MULTI-ETHNIC REFERENCE VALUES FOR SPIROMETRY FOR THE 3-95 YEAR AGE RANGE: THE GLOBAL LUNG FUNCTION 2012 EQUATIONS

Report of the Global Lung Function Initiative, ERS Task Force to establish improved Lung Function Reference Values, endorsed by the ATS, ANZSRS, TSANZ, APSR and the ACCP

Endorsed by the European Respiratory Society (ERS), American Thoracic Society (ATS), Australian and New Zealand Society of Respiratory Science (ANZSRS), Asian Pacific Society for Respirology (APSR), Thoracic Society of Australia and New Zealand (TSANZ) and the American College of Chest Physicians (ACCP).

Philip H. Quanjer[1], Sanja Stanojevic[2], Tim J. Cole[3], Xaver Baur[4], Graham L. Hall[5], Bruce Culver[6], Paul L. Enright[7], John L. Hankinson[8], Mary S.M. Ip[9], Jinping Zheng[10], Janet Stocks[11] and the ERS Global Lung Function Initiative (GLI)

The ERS Global Lungs Initiative (see www.lungfunction.org):

Chairs:
J. Stocks, X. Baur, G.L. Hall, B. Culver

Analytical team:
P.H. Quanjer, S. Stanojevic, T.J. Cole, J. Stocks

Additional members of Steering committee: J.L. Hankinson, P.L. Enright, J.P. Zheng, M.S.M. Ip

Statistical reviewer:
C. Schindler

[1]Department of Pulmonary Diseases, and Department of Paediatrics, Erasmus Medical Centre, Erasmus University, Rotterdam, the Netherlands; [2]Portex Respiratory Unit, UCL Institute of Child Health, London, UK, and Child Health Evaluative Sciences & Respiratory Medicine, The Hospital for Sick Children, Toronto, Canada; [3]MRC Centre of Epidemiology for Child Health, UCL Institute of Child Health, London, UK; [4]Universitätsklinikum Hamburg-Eppendorf, Zentralinstitut für Arbeitsmedizin und Maritime Medizin, Hamburg, Germany; [5]Respiratory Medicine, Princess Margaret Hospital for Children, and School of Paediatric and Child Health and Telethon Institute for Child Health Research, Centre for Child Health Research, University of Western Australia, Perth, Australia; [6]Division of Pulmonary and Critical Care Medicine, Department of Medicine, University of Washington, Seattle, Washington, USA; [7]Division of Public Health Sciences, The University of Arizona, Tucson, Arizona, USA; [8]Hankinson Consulting, Athens, Georgia, USA; 9Department of Medicine, The University of Hong Kong, Queen Mary Hospital, Hong Kong, China; [10]Guangzhou Institute of Respiratory Disease, State Key Laboratory of Respiratory Disease, Guangzhou, China; [11]Portex Respiratory Unit, UCL Institute of Child Health, London, UK.

Corresponding author:
Philip Quanjer, e-mail: pquanjer@xs4all.nl

Persons and centres contributing data to this manuscript:
O.A. Al-Rawas, Department of Medicine, College of Medicine and Health, Sciences, Sultan Qaboos University, Muscat, Sultanate of Oman; H.G.M. Arets; Department of Pediatric Pulmonology, Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, The Netherlands; C. Bárbara , The Portuguese Society of Pneumology, Lisbon, Portugal; R.G. Barr, Columbia University Medical Center, New York, NY, USA and the MESA study; E. Bateman, University of Cape Town Lung Institute, Cape Town, South Africa; C.S. Beardsmore, Department of Infection, Immunity and Inflammation (Child Health), University of Leicester, Leicester, UK; H. Ben Saad, Laboratory of Physiology, Faculty of Medicine, Sousse, University of Sousse, Tunisia; B. Brunekreef, Institute for Risk Assessment Sciences, Universiteit Utrecht, Utrecht, the Netherlands; P.G.J. Burney, National Heart and Lung Institute, Imperial College, London; R.B. Dantes, Philippine College of Chest Physicians, Manila, Philippines; W. Dejsomritrutai, Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand; D. Dockery, Department of Environmental Health, Department of Epidemiology, Boston, MA, USA; H. Eigen, Section of Pulmonology and Intensive Care, James Whitcomb Riley Hospital for Children, Indiana University School of Medicine, Indianapolis, IN, USA; E. Falaschetti, [Health Survey for England 1995-1996 (HSE)], International Centre for Circulatory Health, National Heart and Lung Institute, Imperial College, UK; B. Fallon, Respiratory Laboratory, Nepean Hospital, Penrith, Australia; A. Fulambarker, Pulmonary Division, Rosalind Franklin University of Medicine and Science, The Chicago Medical School, Chicago, IL, USA; M. Gappa [LUNOKID study group], Children's Hospital and Research Institute, Marienhospital Wesel, Germany; M.W. Gerbase, Division of Pulmonary Medicine, University Hospitals of Geneva, Geneva, Switzerland, and the SAPALDIA cohort study; T. Gislason, Landspitali University Hospital, Dept. of Allergy, Respiratory Medicine and Sleep, Reykjavik, Iceland; M. Golshan, Bamdad Respiratory Research Institute, Isfahan, Iran; C.J. Gore, Physiology Department, Australian Institute of Sport, Belconnen, Australia; A. Gulsvik, Department of Thoracic Medicine, Institute of Medicine, University of Bergen, Bergen, Norway; G.L. Hall, Respiratory Medicine, Princess Margaret Hospital for Children, Perth, Australia; J.L. Hankinson, [NHANES, NHANES III Special data sets], Hankinson Consulting, Valdosta, GA, USA; A.J. Henderson, [ALSPAC, http://www.bris.ac.uk/alspac], University of Bristol, Bristol, UK; E. Hnizdo, Division of Respiratory Disease Studies , National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Morgantown, WV, USA [NHANES IV]; M.S.M. Ip, Dept of Medicine, The University of Hong Kong, Pokfulam, Hong Kong SAR, China; C. Janson, Department of Medical Sciences: Respiratory Medicine & Allergology, Uppsala University, Sweden; C. Jenkins, Woolcock Institute of Medical Research, Sydney, Australia; A. Jithoo, University of Cape Town Lung Institute, Cape Town, South Africa; S. Karrasch, Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, Hospital of the Ludwig-Maximilians-University, Munich, Germany [KORA study]; G.S. Kerby (Lung Function Measures in Preschool Children with Cystic Fibrosis study

group), University of Colorado School of Medicine and Children's Hospital Colorado, Aurora, CO, USA; J. Kühr, Klinik für Kinder- und Jugendmedizin, Städtisches Klinikum Karlsruhe, Karlsruhe, Germany; S. Kuster, Lungenliga Zürich, Zürich, Switzerland [LuftiBus study]; A. Langhammer, [The HUNT Study] HUNT Research Centre, NTNU, Verdal, Norway; S. Lum, Portex Respiratory Unit, UCL, Institute of Child Health, London, UK; D.M. Mannino, University of Kentucky, Lexington, Kentucky, USA; G. Marks, Woolcock Institute of Medical Research, Sydney, Australia; A. Miller, Beth Israel Medical Center, New York, NY, USA; G. Mustafa, Department of Pediatrics, Nishtar Medical College, Multan, Pakistan; E. Nizankowska-Mogilnicka, Division of Pulmonary Diseases, Department of Medicine, Jagiellonian University School of Medicine, Cracow, Poland; W. Nystad, Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway; Y-M. Oh [Korean NHANES], Department of Pulmonary and Critical Care Medicine; Asthma Center; Clinical Research Center for Chronic Obstructive Airway Diseases, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea; W-H. Pan, Institute of Medical Sciences, Academica Sinica, Taipei, Taiwan; R. Pérez-Padilla, Instituto Nacional de Enfermedades Respiratorias, Mexico DF, Mexico, [PLATINO study]; P. Piccioni, SC Pneumologia CPA ASL Torino 2, Torino, Italy; F. Pistelli, Pulmonary and Respiratory Pathophysiology Unit, Cardiothoracic Department, University Hospital of Pisa and Pulmonary Environmental Epidemiology Unit, CNR Institute of Clinical Physiology, Pisa, Italy; Prasad KVV, Government Vemana Yoga Research Institute, Ameerpet, Hyderabad, India; P.H. Quanjer, Department of Pulmonary Diseases, and Department of Pediatrics, Erasmus Medical Centre, Erasmus University, Rotterdam, The Netherlands; M. Rosenthal, Royal Brompton Hospital, London, UK; H. Schulz, Institute of Epidemiology I, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany [KORA study]; S. Stanojevic, Portex Respiratory Unit, UCL Institute of Child Health, London, UK, [Asthma UK Growing Lungs Initiative (http://www.growinglungs.org.uk)], and Child Health Evaluative Sciences & Respiratory Medicine, The Hospital for Sick Children, Toronto; J.B. Soriano, Program of Epidemiology and Clinical Research, CIMERA, Recinte Hospital Joan March, Illes Balears, Spain [Framingham study]; W.C. Tan, iCapture Center for Cardiovascular and Pulmonary Research, University of British Columbia, Vancouver, BC, Canada; W. Tomalak, Dept. Physiopathology of Respiratory System, National Institute for TBC & Lung Dis., Rabka Branch, Poland; S.W. Turner [The SEATON study group], Department of Child Health, University of Aberdeen, Aberdeen, UK; Vilozni, Pediatric Pulmonary Units of The Edmond and Lili Safra Children's Hospital, Sheba Medical Center Ramat-Gan, affiliated with the Sackler Medical School, Tel-Aviv University, Tel-Aviv, Israel; H. Vlachos, Department of Pediatrics, Division of Respiratory Medicine, University of Sherbrooke, Quebec, Canada; S. West, Respiratory Function Laboratory, Westmead Hospital, Australia; E.F.M. Wouters, Maastricht University Medical Center, Maastricht, the Netherlands; Y. Wu, Department of Occupational Health, School of Public Health, Harbin Medical University, Harbin, China; D. Zagami, Lung Function Laboratory, Gold Coast Hospital, Southport, QLD, Australia; Z. Zhang, Department of Occupational Health, School of Public Health, Harbin Medical University, China; J.P. Zheng, Guangzhou Institute of Respiratory Disease, State Key Laboratory of Respiratory Disease, Guangzhou, China.

1 NUMBERS OF SUBJECTS IN ORIGINAL AND FINAL GROUPS

Of the datasets from 33 countries that were shared with the Global Lung Function Initiative, 26 could be included in the four groups that were eventually formed. The original number of subjects, *i.e.* before exclusion due to missing data or outliers, are presented in Table E1. Of the Mexican data, those in children and adolescents from Mexico City could not be included because the high predicted values did not fit in any of the groups.

Table E1 – Overview of countries sharing data with the Global Lung Function Initiative. The numbers below include data that were subsequently excluded for reasons outlined in the printed text and in Table E2.

Datasets used in final analyses				Datasets not used in final analyses	
Country	N	Country	N	Country	N
Algeria	273	Netherlands	3319	China	3483
Australia	982	Norway	1535	France	63376
Austria	333	Poland	220	India	2548
Brazil	178	Portugal	137	Iran	6137
Canada	329	Sweden	123	Oman	1256
Chile	102	Switzerland	11756	Pakistan	2928
China	5114	Taiwan	2806	Philippines	316
Germany	4708	Thailand	3262	South Africa	146
Iceland	164	Tunisia	870	Total	80190
Israel	124	UK	16888		
Italy	1818	Uruguay	156		
Korea	2252	USA	18212		
Mexico	4236	Venezuela	243		
Total			80140		

After leaving out various datasets from the regression analysis (Table E2) for reasons delineated in the printed manuscript, the breakdown of numbers by age in each of the four groups is as in Table E3.

2 MODELLING SPIROMETRIC INDICES

The LMS method, implemented in the GAMLSS package [1] in the statistical software R [Version 2.14.1; R Foundation, http://www.r-project.org] allows modelling the expected mean (μ or M), the coefficient of variation (σ or S), and skewness (λ or L). A continuous, smooth fit over the entire age range is obtained by the use of splines. Applying the methodology described by Cole *et al.* [2] the best fit was estimated using untransformed

Table E2 – Number of subjects that could not be included in the final analyses for reasons delineated in the printed manuscript.

Data submitted	160,330	
Unknown ethnicity		63,865
Suspected asthma		805
Forced expiratory time < 1 s		123
Cannot be fitted in groups		
Indian and Pakistani		5,476
Omani		1,256
South African		146
Philippino		316
Mexico City		4,009
Iran		6,137
No permission to publish		3,483
Outliers		527
Remaining data		74,187

Table E3 – Breakdown by age of the numbers of subjects in each of four groups. Caution is required when comparing spirometric test results in those over 80 years with the present GLI prediction equations.

	Males				Females				
Age (yr)	Cauc.	Afr.Am.	N E Asia	S E Asia	Cauc.	Afr.Am.	N E Asia	S E Asia	Total
0-5	470			55	446			46	1017
5-10	8452	221		150	8307	227		153	17510
10-15	3924	318		271	3782	340		257	8892
15-20	1677	206	112	227	1557	235	115	292	4421
20-25	1090	137	145	251	1242	206	208	397	3676
25-30	1060	112	124	229	1203	149	217	343	3437
30-35	1106	99	114	248	1467	138	288	426	3886
35-40	1158	72	147	305	1495	132	513	485	4307
40-45	1296	68	121	274	1842	112	506	429	4648
45-50	1203	49	127	245	1782	59	389	404	4258
50-55	1077	58	126	211	1764	73	323	437	4069
55-60	912	40	126	188	1668	70	272	383	3659
60-65	760	44	120	164	1536	89	267	439	3419
65-70	605	34	63	141	1267	67	245	333	2755
70-75	519	34	66	94	1079	65	142	220	2219
75-80	317	18	14	26	727	50	71	91	1314
80-85	148	9	7	15	294	10	18	18	519
85-90	46	1	1	1	93	3	4	6	155
90-95	6		1		16			1	24
95	1				1				2
Total	25827	1520	1414	3095	31568	2025	3578	5160	74187

Cauc. = Caucasian; Afr.Am. = African American; S E Asia = South East Asian; N E Asia = North East Asian.

and log-transformed dependent and explanatory variables, using the Box-Cox-Cole-Green (BCCG) distribution. The optimal degrees of freedom (df) for the spline curve was chosen to minimise the Schwarz Bayesian Criterion (SBC), where adding one df to the model penalises the deviance by ln(N) units, N being the sample size. As N for males and females ~30,000-40,000, the penalty for an extra df is ~10.3-10.6 deviance units. Thus a parsimonious model with an optimal spline curve was obtained.

Over the 2.5 to 95 year age range the age-specific contribution of the spline in age leads to extensive look-up tables. This is exacerbated if separate equations are derived for different ethnic or geographical groups (henceforth called group). Two simplifying approaches were explored. The first one was to include group (a dummy variable) as a function of age. Thus the general form of the equation was:

$$Y = a + b \cdot H + c \cdot A + spline + d_1 \cdot group + d_2 \cdot group \cdot A$$

where Y = dependent variable, H = standing height (cm); A = age (yr); a, b, c, d_1 and d_2 are coefficients which vary for for each index, and d_1 and d_2 vary for each group; *spline* is an age-specific contribution from the spline function. In practice the best fit was obtained after log transformation of the spirometric index (FEV_1, FVC, FEV_1/FVC ratio), height and age. Group-age interaction did not improve the fit, so that the equation simplified to:

$$Y = a + b \cdot H + c \cdot A + spline + d_1 \cdot group$$

The Box-Cox Power Exponential (BCPE) distribution, which also models kurtosis, invariably provided a somewhat better fit than the BCCG distribution. However, the improvement was limited to the tails of the distribution, *i.e.* at Z-scores beyond +3 and -3, thus affecting 0.14% of data at either end; we settled for the BCCG distribution, as slightly greater accuracy is clinically meaningless, and implementation of equations simpler.

3 SIMPLIFYING LOOK-UP TABLES

As delineated above, a term *spline* which varies with age arises from fitting a smoothing spline. Please note that age splines were fitted to L, M and S, hence denoted as Mspline, Sspline and Lspline. Figure E1 depicts how *spline* for predicted FEV_1 (hence Mspline) varies with age in healthy white males and females. Particularly in (pre)school children the table for *spline* needs to be quite detailed. Thus the age-specific table for the 3-95 year age range for L, M and S can each easily have >400 cells for each index (FEV_1 *etc.*). An effort was made to replace each table by an equation, potentially decreasing the complexity of implementing equations in pulmonary function test devices with limited memory dramatically.

It was not possible to satisfactorily fit *spline* over the entire age range. From age 25 years and above a satisfactory polynomial fit could be obtained.

$$spline \approx b_0 + b_1 \cdot (A/100) + b_2 \cdot (A/100)^2 + b_3 \cdot (A/100)^3 + b_4 \cdot (A/100)^4 + b_5 \cdot (A/100)^5$$

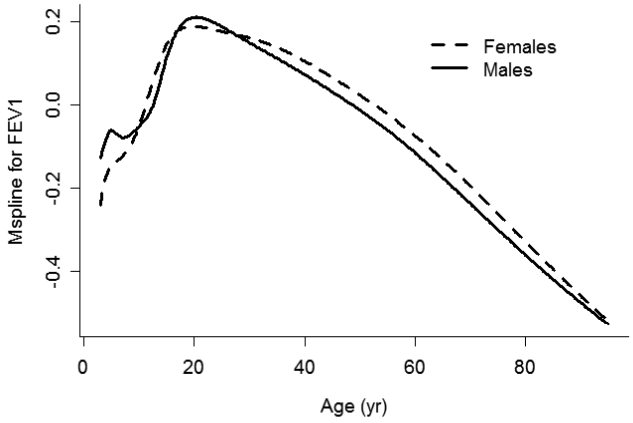

Figure E1 – The contribution of the age-spline to predicted FEV_1 in males and females varies with age.

where A = age in years; seventh degree polynomials were also used to fit Lspline. Coefficients were estimated using linear regression. Thus differences between values predicted using the equation or splines derived by GAMLSS were reduced to a maximum of 0.2%.

Detailed look-up tables are required in children and adolescents, as a few months age difference can affect the predicted values by up to 6%. This is because the predicted values are a power function of age. For example, the linear age coefficient for FEV_1 in males is 0.0574; the contribution from the spline at ages 14.75 and 14.0 years is 0.0958 and 0.0684, respectively. Therefore, if one substitutes 14 years into the equation instead of 14.5, the predicted value will be biased by 100·(1 - exp(0.0574·(log(14.5) - log(14)) + 0.0958 - 0.0684)) = -3%.

The look-up tables for the 3-95 years age range, as well as the equations that replace the tables from 25 years up, can be downloaded from www.lungfunction.org/files/lookuptables.xls.

4 WORKED EXAMPLES OF CALCULATING PREDICTED VALUES

4.1 Introduction

Predicted values depend on three quantities L, M and S, which are functions of sex, age, height and ethnic group. L measures t - dicted value of FEV_1, FVC or FEV_1/FVC. The lower limit of normal (LLN) and the Z-score are calculated from L, M and S as follows:

Predicted value = M
LLN (5^{th} centile) = $\exp(\ln(1 - 1.644 \cdot L \cdot S)/L + \ln(M))$
Z-score = $((\text{measured}/M)^L - 1)/(L \cdot S)$ (for $L \neq 0$)
% predicted = (measured/M)·100

The LMS equations for Caucasians, African Americans, South and North East Asians, valid from 3-95 years, are of this form, with volumes in litres, age in years, and height in cm:

$L = q_0 + q_1 \cdot \ln(\text{Age}) + \text{Lspline}$
$M = \exp(a_0 + a_1 \cdot \ln(\text{Height}) + a_2 \cdot \ln(\text{Age}) + a_3 \cdot \text{black} + a_4 \cdot \text{NEA} + a_5 \cdot \text{SEA} + \text{Mspline})$
$S = \exp(p_0 + p_1 \cdot \ln(\text{Age}) + p_2 \cdot \text{black} + p_3 \cdot \text{NEA} + p_4 \cdot \text{SEA} + \text{Sspline})$

Ta – Regression coefficients for splines for FEV_1 in males that may be used to replace look-up tables for ages 25-95 years in systems with limited memory.

		Coefficients				
		M		S		L
Intercept	a_0	-10.3420	p_0	-2.3268	q_0	0.8866
Height	a_1	2.2196				
Age	a_2	0.0574	p_1	0.0798	q_1	0.0850
Afr.Am.	a_3	-0.1589	p_2	0.1096		
N East Asia	a_4	-0.0351	p_3	-0.3973		
S East Asia	a_5	-0.0881	p_4	0.0327		
	+	Mspline	+	Sspline	+	Lspline

Height (cm), age (yr). Afr.Am. = African American.

where

ln() natural log transformation
white = 1 if a subject is Caucasian, otherwise = 0
black = 1 if a subject is African American, otherwise = 0
NEA = 1 if a subject is North East Asian*, otherwise = 0
SEA = 1 if a subject is South East Asian*, otherwise = 0
coefficients L_0 … q depend on the measurement and sex
Mspline, Sspline, Lspline: age-varying coefficients

* Mongoloid people, does not apply to Indian subcontinent.

For 3-95 years:
Linearly interpolate Lspline, Mspline and Sspline from lookup tables as follows:

$$\text{Xspline}(\text{age}) \approx [(\text{age}_2 - \text{age}) \cdot \text{Xspline}(\text{age}_1) + (\text{age} - \text{age}_1) \cdot \text{Xspline}(\text{age}_2)] / (\text{age}_2 - \text{age}_1)$$

where X represents L, M or S, age = actual age, age_1 and age_2 represent the ages between which interpolation should be performed.

For 25-95 years one might use polynomial equations (table E5):

$\text{Mspline} \approx b_0 + b_1 \cdot (\text{Age}/100) + b_2 \cdot (\text{Age}/100)^2 + b_3 \cdot (\text{Age}/100)^3 + b_4 \cdot (\text{Age}/100)^4 + b_5 \cdot (\text{Age}/100)^5$
$\text{Sspline} \approx c_0 + c_1 \cdot (\text{Age}/100) + c_2 \cdot (\text{Age}/100)^2 + c_3 \cdot (\text{Age}/100)^3 + c_4 \cdot (\text{Age}/100)^4 + c_5 \cdot (\text{Age}/100)^5$
$\text{Lspline} \approx d_0 + d_1 \cdot (\text{Age}/100) + d_2 \cdot (\text{Age}/100)^2 + d_3 \cdot (\text{Age}/100)^3 + d_4 \cdot (\text{Age}/100)^4 + d_5 \cdot (\text{Age}/100)^5 + d_6 \cdot (\text{Age}/100)^6 + d_7 \cdot (\text{Age}/100)^7$

4.2 Example coefficients for FEV_1 in males

For the purpose of demonstrating how to use the equations in calculating predicted and derived values, we reproduce tables from the look-up tables for calculating FEV_1 in males.

In the case of FEV_1 in males, L is a constant, independent of age (Table E4). Table E5 lists the coefficients required to calculate the contribution of splines at specific ages, for subjects between 25-95 years. One should never extrapolate beyond these age ranges!

Ta – Regression coefficients for splines for FEV_1 in males that may be used to replace look-up tables for ages 25-95 years in systems with limited memory.

	Coefficients					
	Mspline		Sspline		Lspline	
Intercept	b_0	0.3901	c_0	-1.6902	d_0	0.8866
Age	b_1	-1.0579	c_1	17.0986	d_1	0.0850
Age_2	b_2	1.4743	c_2	-68.1649	d_2	0
Age_3	b_3	-2.1077	c_3	127.1964	d_3	0
Age_4	b_4	-0.1215	c_4	-109.6777	d_4	0
Age_5	b_5	0.8873	c_5	35.6832	d_5	0
Age_6					d_6	0
Age_7					d_7	0

4.3 Five worked examples

4.3.1 White boy

We wish to calculate the predicted FEV_1, % predicted, 5th centile LLN and Z-score for a white boy age 4.8 yr, height 107 cm, FEV_1 = 0.800 L.
For linear interpolation for age 4.8 yr we consult the look-up table:

Age (yr)	Mspline	Sspline
4.75	-0.0769	0.1592
5	-0.0752	0.1535

Mspline = -0.0769 + ((4.8-4.75)/0.25)·(-0.0752 - (-0.0769)) = -0.0766
Sspline = 0.1592 + ((4.8-4.75)/0.25)·(0.1535 - 0.1592) = 0.1581
M = **FEV_1 predicted** = exp(-10.3420 + 2.2196·ln(107) + 0.0574·ln(4.8) - 0.0766) = 1.0442 L
% predicted = (0.800/1.0442)·100= 76.6%
S = exp(-2.3268 + 0.0798·ln(4.8) + 0.1581) = 0.1296
L = 0.8866 + 0.0850·ln(4.8) = 1.0199
LLN = exp(ln(1 - 1.644·L·S)/ L + ln(M))
= exp(ln(1 - 1.644·1.0199·0.1296)/1.0199 + ln(1.0442))
= 0.8212 L
Z-score = ((measured/predicted)^L - 1)/(L·S)
= ((0.800/1.0442)$^{1.0199}$ - 1)/(1.0199·0.1296) = -1.80

4.3.2 African American boy

We wish to calculate the predicted FEV_1, % predicted, 5th centile LLN and Z-score for an African American boy age 12.2 yr, height 152 cm, FEV_1 = 2.405 L.

For linear interpolation to age 12.2 yr we consult the look-up table:

Age (yr)	Mspline	Sspline
12	-0.0176	-0.0387
12.25	-0.0101	-0.0339

Mspline = -0.0176 + ((12.2-12)/0.25)·(-0.0101- (-0.0176)) = -0.0116
Sspline = -0.0387 + ((12.2-12)/0.25)·(-0.0339 - (-0.0387)) = -0.0349
M = **FEV_1 predicted** = exp(-10.3420 + 2.2196·ln(152) + 0.0574·ln(12.2) - 0.1589 - 0.0116) = 2.1860 L
% predicted = (2.405/2.1860)·100 = 110.0%
S = exp(-2.3268 + 0.0798·ln(12.2) + 0.1096 - 0.0349) = 0.1284
L = 0.8866 + 0.0850·ln(12.2) = 1.0992
LLN = exp(ln(1 - 1.644·L·S)/L + ln(M))
= exp(ln(1 - 1.644·1.0992·0.1284)/1.0992 + ln(2.1860))
= 1.7193 L
Z-score = ((measured/predicted)^L - 1)/(L·S)
= ((2.405/2.1860)$^{1.0992}$ - 1)/(1.0992·0.1284) = 0.78

4.3.3 South East Asian adult male

We wish to calculate the predicted FEV_1, % predicted, 5th centile LLN and Z-score for a South East Asian male age 53 yr, height 175 cm, FEV_1 = 2.410 L.

Calculations can be performed using the look-up table for L, M and S, or by replacing the latter with the equations in Table E5. First using the look-up tables:

Age (yr)	Mspline	Sspline
53	-0.0404	0.0003

Mspline = -0.0404
Sspline = 0.0003
M = **FEV_1 predicted** = exp(-10.3420 + 2.2196·ln(175) + 0.0574·ln(53) - 0.0881 - 0.0404) = 3.3911 L
% predicted = (2.410/3.3911)·100 = 71.1%
S = exp(-2.3268 + 0.0798·ln(53) + 0.0327 + 0.0003) = 0.1395
L = 0.8866 + 0.0850·ln(53) = 1.2241
LLN = exp(ln(1 - 1.644·L·S)/ L + ln(M))
= exp(ln(1 - 1.644·1.2241·0.1395)/1.2241 + ln(3.3911))
= 2.5908 L
Z-score = ((measured/predicted)^L - 1)/(L·S)
= ((2.410/3.3911)$^{1.2241}$ - 1)/(1.2241·0.1395) = -2.00

We can also replace the look-up tables by the equations in Table E5:
Mspline = 0.3901 - 1.0579·(53/100) + 1.4743·(53/100)^2 - 2.1077·(53/100)^3 - 0.1215·(53/100)^4 + 0.8873· (53/100)^5 = -0.0427
Sspline = -1.6902 + 17.0986·(53/100) - 68.1649·(53/100)^2 + 127.1964·(53/100)^3 - 109.6777·(53/100)^4 + 35.6832·(53/100)^5 = -0.0007
M = **FEV_1 predicted** = exp(-10.3420 + 2.2196·ln(175) + 0.0574·ln(53) - 0.0881 - 0.0427) = 3.3833 L
% predicted = (2.410/3.3833)·100 = 71.2%
S = exp(-2.3268 + 0.0798·ln(53) + 0.0327 - 0.0007) = 0.1383
L = 0.8866 + 0.0850·ln(53) = 1.2241
LLN = exp(ln(1 - 1.644·L·S)/ L + ln(M))
= exp(ln(1 - 1.644·1.2241·0.1383)/1.2241 + ln(3.3833))
= 2.5919 L
Z-score = ((measured/predicted)^L - 1)/(L·S)
= ((2.410/3.3833)$^{1.2241}$ - 1)/(1.2241·0.1383) = -2.01

4.3.4 South East Asian adult female

We wish to calculate the predicted FEV_1, % predicted, 5th centile LLN and Z-score for a South East Asian female age 39.1 yr, height 165 cm, FEV_1 = 2.210 L.

Calculations can be performed using the look-up table for L, M and S, or by replacing the latter with the equations in table E6. First using the look-up tables:

Ta – Regression coefficients for FEV_1 for calculating M (predicted value), S (coefficient of variation) a). The contributions of splines must be added to the calculated values; they are available in look-up tables.

	Coefficients					
	Mspline		Sspline		Lspline	
Intercept	a_0	-9.6987	p_0	-2.3765	q_0	1.1540
Height	a_1	2.1211	p_1			
Age	a_2	-0.0270	p_2	0.0972	q_1	0
Afr.Am.	a_3	-0.1484	p_3	0.1016		
N East Asia	a_4	-0.0149	p_4	-0.0109		
S East Asia	a_5	-0.1206	p_5	0.0733		
	+	Mspline	+	Sspline	+	Lspline

Table E7 – Regression coefficients for splines for FEV_1 in females that may be used to replace look-up tables for ages 25-95 years in systems with limited memory.

	Coefficients					
	Mspline		Sspline		Lspline	
Intercept	b_0	0.0552	c_0	-0.0825	d_0	1.1540
Age	b_1	1.6032	c_1	1.4104	d_1	0
Age_2	b_2	-6.4855	c_2	-11.2699	d_2	0
Age_3	b_3	10.2741	c_3	29.4400	d_3	0
Age_4	b_4	-9.8646	c_4	-29.5505	d_4	0
Age_5	b_5	3.8808	c_5	10.4405	d_5	0
Age_6					d_6	0
Age_7					d_7	0

Age(yr)	Mspline	Sspline
39	0.1112	-0.0919
39.25	0.1094	-0.0913

Mspline = 0.1112 + (0.1/0.25)·(0.1094 - 0.1112) = 0.1105
Sspline = -0.0919 + (0.1/0.25)·(-0.0913 - (-0.0919)) = -0.0917
Lspline = 0

M = **FEV_1 predicted** = exp(-9.6987 + 2.1211·ln(165) - 0.0270·ln(39.1) - 0.1206 + 0.1105) = 2.7800 L
% predicted = (2.210/2.7800)·100 = 79.5 %
S = exp(-2.3765 + 0.0972·ln(39.1) + 0.0733 - 0.0917) = 0.1302
L = 1.1540
LLN = exp(ln(1 - 1.644·L·S)/ L + ln(M))
= exp(ln(1 - 1.644·1.1540·0.1302)/1.1540 + ln(2.7800))
= 2.1741 L
Z-score = ((measured/predicted)^L - 1)/(L·S)
= ((2.210/2.7800)$^{1.1540}$ - 1)/(1.1540·0.1302) = -1.55

We can also replace the look-up tables by equations (Table E7):

Mspline = 0.0552 + 1.6032·(39.1/100) - 6.4855·(39.1/100)^2 + 10.2741·(39.1/100)^3 - 9.8646·(39.1/100)^4 + 3.8808·(39.1/100)^5 = 0.1096
Sspline = -0.0825+ 1.4104·(39.1/100) - 11.2699·(39.1/100)^2 + 29.4400·(39.1/100)^3 - 29.5505·(39.1/100)^4 + 10.4405·(39.1/100)^5 = - 0.0894
M = **FEV_1 predicte**d = exp(-9.6987 + 2.1211·ln(165) - 0.0270·ln(39.1) - 0.1206 + 0.1096) = 2.7775 L
% predicted = (2.210/2.7775)·100 = 79.6%
S = exp(-2.3765 + 0.0972·ln(39.1) + 0.0733 - 0.0894) = 0.1305
Lspline = 0
L = 1.1540
LLN = exp(ln(1 - 1.644·L·S)/ L + ln(M))
= exp(ln(1 - 1.644·1.1540·0.1305)/1.1540 + ln(2.7775))
= 2.1707 L
Z-score = ((measured/predicted)^L - 1)/(L·S)
= ((2.210/2.7775)$^{1.1540}$ - 1)/(1.1540·0.1305) = -1.54

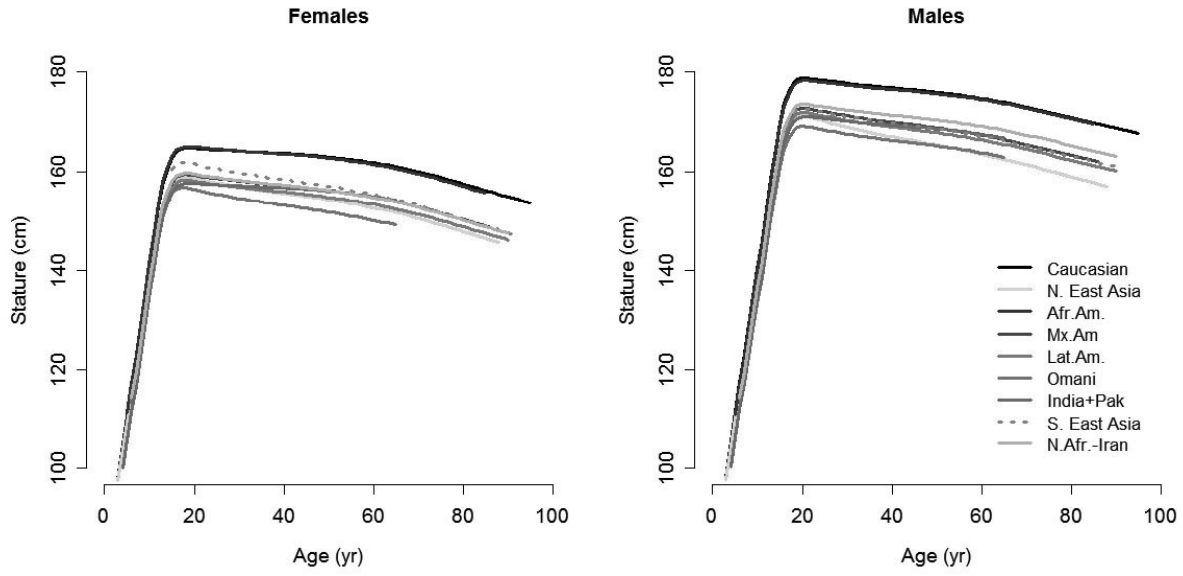

Figure E2 – Standing height as a function of age in groups from various parts of the world. Caucasians and African Americans are up to 10 cm taller than other ethnic groups.
N East Asia = North East Asia; Afr.Am. = African American; Mx.Am. = Mexican American; Lat.Am. = Latin American; India+Pak = India and Pakistan; S East Asia = South East Asia; N.Afr.+Iran = Algeria, Tunisia and Iran.

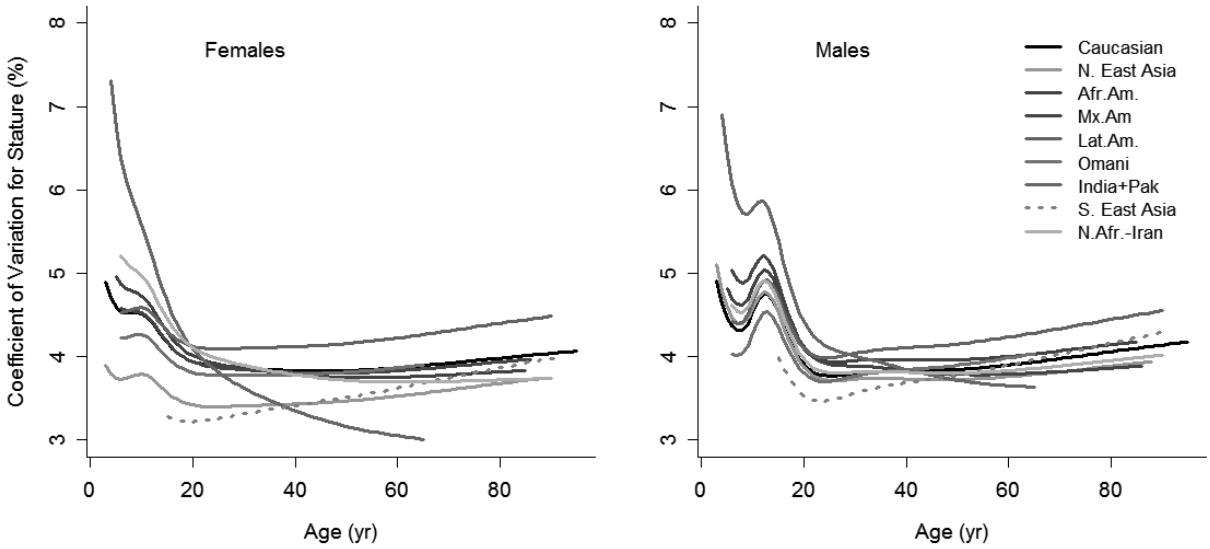

Figure E3 – Between-subject coefficient of variation of standing height as a function of age in groups from various parts of the world. The between-person variability is particularly large in children from India and Pakistan; we could not elucidate the cause of these results.
N East Asia = North East Asia; Afr.Am. = African American; Mx.Am. = Mexican American; Lat.Am. = Latin American; India+Pak = India and Pakistan; S East Asia = South East Asia; N.Afr.+Iran = Algeria, Tunisia and Iran.

As demonstrated above, there are slight differences in the results calculating values using the look-up table or the equations replacing them. The discrepancies are not clinically important; nevertheless use of look-up tables and linear interpolation is the recommended procedure.

5 DIFFERENCES IN STANDING HEIGHT BETWEEN GROUPS

Stature, the main determinant of pulmonary function, differed significantly between populations (Figure E2). The coefficient of variation (CoV) for height is largest in preschool children (Figure E3), declining rapidly towards adolescence followed by an increase, reflecting differences in the age of onset and end of the pubertal growth spurt. There is then a drop until about age 30 years, followed by a small but steady increase towards old age. With one exception, the maximum difference between groups is about 0.01 (1%). The CoV is considerably larger in Indian and Pakistani schoolchildren and adolescents than in others (Figure E3), probably due to differences in socio-economic conditions; differences in stature and CoV between Pakistani and Indian children were small.

6 DATA

6.1 Latin-American data

Five datasets (Mexico City, Sao Paulo, Caracas, Montevideo, Santiago) [3] related to adults, one to Mexican children and young adults [4]. There were pronounced differences in height for age between centres, in males ranging from 12 cm at age 45 to 7 cm at age 80 years (Figure E4), people in Mexico City being smallest. Predicted values for spirometry derived from each individual dataset also produced appreciable differences (Figure E5).

6.2 East Asian data

Nine datasets were available from Hong Kong [5,6], Taiwan [7], Thailand [8], the USA [9], Korea [10], China [11] (data

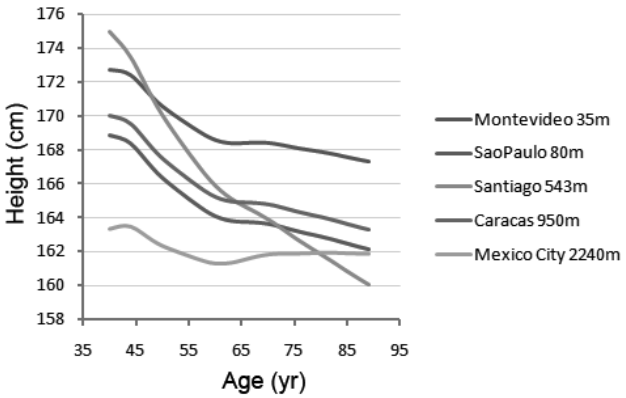

Figure E4 – Differences in standing height in males in 5 centres participating in the PLATINO study [3].

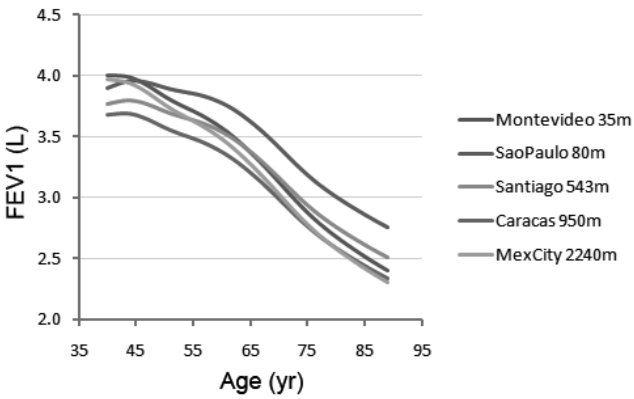

Figure E5 – Predicted FEV_1 in males in 5 centres participating in the PLATINO study [3].

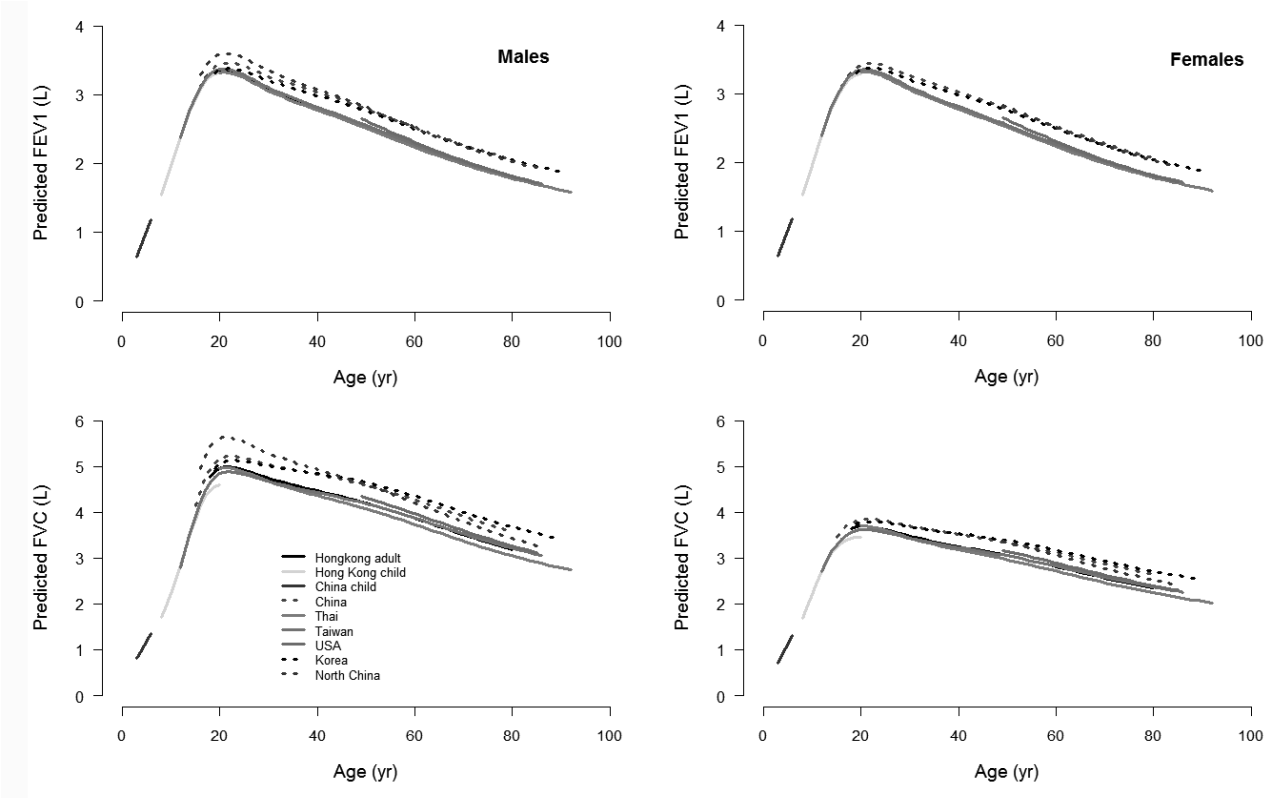

Figure E6 – Predicted FEV_1 and FVC in 9 different datasets from East Asian subjects. Data from South East Asia are systematically shorter than those from North East Asia.

not used for present analyses, see printed text), North China [12], and a dataset on children aged 3-6 years [13]. Regression analysis revealed significant differences for spirometric indices between centres. Using GAMLSS, a best fit of height for age was derived from the collated data of East Asians; calculated height for age was then used to derive predicted values for spirometric indices for each centre, allowing comparison between centres that was unaffected by differences in height. Predicted values for FEV_1 and FVC from mainland China (unpublished data from North and South China collected in 2008), North China [12] and Korea [10] came out systematically higher than in the remaining 6 datasets (data collected between 1996-2002 in Hong Kong, Taiwan, Thailand, USA and China), which agreed remarkably well (Figure E6). No evidence was found that this related to methodological differences, or to unrepresentative samples arising from small sample size. There were significant differences in standing height between centres, Chinese in the USA and the North of China (*i.e.* north of the Huai River and Qinling Mountains) being tallest, and people in Thailand shortest: at age 50 yr maximum differences were 5.4 cm in females, and 5.7 cm in males. Data collection spanned 18 years (1990-2008), during which time changes occurred in Oriental societies which significantly affected the availability of food and education. Better health conditions are associated with an increase in standing height, mainly due to an increase in leg length [14], although the latter finding could not be reproduced recently [15]. Average height for someone aged for example 40 yr in 1990 may differ from that of a 40 yr old person in 2008. Displaying standing height as a function of birth date, rather than calendar age, removes temporal differences, allowing to better depict differences between populations (Figure E7). The slopes are steepest for males and females in the USA and Korea, and flattest for Taiwan, leading to curves crossing over at about the year 1960 (Figure E7). These slopes reflect a combination of diminishing height with advancing age and secular trends in height. For example, for mainland China the estimated rise in standing height between 1979-1995 for a 17 years old

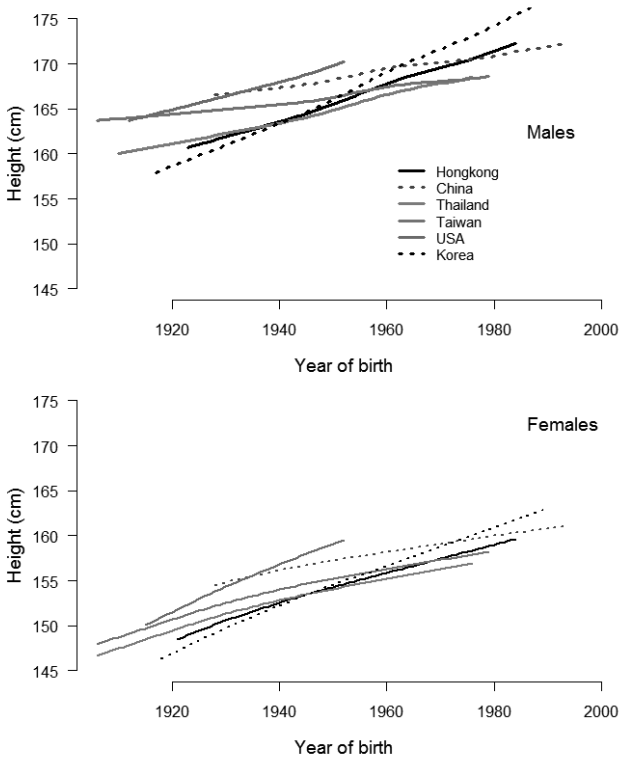

Figure E7 – Relationship between year of birth and standing height in adult East Asian subjects born in Hong Kong, China, Thailand, Taiwan,the USA and Korea. The slope of the line reflects both a secular trend and a change in height with ageing.

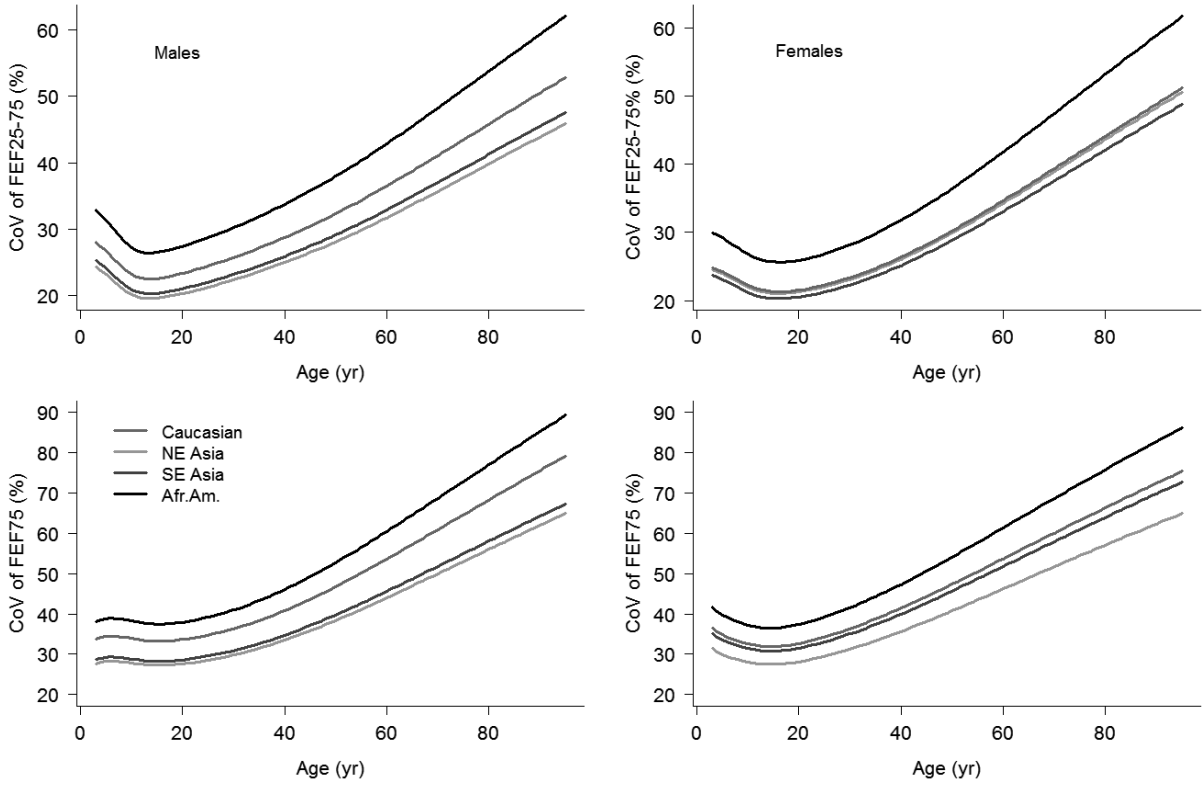

Figure E8 – Coefficient of variation (CoV) of $FEF_{25-75\%}$ and FEF_{75} in males and females. Note the very large variability, particularly in adults, which severely limits the use of these indices for diagnostic purposes. However, this does not preclude a more favourable coefficient of variation for within-subject comparisons, nor the use of these indices in etiological studies where differences between groups may provide valuable clues.
NE Asia = North East Asia; SE Asia = South East Asia, Afr.Am. = African American.

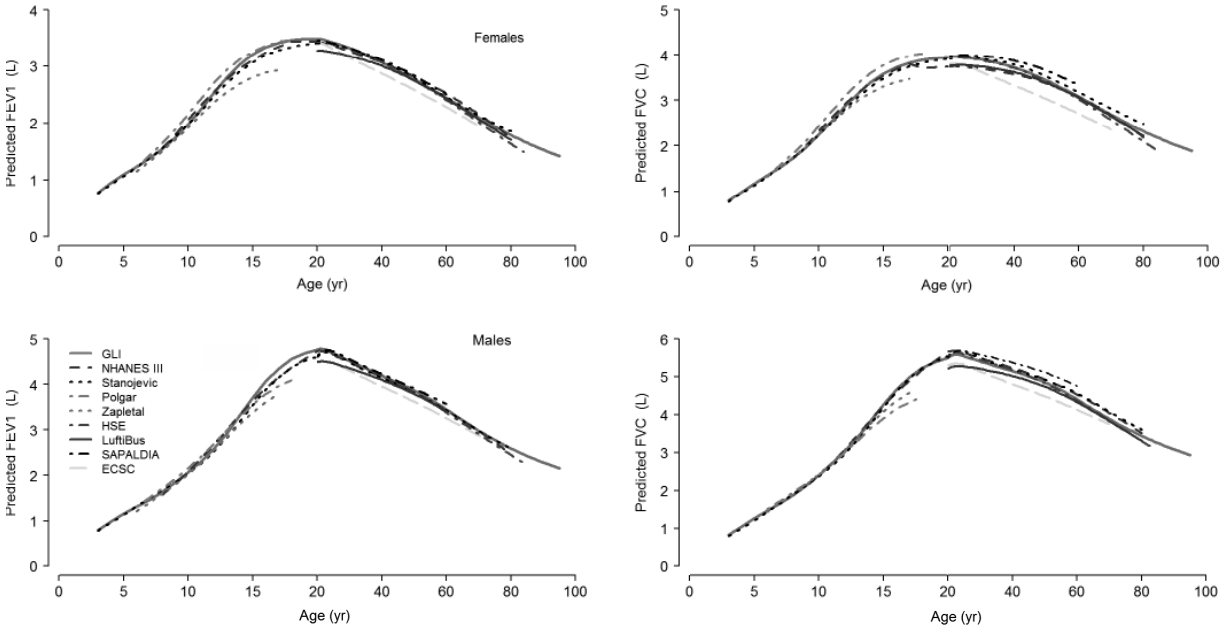

Figure E9 – Comparison of predicted values for FVC (right panels) and FEV_1 (left panels) in males and females. GLI = Global Lungs Initiative (present study), Stanojevic [27], NHANES III [22], ECSC = European Community for Steel and Coal [26], HSE = Health Survey for England 1995-1996 [23], LuftiBus [24], SAPALDIA [25], Polgar [28], Zapletal [29]. Predicted values from the ECSC in adults, from the LuftiBus study in young adults, and from Polgar and Zapletal in children and adolescents, do not agree well with the others. Data from Stanojevic and the GLI practically overlap. Note the different scaling from 0-20 years and 20-100 years, so as to show greater detail in the paediatric age range. Graphs were generated using mean height for age in Caucasians, hence the above differences in predicted values cannot be accounted for by differences in height.

person was 1.0 and 2.0 cm per decade for urban and rural males, respectively, and 0.5 and 0.7 cm per decade for urban and rural females, respectively [16]. At the same time the coefficient of variation for stature widened, indicating greater regional variation in height for age [16,17]. However, the well-known average height difference between southern and northern Chinese for urban populations has narrowed from 6-7 cm in the 1920s to 2-3 cm in the 1990s [17]. Figure E7 shows that the subjects from mainland China and Korea were on average taller, yet shorter than Orientals living in the USA. Differences in pulmonary function remaining after standardising for the same height for age (Figure E6) might therefore reflect differences in body build. This is in keeping with various reports [15-21]. The scale of differences in FEV_1 and FVC, and the limited time span between data collections are not compatible with a secular trend.

6.3 Handling of longitudinal datasets

Two longitudinal studies of schoolchildren [18] and adolescents [19] were transformed into a cross-section by selecting at random one record from a person's available measurements so that the new cross-sectional data set had a representative age distribution. In another follow-up study of school children [20] only the data collected at the first occasion were selected. In a longitudinal study of adults [21] the third record was selected.

7 $FEF_{25-75\%}$ AND FEF_{75}

The coefficients of variation of $FEF_{25-75\%}$ and FEF_{75} rise to very high values in adults (Figure E8).

8 COMPARISON OF PREDICTED VALUES

The present set of prediction equations, using collated data, was compared with those from four recent large datasets [22-25]. Regression equations were derived using GAMLSS [1] in R [Version 2.14.1; R Foundation, http://www.r-project.org]. Predicted values and their lower limits of normal were calculated as a function of age, using the mean height for age in Caucasians; the latter was obtained by regressing height on age, using GAMLSS. Predicted values according to the ECSC/ERS [26], widely used in Europe, and predicted values from Stanojevic [27] which are being used increasingly, were added to the comparison (see also Table 5 in printed paper), as were those from Polgar and Zapletal [28,29] (Figure E9). Except for the ECSC/ERS predicted values and those from Polgar and Zapletal, which differ systematically from the other ones, differences are marginal. Hence, the use of collated data has not led to loss of accuracy and precision.

Acknowledgement

The authors are extremely grateful to all individuals and organisations who contributed data and information to the Global Lungs Initiative. Without their help, contributions and mutual trust this project would have been impossible. The extensive statistical review by C. Schindler is also gratefully acknowledged. This study is based on data from 70 centres, including the MESA study. The MESA and MESA Lung Studies are conducted and supported by the National Heart, Lung and Blood Institute (NHLBI) in collaboration with the MESA and MESA Lung Investigators. In addition to review by representatives of all bodies contributing data to the GLI, this manuscript has been reviewed by the MESA investigators for scientific content and consistency of data interpretation with previous MESA publications and significant comments incorporated prior to submission for publication. A full list of participating MESA Investigators and institutions can be found at http://www.mesa-nhlbi.org/.

References

1. Rigby R A, Stasinopoulos DM. Generalized additive models for location, scale and shape (with discussion). *Appl Statist* 2005; 54: 507-554.
2. 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. *Statist Med* 2009; 28: 880-898.
3. Pérez-Padilla R, Valdivia G, Muiño A, *et al*. Spirometric reference values in 5 large Latin American cities for subjects aged 40 years or over. *Arch Bronconeumol* 2006; 42: 317-325.
4. Pérez-Padilla R, Regalado-Pineda J, Rojas M, *et al*. Spirometric function in children of Mexico City compared to Mexican-American children. *Pediatr Pulmonol* 2003; 35: 177–183.
5. Ip MSM, Karlberg EM, Karlberg JPE, Luk KDK, Leong JCY. Lung function reference values in Chinese children and adolescents in Hong Kong. I. Spirometric values and comparison with other populations. *Am J Respir Crit Care Med* 2000; 162: 424–429.
6. Ip MSM, Ko FWS, ACW Lau, Yu WC, Tang KS, Choo K, Chan-Yeung MMW. Updated spirometric reference values for adult Chinese in Hong Kong and implications on clinical utilization. *Chest* 2006; 129: 384–392.
7. Pan WH, Chen JY, Haung SL, *et al*. Reference spirometric values in healthy Chinese neversmokers in two townships of Taiwan. *Chin J Physiol* 1997; 40: 165-174.
8. Dejsomritrutai W, Nana A, Maranetra KN, *et al*. Reference spirometric values for healthy lifetime nonsmokers in Thailand. *J Med Assoc Thai* 2000; 83: 457-466.
9. MESA Lung Study, courtesy dr G. Barr.
10. Jung Keun Choi, Domyung Paek, Jeoung Oh Lee. Normal predictive values of spirometry in Korean population. *Tuberc Respir Dis* 2005; 58: 230-242.
11. Chinese adults, courtesy Prof. J.P. Zheng.
12. Wu Y, Zhang Z, Gang B, Love EJ. Predictive equations for lung function based on a large occupational population in North China. *J Occup Health* 2009; 51: 417-477.
13. Zhang QL, Zheng JP, Yuan BT, He H, Wang J, An JY, Zhang M, Luo DF, Chen GL. Feasibility and predicted equations of spirometry in Shenzhen preschool children. Zhonghua Er Ke Za Zhi (*Chinese J Pediatr*) 2005; 43: 843-848.
14. Tanner JM, Hayashi T, Preece MA, *et al*. Increase in length of leg relative to trunk in Japanese children and adults from 1957 to 1977: comparison with British and with Japanese Americans. *Ann Hum Biol* 1982; 9: 411–423.
15. Ashizawa K, Tanamachi N, Kato S, Kumakura C, Zhcou X, Jin F, Li Y, Lu S. Growth of height and leg length of children of Beijing and Xilinhot, China. *Anthropol Sci* 2008; 116: 67-76.
16. Stature and Famine in China: The Welfare of the Survivors of the Great Leap Forward Famine, 1959-61, http://ehsanz.econ.usyd.edu.au/papers/Morgan.pdf. Last accessed March 30, 2012.
17. Morgan SL. Richer and taller: stature and living standards in China, 1979-1985. *The China Journal* 2000; 44: 1-39.
18. Etler DA.Recent developments in the study of human biology in China: a review. *Hum Biol* 1992; 64: 567-585.
19. Zhang HG, Chen YF, Ding M, *et al*. Dermatoglyphics from all Chinese ethnic groups reveal geographic patterning. *PLoS One* 2010; 20; 5:e8783.
20. Lin W-S, Zhu F-C, Chen ACN, Xin W-H, Su Z, Li J-Y, Ye G-S. Physical growth of Chinese school children 7-18 years, in 1985. *Ann Hum Biol* 1992; 19: 41-55.
21. Weitz CA, Garruto RM, Chin C, Liu J. Morphological growth and thorax dimensions among Tibetan compared to Han children,

adolescents and young adults born and raised at high altitude. *Ann Hum Biol* 2004; 31: 292–310.
22. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general US population. *Am J Respir Crit Care Med* 1999; 159: 179–187.
23. Falaschetti E, Laiho J, Primatesta P, Purdon S: Prediction equations for normal and low lung function from the Health Survey for England. *Eur Respir J* 2004; 23: 456-463.
24. Kuster SP, Kuster D, Schindler C, Rochat MK, Braun J, Held L, Brändli O. Reference equations for lung function screening of healthy never smoking adults aged 18-80 years. *Eur Respir J* 2008; 31: 860-868.
25. Brändli O, Schindler Ch, Künzli N, Keller R, Perruchoud AP, and SAPALDIA team. Lung function in healthy never smoking adults: reference values and lower limits of normal of a Swiss population. *Thorax* 1996; 51: 277-283.
26. Quanjer PH, Tammeling GJ, Cotes JE, *et al*. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J* 1993; 6 (Suppl. 16): 5-40.
27. Stanojevic S, Wade A, Cole TJ, *et al*. Spirometry centile charts for young Caucasian children: the Asthma UK Collaborative Initiative. *Am J Respir Crit Care Med* 2009; 180: 547–552.
28. Polgar G, Promadhat V. Pulmonary function testing in children: Techniques and standards. Philadelphia, Saunders, 1971.
29. Zapletal A, Paul T, Samanek N. Die Bedeutung heutiger Methoden der Lungenfunktionsdiagnostik zur Feststellung einer Obstruktion der Atemwege bei Kindern und Jugendlichen. *Z Erkrank Atm-Org* 1977; 149: 343-371.