Respiration

Effects of Adopting the New Global Lung Function Initiative 2012 Reference Equations on the Interpretation of Spirometry

Danny J. Brazzale[a, b]   Graham L. Hall[c, d]   Jeffrey J. Pretto[b, e, f]

[a]Department of Respiratory and Sleep Medicine, Austin Hospital and [b]Institute for Breathing and Sleep, Melbourne, Vic., [c]Telethon Institute for Child Health Research, Centre for Child Health Research, University of Western Australia and [d]Respiratory Medicine, Princess Margaret Hospital for Children, Perth, W.A., [e]Department of Respiratory and Sleep Medicine, John Hunter Hospital and [f]School of Medicine and Public Health, University of Newcastle, Newcastle, N.S.W., Australia

Abstract
Background: The recently generated spirometry reference equations from the Global Lung Function Initiative (GLI2012) provide a long-awaited opportunity for the adoption of a globally applicable set of normal reference values. *Objective:* The aim of this study was to document the likely interpretative effects of changing from commonly used current spirometry reference equations to the GLI2012 equations on interpretation of test results in a clinical spirometry dataset. *Methods:* Spirometry results from 2,400 patients equally distributed over the age range of 5–85 years were obtained from clinical pulmonary function laboratories at three public hospitals. The frequency of obstruction [FEV$_1$/FVC below the lower limits of normal (LLN)] and spirometric restriction (FVC below the LLN) was assessed using the GLI2012, the National Health and Nutrition Assessment Survey (NHANES III), the European Community of Steel and Coal (ECSC) and the Stanojevic all-ages reference equations. *Results:* The rates of obstruction (range 20.0–28.5%) and spirometric restriction (range 14.2–25.8%) were similar across the four sets of reference equations. The highest level of agreement with the new GLI2012 equations was seen with the NHANES III equations (97.6% for obstruction and 93.6% for spirometric restriction) and the lowest with those from the ECSC (96.0 for obstruction and 92.0% for restriction). These data can be used to estimate likely diagnostic spirometry interpretation effects in the clinical setting when switching to GLI2012 spirometry reference data. *Conclusions:* We have found the effects on interpretation of changing to GLI2012 reference data to be minimal when changing from NHANES III and most significant when changing from ECSC reference data.

Copyright © 2013 S. Karger AG, Basel

Introduction

Spirometry is the most widely used respiratory function test [1] and it plays a vital role in the diagnosis and management of respiratory disease [2–4]. An important step in the interpretation of spirometry is the comparison of an individual patient's results with that of reference values derived from a healthy population sample. However, the selection of appropriate reference values to en-

Danny J. Brazzale
Respiratory Laboratory
Austin Hospital, PO Box 5555
Heidelberg, VIC 3084 (Australia)
E-Mail danny.brazzale @austin.org.au

sure accurate interpretation is problematic. This is at least partly due to the wide choice of reference value studies available, with at least 53 published between the years of 1995 and 2004 alone [5].

These various reference equations produce surprisingly different predicted spirometry values for the same individual [6, 7] leading to significant differences in interpretation of the same set of spirometry results [8–10]. These interpretative differences could potentially lead to significant changes in clinical outcomes, not only in identifying the presence or absence of disease, but also in the classification of disease severity.

The most recent American Thoracic Society/European Respiratory Society (ATS/ERS) guidelines [5] recommend the use of the National Health and Nutrition Assessment Survey (NHANES III) [11] reference equations within the USA, and these normal reference values are widely used within the Australasian region. The NHANES III data have since been supplemented with other data to produce new prediction equations with a wider applicability in terms of age range and a smoother transition from childhood to adult values [12, 13]. While the ATS/ERS make no specific recommendations for spirometry reference equations outside of the USA, the European Community of Steel and Coal (ECSC) predicted values [14] are commonly used in Europe [5].

The recent spirometry reference values from the Global Lung Function Initiative (GLI2012) [15] represent the largest collection of normal spirometric data ever collated. This rigorously collected and analysed dataset provides a long-awaited opportunity for the adoption of a globally applicable set of reference equations for spirometry. The GLI2012 reference values have been endorsed by multiple professional respiratory societies worldwide and are therefore likely to be widely implemented for routine clinical use.

Implementation of any new reference values into clinical practice can be problematic in that the interpretation of the presence and severity of lung function abnormalities may be altered. Given that these interpretation differences can have significant clinical consequences, it is important to document the likely effects of switching reference values on the interpretation of spirometry results [16]. While these reference data have been shown to be applicable in a contemporary Australasian Caucasian population of normal subjects [17], assessment in a group of patient results will provide additional useful information.

Similar to in our previous analysis [10], the aim of this study was to document the likely interpretative effects of making a change in spirometric reference values. Specifically, we wished to describe the effect of changing from commonly used current reference equations, i.e. NHANES III [11], Stanojevic all-ages [13] or ECSC [14], to the GLI2012 equations [15] on the interpretation of test results in the clinical setting.

Materials and Methods

Data were obtained from clinical pulmonary function laboratory databases at three Australian public hospitals (Austin Hospital, Vic., John Hunter Hospital, N.S.W. and Princess Margaret Hospital for Children, W.A.). All three are large university-affiliated tertiary referral centres involved in managing a broad range of respiratory diseases. In order to minimise potential biases and thereby enable more widespread applicability of results, we analysed equal patient numbers from each decade from 5 to 85 years of age with equal gender representation in each decade. Given that a sample size of at least 300 subjects is required to confirm that a population is truly different from a reference group [18], we obtained 300 spirometry records for each decade resulting in a total of 2,400 test results. Consecutive patients' spirometry data from the three laboratories were retrieved backwards from July 2011 until the required patient numbers were obtained (March 2009). Our dataset was limited to Caucasian subjects so that race-appropriate predicted values could be compared. All testing was performed in accordance with the relevant guidelines [19] with only pre-bronchodilator data being analysed and only the most recent test results being included on patients who had been tested on more than one occasion. Any tests not meeting ATS/ERS guidelines [19] were excluded. Local ethics committee approval was obtained.

In order to span the paediatric-adult age range, the adult reference equations of the ECSC were combined with the paediatric reference equations from Quanjer et al. [20], while the NHANES III equations were combined with those from Wang et al. [21]. The Stanojevic equations already cover the paediatric and adult age range. Given that this is common practice in the clinical setting, equations were extrapolated to cover the age range of the older patients. It is, however, not possible to extend the reference equations of Stanojevic beyond the age of 80 years, so this specific analysis was limited to patients aged between 5 and 80 years (reducing the dataset for this comparison to 2,287). Similarly, the predictions by Wang et al. [21] are not possible for 5-year-old children, so this analysis was also limited (n = 2,391).

Mean predicted values and lower limits of normal (LLN) based on the 95th percentile were calculated for each set of reference equations. An obstructive ventilatory defect was defined as an FEV_1/FVC ratio below the LLN. While it is recognised that a reduced TLC is required to define true restriction, a reduced FVC is commonly interpreted as suggestive of a restrictive defect [22]. For this reason, spirometric restriction was defined as an FVC below the LLN (irrespective of the FEV_1/FVC ratio).

Statistical Analysis

Differences for predicted values and LLNs across the four equations were analysed using one-way ANOVA. Post hoc comparisons between the GLI2012 equations and the other equations

Table 1. Summary of measured spirometric data and predicted values for the entire patient group

	Mean ± SD (range) of measured data (n = 2,400)	Mean (LLN) predicted value (n = 2,278)			
		NHANES III*	ECSC*	Stanojevic*	GLI2012
FEV_1 (L)	2.60±1.09 (0.40–6.33)	3.131 (2.485)	2.979 (2.301)	3.125 (2.463)	3.135 (2.442)
FVC (L)	3.47±1.31 (0.42–7.98)	3.909 (3.145)	3.584 (2.793)	3.906 (3.123)	3.860 (3.018)
% FEV_1/FVC	75±13 (24–100)	80.42 (70.69)	80.33 (69.50)	81.13 (67.53)	81.89 (70.08)

There are 122 patients outside the prediction age range of 6–80 years reducing the predicted value calculations from 2,400 to 2,278.
* $p < 0.001$ for mean predicted values and LLN for all comparisons with GLI2012.

Table 2. Patients classified as obstructed (FEV_1/FVC below the LLN) for each of the four sets of predicted values, stratified by gender

Reference equation	Total n	Males n (%)	Females n (%)	Overall %
NHANES III	2,391	353 (29.5)	328 (27.4)	28.5
ECSC	2,400	299 (24.9)	286 (23.8)	24.4
Stanojevic	2,287	207 (18.2)	251 (21.9)	20.0
GLI2012	2,400	331 (27.6)	300 (25.0)	26.3

Table 3. Patients classified as having spirometric restriction (FVC below the LLN) for each of the four sets of predicted values, stratified by gender

Reference equation	Total n	Males n (%)	Females n (%)	Overall %
NHANES III	2,391	328 (27.4)	288 (24.1)	25.8
ECSC	2,400	211 (17.6)	129 (10.8)	14.2
Stanojevic	2,287	314 (27.6)	253 (22.0)	24.8
GLI2012	2,400	262 (21.8)	230 (19.2)	20.5

were then made using two-tailed paired Student's t tests. Levels of agreement in interpretation between different equations were based on the proportion of matching classifications, and also using the kappa statistic. SPSS (IBM, Version 17) was used for statistical analyses.

Results

The data collection technique meant that the average age was 45 years with 150 people of each gender in each decade from 5 to 85 years of age. The median height of all patients was 166 cm, with an interquartile range of 158–174 cm. Table 1 shows the spirometric values for all patients (1,200 males) and the mean predicted values and LLNs for the 2,278 patients (aged 6–80 years) for which it was possible to calculate predicted values using all four reference equations. The mean predicted values and LLNs for FEV_1, FVC and FEV_1/FVC for the NHANES III, Stanojevic and ECSC equations were statistically different from those produced by the GLI2012 equations. The mean differences between the GLI2012 equations and the other reference equations for FEV_1 ranged between 4 ml (NHANES III) and 156 ml (ECSC). The mean difference

for FVC was between 46 ml (Stanojevic) and 276 ml (ECSC).

Table 2 and table 3 summarise the number of test results designated as showing obstruction and spirometric restriction for each of the reference equations used. The highest rate of airflow obstruction was found using the reference equations of NHANES III (28.5%), while the lowest rate was found using the Stanojevic equations (20.0%). Changing to GLI2012 equations caused an 8% reduction, a 32% increase and an 8% increase in rates of obstructive interpretation when compared to the NHANES III, Stanojevic and ECSC equations, respectively. The spirometric restriction rate was highest using the NHANES III equations and lowest using the equations from the ECSC. Changing from the ECSC reference equations to the GLI2012 equations causes a 45% increase in the rate of restrictive spirometry classification. The corresponding decreases when changing from the NHANES III or Stanojevic equations to the GLI2012 equations are 20 and 19%, respectively.

We examined whether gender had a significant impact on the number of people who were classified differently by the different equations; in terms of airflow obstruction, males indeed tended to be more likely to be classified

Table 4. Levels of agreement for obstruction and spirometric restriction comparing each of the three older reference value sets with GLI2012

Equation vs. GLI2012	% agreement (kappa statistic)	
	$FEV_1/FVC < LLN$	$FVC < LLN$
NHANES III	97.6 (0.94)	93.6 (0.82)
ECSC	96.0 (0.89)	92.0 (0.72)
Stanojevic	93.4 (0.82)	95.7 (0.88)

Fig. 1. Differing rates of airflow obstruction ($FEV_1/FVC < LLN$) in different age categories (children, adults and older people), using the four sets of predicted values.

Fig. 2. Differing rates of spirometric restriction ($FVC < LLN$) in different age categories (children, adults and older people), using the four sets of predicted values.

Fig. 3. Histogram showing distribution of the degree of impairment for those with abnormal spirometry using the four sets of predicted values.

differently. Of the 71 individuals who were obstructed using GLI2012 equations but not ECSC equations, 47 (66%) were male. Of the 148 individuals classified as obstructed using GLI2012 but not Stanojevic equations, 108 (73%) were male.

Table 4 summarises the levels of agreement with the GLI2012 equations in terms of interpretation of obstruction and spirometric restriction. Overall, these data show that more than 90% of test interpretations are in agreement when the other three reference equations are compared to GLI2012. Combining the agreement for both obstructive and restrictive pattern spirometry, agreement with GLI2012 is highest with NHANES III and lowest with ECSC.

Age-related differences in obstructive and spirometric restriction interpretation rates were also examined. Generally, those individuals classified differently by the other equations compared to the GLI2012 equations were evenly spread across the decades. The only exceptions were that the patients classified as spirometry-restricted using the NHANES III (n = 138) and Stanojevic reference equations (n = 99), but not with the GLI2012 equations, tended to be older with 83 and 79%, respectively, being over the age of 45 years. The differences in rates of obstruction (fig. 1) and spirometric restriction (fig. 2) were also examined when the patient data were grouped into 3 age categories: children (6–20 years), adults (20–65 years) and older people (65–80 years). While the differences in rates of obstruction were similar across the 3 age groups, the rates of spirometric restriction are more similar in the youngest age group and the differences between the equations become more evident above the age of 20 years.

We also examined the effects of changing predicted values on the categorization of disease severity using the ATS/ERS guidelines [5]. Figure 3 illustrates how the different reference equations categorize disease severity in those with abnormal spirometry (FEV_1/FVC below the LLN and/or FVC below the LLN).
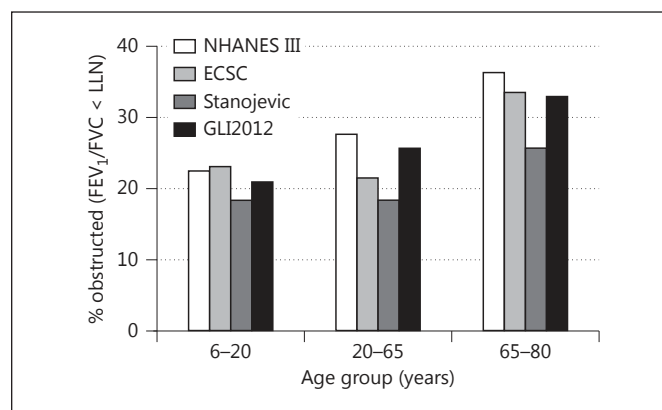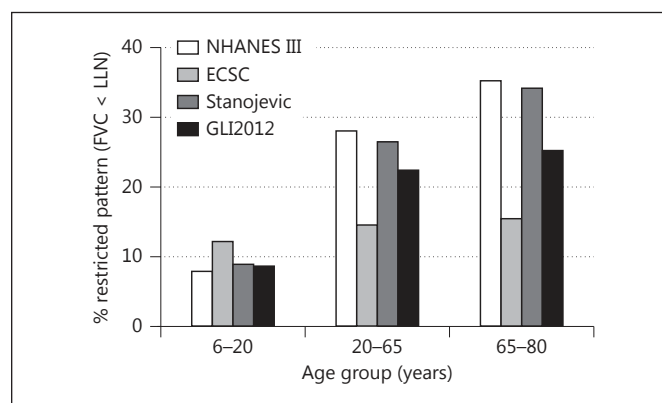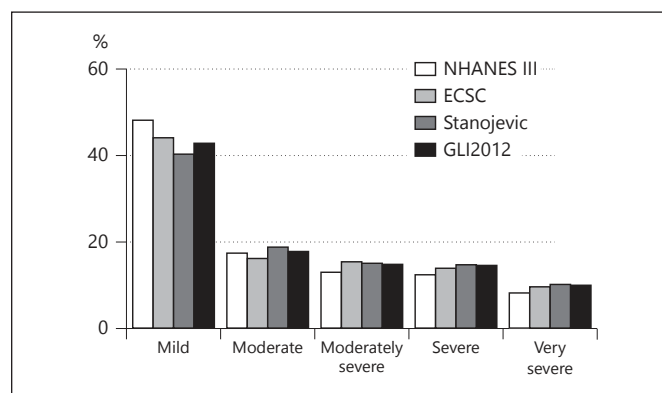
Discussion

The publication of the GLI2012 reference data may cause many respiratory laboratory directors to reappraise the choice of reference equations used for spirometry. Changing reference equations is problematic for any clinical laboratory, with the characterisation of the presence of disease and disease severity potentially altered. For this reason, it is important to examine the potential effects of changes in reference value selection on the interpretation of test results. These data provide an overview of the effects on the interpretation of spirometry of changing reference values to the new GLI2012 equations in a clinical population undergoing spirometry testing.

Our data show that the number of patients classified as having obstructive or restrictive spirometry will be altered significantly on changing from the ECSC to the GLI2012 equations. This alteration was smaller if changing from the NHANES III or Stanojevic equations. Previous studies [8–10] have evaluated the effects on spirometric interpretation of changing reference values and have found the impact can be significant. Our data show a similar pattern, with the number of subjects affected dependent on the equations previously used and also on whether restrictive or obstructive rates are being examined.

Given that the Stanojevic equations predominantly utilise data from the NHANES III dataset in the adult age range, it is not surprising that these 2 studies show similar levels of agreement when compared to GLI2012, in terms of mean predicted values. The lower predicted values reported here for the ECSC equations are similar to previous reports [7, 10] and reinforce further the perception that the ECSC predicted values do not match contemporary healthy spirometry data.

We report a change in prevalence rates of airflow obstruction of up to 32% between the 3 newer studies (NHANES III, Stanojevic and GLI2012), largely due to differences in LLN rather than in mean values for FEV_1/FVC (table 1). We suspect that these differences are partly due to different statistical approaches used in both the GLI2012 and Stanojevic studies. Both studies used the lambda-mu-sigma method to derive predicted and LLN values which allows modelling of variability and skewness of data and uses splines to allow for the interactive effects of age, height and sex [23]. Furthermore, the large sample size of GLI2012 (n = 57,395) is the likely cause of the tighter LLN for FEV_1/FVC compared with Stanojevic (n = 3,598). Comparing the FEV_1/FVC LLN of GLI2012 and
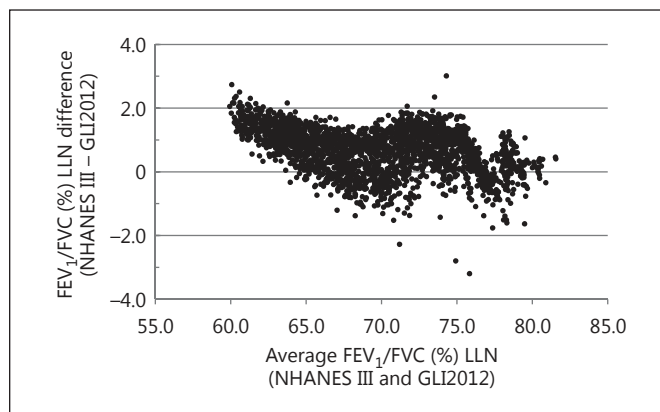

Fig. 4. Differences (n = 2,391) between LLN FEV_1/FVC for NHANES III compared with GLI2012 illustrating the higher LLN for NHANES III, particularly for low FEV_1/FVC values.

NHANES III reveals a systematic difference at lower values (fig. 4); as a result, this effect is more likely to be evident for elderly patients.

Figure 3 reveals a tendency for the NHANES III equations to result in the interpretation of more mild disease than with the other equations. This is largely a reflection of the overall higher rates of spirometric abnormalities using NHANES III – those individuals just falling into the abnormal range would commonly be categorized as having mild abnormality.

While the comparison of average predicted values from different reference studies can elucidate population-level effects of spirometry outcomes, it cannot adequately describe the consequences of changing reference ranges at an individual patient level. We believe that using real spirometric data from patients undergoing lung function testing for clinical purposes allows more appropriate characterisation of the likely effects of changing reference data. This approach has allowed us to quantify not only the average change in predicted values (which could be estimated by simple modelling), but, more importantly, the likely effects on clinical interpretation of adopting new spirometry reference values. Our analysis has also revealed that there is little change in the severity rating of the spirometric abnormalities detected.

One clear advantage of the GLI2012 equations is that they are applicable over a wider age range than any other published reference data (3–95 years of age) and that robust predicted values for multiple ethnic groups are available. This ensures that patients being tested longitudinally can be evaluated using a single reference dataset, thus

removing the potential errors associated with switching equations [24], which is particularly problematic in the childhood to adulthood transition [25]. It also minimises the need for the potentially erroneous practice of extrapolating prediction equations beyond the age range in which the data was collected.

One potential limitation of our analysis is the assumption that the LLN clearly separates normal from abnormal values. We acknowledge that this is a simplistic approach and that a more appropriate interpretative approach would be to apply caution when results are in this marginal region close to the LLN. It is likely that most of the test results which changed between reference sets were close to the LLN range. In these cases, the veracity of the interpretation is less certain and additional information, such as pre-test probability, clinical history, symptoms and other test results would assist with interpretation. Despite this obvious limitation, the LLN are the values that underpin interpretative strategies (such as those recommended by ATS/ERS [5]) and we believe that our approach is justified in describing the likely effects of changing predicted values.

Another potential limitation of the study is the fact that paediatric equations were required to be combined with adult equations in order to cover the entire age range. We believe that this approach is justified, given that it is commonly adopted in clinical laboratories where the patients tested span the paediatric-adult age range. The comparisons we have made by combining paediatric and adult equations have allowed us to examine the effect of changing predicted values on a wider age range. We did also consider the possibility that some of the comparisons may have been influenced by the patients who were excluded because of the age limitations of the equations (i.e. patients aged 5 and >80 years). However, when comparisons were repeated that took this into account, no significant differences were found in the results.

The GLI2012 spirometry reference values provide a unique approach to enabling the accurate prediction of normal spirometric values, and their widespread uptake into clinical practice appears likely. We have shown that changing to the GLI2012 values will impact on the interpretation of spirometry test results and that the magnitude and direction of change is dependent upon which reference data are being currently used at the level of individual practice. Our analysis provides an all-age summary of the changes to be expected, and we believe that this information might assist in the uptake of these reference data. This would facilitate further standardisation in the performance and interpretation of this valuable diagnostic tool.

Acknowledgements

References

1 Crapo RO: Pulmonary function testing. New Eng J Med 1994;331:25–30.
2 Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, Colby TV, Cordier JF, Flaherty KR, Lasky JA, et al: An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. Am J Resp Crit Care Med 2011;183:788–824.
3 GINA: Global Strategy for Asthma Management and Prevention, 2011. http://www.ginasthma.org/uploads/users/files/GINA_Report2011_May2014.pdf.
4 Pauwels RA, Buist AS, Calverley PM, Jenkins CR, Hurd SS: Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. Am J Respir Crit Care Med 2001;163:1256–1276.
5 Pellegrino R, Viegi G, Brusasco V, Crapo R, Burgos F, Casaburi R, Coates A, van der Grinten C, Gustafsson P, Hankinson J, et al: Interpretative strategies for lung function tests. Eur Respir J 2005;26:948–968.
6 Roca J, Burgos F, Sunyer J, Saez M, Chinn S, Anto JM, Rodriguez-Roisin R, Quanjer PH, Nowak D, Burney P: References values for forced spirometry. Eur Resp J 1998;11:1354–1362.
7 Langhammer A, Johnsen R, Gulsvik A, Holmen TL, Bjermer L: Forced spirometry reference values for Norwegian adults: the Bronchial Obstruction in Nord-Trondelag Study. Eur Resp J 2001;18:770–779.
8 Sood A, Dawson BK, Henkle JQ, Hopkins-Price P, Quails C: Effect of change of reference standard to NHANES III on interpretation of spirometric 'abnormality'. Int J COPD 2007;2:361–367.
9 Collen J, Greenburg D, Holley A, King CS, Hnatiuk O: Discordance in spirometric interpretations using three commonly used reference equations vs. the National Health and Nutrition Examination Study III. Chest 2008;134:1009–1016.
10 Brazzale DJ, Upward AL, Pretto JJ: Effects of changing reference values and definition of the normal range on interpretation of spirometry. Respirology 2010;15:1098–1103.
11 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;159:179–187.
12 Stanojevic S, Wade A, Stocks J, Hankinson J, Coates AL, Pan H, Rosenthal M, Corey M, Lebecque P, Cole TJ: Reference ranges for spirometry across all ages: a new approach. Am J Respir Crit Care Med 2008;177:253–260.
13 Stanojevic S, Wade A, Cole TJ, Lum S, Custovic A, Silverman M, Hall GL, Welsh L, Kirkby J, Nystad W, et al: Spirometry centile charts for young Caucasian children: the Asthma UK Collaborative Initiative. Am J Respir Crit Care Med 2009;180:547–552.

14 Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC: Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Eur Respir J 1993;16:5–40.

15 Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, Enright PL, Hankinson JL, Ip MS, Zheng J, Stocks J: Multi-ethnic reference values for spirometry for the 3–95-year age range: the Global Lung Function 2012 equations. Eur Respir J 2012;40:1324–1343.

16 Stanojevic S, Hall GL: Reference values for spirometry: the way forward for our patients. Respirology 2011;16:869.

17 Hall GL, Thompson BR, Stanojevic S, Abramson MJ, Beasley R, Coates A, Dent A, Eckert B, James A, Filsell S, et al: The Global Lung Initiative 2012 reference values reflect contemporary Australasian spirometry. Respirology 2012;17:1150–1151.

18 Quanjer PH, Stocks J, Cole TJ, Hall GL, Stanojevic S: Influence of secular trends and sample size on reference equations for lung function tests. Eur Respir J 2011;37:658–664.

19 Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P, van der Grinten CP, Gustafsson P, et al: Standardisation of spirometry. Eur Respir J 2005; 26:319–338.

20 Quanjer PH, Borsboom GJ, Brunekreef B, Zach M, Forche G, Cotes JE, Sanchis J, Paoletti P: Spirometric reference values for white European children and adolescents: Polgar revisited. Pediatr Pulmonol 1995;19: 135–142.

21 Wang X, Dockery DW, Wypij D, Fay ME, Ferris BG Jr: Pulmonary function between 6 and 18 years of age. Pediatr Pulmonol 1993; 15:75–88.

22 Aaron SD, Dales RE, Cardinal P: How accurate is spirometry at predicting restrictive pulmonary impairment? Chest 1999;115:869–873.

23 Cole TJ, Stanojevic S, Stocks J, Coates AL, Hankinson JL, Wade AM: Age- and size-related reference ranges: a case study of spirometry through childhood and adulthood. Stat Med 2009;28:880–898.

24 Quanjer PH, Hall GL, Stanojevic S, Cole TJ, Stocks J: Age- and height-based prediction bias in spirometry reference equations. Eur Respir J 2012;40:190–197.

25 Kirkby J, Aurora P, Spencer H, Rees S, Sonnappa S, Stocks J: Stitching and switching: the impact of discontinuous lung function reference equations. Eur Respir J 2012;39:1256–1257.