the choice of reference equation and caution when interpreting lung function is required in different ethnic groups.

Acknowledgements The authors thank all the participants in this study.

Contributors HK compiled the clinical data, conducted the clinical review and analysis, and drafted the manuscript. AJ conducted the research, compiled the clinical data, conducted the clinical review, and drafted the manuscript. CN, HO and JKYW conducted the research, compiled the clinical data, developed quality control and assurance of the data and edited the manuscript. CMR developed the research protocol, ensured quality control of the pulmonary function data and edited the manuscript. C-WC developed the concept, study protocol, oversaw all aspects of the project and is the guarantor for the study.

Funding This work was supported by the Lung Health Foundation, Canadian Institutes for Health Research (CIHR; grant number 4518060 and the CIHR-Natural Sciences and Engineering Research Council Collaborative Health Research Programme; grant number 415013). HK is supported by the scholarship funded by the Nakayama Foundation for Human Science.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Consent obtained directly from patient(s)

Ethics approval The study was approved by the University Health Network Research Ethics Board under protocol numbers 19-5582 and 17-5652. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Data will be made available on request after review and approval by the respective institutions where the data are gathered and stored.

ORCID iDs
Haruna Kitazawa http://orcid.org/0000-0001-7234-1275
Annie Jiang http://orcid.org/0000-0002-6540-1409
Cynthia Nohra http://orcid.org/0000-0002-3607-5376
Honami Ota http://orcid.org/0000-0002-1935-3768
Joyce K Y Wu http://orcid.org/0000-0002-4113-2531
Chung-Wai Chow http://orcid.org/0000-0001-9344-8522

REFERENCES

1 Pellegrino R, Viegi G, Brusasco V, *et al*. Interpretative strategies for lung function tests. *Eur Respir J* 2005;26:948–68.
2 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;196:1463–72.
3 Cooper BG, Stocks J, Hall GL, *et al*. The global lung function initiative (Gli) network: bringing the world's respiratory reference values together. *Breathe* 2017;13:e56–64.
4 Kirkby J, Aurora P, Spencer H, *et al*. Stitching and switching: the impact of discontinuous lung function reference equations. *Eur Respir J* 2012;39:1256–7.
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;40:1324–43.
6 Bonner R, Lum S, Stocks J, *et al*. Applicability of the global lung function spirometry equations in contemporary multiethnic children. *Am J Respir Crit Care Med* 2013;188:515–6.
7 Langhammer A, Johannessen A, Holmen TL, *et al*. Global lung function initiative 2012 reference equations for spirometry in the Norwegian population. *Eur Respir J* 2016;48:1602–11.
8 Linares-Perdomo O, Hegewald M, Collingridge DS, *et al*. Comparison of NHANES III and ERS/GLI 12 for airway obstruction classification and severity. *Eur Respir J* 2016;48:133–41.
9 Huprikar NA, Holley AB, Skabelund AJ, *et al*. A comparison of global lung initiative 2012 with third National health and nutrition examination survey spirometry reference values. Implications in defining obstruction. *Ann Am Thorac Soc* 2019;16:225–30.
10 Brazzale DJ, Hall GL, Pretto JJ. Effects of adopting the new global lung function initiative 2012 reference equations on the interpretation of spirometry. *Respiration* 2013;86:183–9.
11 Ratomaharo J, Linares Perdomo O, Collingridge DS, *et al*. Spirometric reference values for Malagasy adults aged 18-73 years. *Eur Respir J* 2015;45:1046–54.
12 Hall GL, Thompson BR, Stanojevic S, *et al*. The global lung initiative 2012 reference values reflect contemporary Australasian spirometry. *Respirology* 2012;17:1150–1.
13 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;205:700–10.
14 Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999;159:179–87.
15 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;205:819–29.
16 Gutierrez C, Ghezzo RH, Abboud RT, *et al*. Reference values of pulmonary function tests for Canadian Caucasians. *Can Respir J* 2004;11:414–24.
17 Statistics Canada. *Toronto [Census metropolitan area], Ontario and Ontario [Province] (table). Census Profile. 2016 Census. Statistics Canada Catalogue no. 98-316-X2016001*. Ottawa, 2017. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E
18 Karunanayake CP*et al*. Reference values of pulmonary function tests for rural Canadians. *Int J Respir Pulm Med* 2015;2:021.
19 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:e70–88.
20 ERS e-Learning Resources. Global lung function initiative. what reference equations do I apply for non-Caucasians? Available: https://www.ers-education.org/guidelines/global-lung-function-initiative/faq/what-reference-equations-do-i-apply-for-non-caucasians/ [Accessed 11 Mar 2022].
21 Backman H, Lindberg A, Sovijärvi A, *et al*. Evaluation of the global lung function initiative 2012 reference values for spirometry in a Swedish population sample. *BMC Pulm Med* 2015;15:26.
22 Tan WC, Bourbeau J, Hernandez P, *et al*. Canadian prediction equations of Spirometric lung function for Caucasian adults 20 to 90 years of age: results from the Canadian obstructive lung disease (cold) study and the lung health Canadian environment (LHCE) study. *Can Respir J* 2011;18:321–6.
23 Pistelli F, Bottai M, Carrozzi L, *et al*. Reference equations for spirometry from a general population sample in central Italy. *Respir Med* 2007;101:814–25.
24 Scanlon PD, Shriver MD. "Race correction" in pulmonary-function testing. *N Engl J Med* 2010;363:385–6.
25 Raju PS, Prasad KVV, Ramana YV, *et al*. Influence of socioeconomic status on lung function and prediction equations in Indian children. *Pediatr Pulmonol* 2005;39:528–36.
26 Quanjer PH. Lung function, race and ethnicity: a conundrum. *Eur Respir J* 2013;41:1249–51.
27 Fulambarker A, Copur AS, Cohen ME, *et al*. Comparison of pulmonary function in immigrant vs US-born Asian Indians. *Chest* 2010;137:1398–404.
28 Zhang J, Hu X, Tian X, *et al*. Global lung function initiative 2012 reference values for spirometry in Asian Americans. *BMC Pulm Med* 2018;18:95.
29 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;205:723–4.
30 Bhakta NR, Balmes JR. A good fit versus one size for all: alternatives to race in the interpretation of pulmonary function tests. *Am J Respir Crit Care Med* 2022;205:616–8.
31 Kaminsky DA. Is there a role for using race-specific reference equations? Yes and NO. *Am J Respir Crit Care Med* 2022;205:746–8.
32 Wan ES, Fortis S, Regan EA, *et al*. Longitudinal phenotypes and mortality in preserved ratio impaired spirometry in the COPDGene study. *Am J Respir Crit Care Med* 2018;198:1397–405.
33 Wijnant SRA, De Roos E, Kavousi M, *et al*. Trajectory and mortality of preserved ratio impaired spirometry: the Rotterdam study. *Eur Respir J* 2020;55:1901217.
34 Wan ES, Hokanson JE, Regan EA, *et al*. Significant Spirometric transitions and preserved ratio impaired spirometry among ever smokers. *Chest* 2022;161:651–61.
35 Schwartz A, Arnold N, Skinner B, *et al*. Preserved ratio impaired spirometry in a spirometry database. *Respir Care* 2021;66:58–65.
36 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.
37 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;202:1724–7.