Online supplement

Adopting Reference Equations for Pulmonary Function Tests in a Clinical Laboratory

Philip H. Quanjer, Departments of Pulmonary Diseases and Pediatrics, Erasmus Medical Centre – Sophia Children's Hospital, Erasmus University, Rotterdam, the Netherlands
Janet Stocks, Portex Respiratory Unit, UCL Institute of Child Health, London, UK
Tim J. Cole, MRC Centre of Epidemiology for Child Health, UCL Institute of Child Health, London, UK
Graham L Hall, 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
Sanja Stanojevic, Portex Respiratory Unit, UCL Institute of Child Health, London, UK

on behalf of the ERS Global Lungs Task Force.

Methods

Four of the datasets contained longitudinal data. In two cases these were transformed into cross-sectional sets by selecting one cross-sectional subset, in two other cases by selecting a single record from a person's available measurements so that the new cross-sectional dataset had an age distribution that was similar to the original dataset.

Statistical analysis

Prior to analysis, data were checked for transcription errors, improbable values and obvious outliers. As the majority of studies had previously been published there were very few errors (<1%). Lung function indices were modelled separately in males and females as a function of age and height, using generalised additive modelling of location, scale and shape (GAMLSS) [1]. This technique offers a choice of distributions, and allows modelling the median, the coefficient of variation and skewness using cubic smoothing splines. In addition it allows modelling additive and multiplicative relationships. All models were fitted using the package GAMLSS [2] in R (version 2.11.1) [3] as described by Cole [4] and recently applied to spirometry [5-6]. We used the Box-Cox-Cole-Green (BCCG) distribution. A stepwise approach [4] was used to determine the best curves for the median, the coefficient of variation, and skewness. The model with the smallest Schwarz Bayesian Criterion (SBC) within a family of models was selected. The final choice of the most appropriate and parsimonious model was also based on inspection of worm plots [7], normal probability plots, the distribution of residuals, and the fit of the data to centiles. The analysis is sensitive to extreme outliers. Therefore, after derivation of prediction equations, the z-scores were inspected for values < -5 and > +5; if present these were removed, and the curve fitting procedure restarted. In total there were 7 such outliers.

Modelling pulmonary function

The best fitting model (BCCG) describes a multiplicative and allometric height relationship of the form ln(index) = a + b*log(height) +c*log(age) + d(log(age)), where d(log(age)) represents the age-specific contribution of the spline function. The volumes are proportional to height raised to the power ~2.2. As reported earlier [5] the coefficient of variation (CoV) for each of the indices varies with age, with a minimum at adolescence; the CoV could be properly modelled as a spline function of age. The Box-Cox power exponential (BCPE)

distribution produced a significant reduction in the Schwarz Bayesian Criterion (SBC), indicating that ideally kurtosis needs to be modelled. As there were no improvements in the 2.5 and 97.5 centiles, we settled for the Box-Cox Cole and-Green (BCCG) distribution.

Relationship between z-scores and centiles

As delineated in the printed text the scatter around predicted values differed between centres. By definition, the SD derived by GAMLSS from collated data will be one. Since this includes differences between centres, the average of the SD of the z-scores from all the individual centres will be slightly lower (0.97 in females, 0.95 in males for FEV_1; 0.96 in females, 0.94 in males for FVC). Hence the lower limit of normal for FEV_1 and FVC from the entire dataset is slightly lower than in the small datasets. This effect was more pronounced for the FEV_1/FVC ratio (0.90 in both females and males). The relationship between z-scores and centiles, obtained by multiplying the SD by 1.6448 is as shown below.

Index	SD z-score		Centile	
	Males	Females	Males	Females
FEV_1	0.95	0.97	6	5.5
FVC	0.94	0.96	6.1	5.7
FEV_1/FVC	0.90	0.90	7.0	7.0

Thus, in smaller populations, such as local controls, the LLN for FEV_1 or FEV_1/FVC in females will, on average, delineate 5.5% or 7% of observations respectively as being below the normal range, rather than the 5% identified if based on the LLN from the entire collated dataset,

The LLN from collated data is mainly determined by the largest datasets, which represent a representative sample from the population. Adopting the LLN from collated data should therefore not lead to any bias.

Table E1 – Summary of numeric data from the 30 centres included in the analysis

	Females			Males				Numbers		Age ranges	
Study	FEV_1	FVC	FEV_1/FVC	FEV_1	FVC	FEV_1/FVC	Study date	Females	Males	Females	Males
1	0.02	-0.01	0.09	0.03	-0.02	0.12	1995-1996	4759	3683	7-94	7-89
2	-0.12	-0.17	0.17	-0.13	-0.18	0.09	2001-2005	5129	3459	18-95	18-95
3	-0.09	-0.10	0.01	-0.11	-0.15	0.05	1999-2000	3345	3383	8.1-10.6	7.4-10.4
4	0.18	0.25	-0.15	0.19	0.25	-0.10	1994-1999	1408	1416	6.1-13.9	6.3-13.6
5	0.19	0.31	-0.25	0.15	0.25	-0.22	1991	1888	1267	18.2-61.8	18.2-61.8
6	-0.15	-0.05	-0.28	-0.25	-0.14	-0.22	2006-2007	1003	881	4.1-18.9	4.0-18.9
7	0.02	0.12	-0.12	-0.05	0.05	-0.18	1988-1994	1364	869	8-80	8-80
8	0.16	0.09	0.03	0.28	0.16	0.07	1984	860	806	5.9-12.7	6.1-13.4
9	-0.30	-0.19	-0.05	-0.35	-0.15	-0.37	1985-1987	314	447	4.6-18.8	4.4-18.7
10	-0.06	0.18	-0.43	-0.08	0.16	-0.29	1980-1982	670	369	8-70	8-65
11	0.40	0.07	0.21	0.35	0.28	0.04	1978-1985	123	365	11.4-19.6	11.5-19.4
12	0.27	0.20	0.17	0.24	0.20	0.10	1995-1997	545	361	19-80	19-80
13	-0.24	0.07	-0.55	-0.30	0.00	-0.39	1991-1995	612	341	31-78	31-78
14	0.07	-0.11	0.84	0.11	-0.12	0.78	1999	232	249	2.5-6.0	2.5-7.0
15	0.24	0.26	0.04	0.37	0.20	0.29	1991-1993	275	181	8-74	8-70
16	0.11	0.11	-0.11	0.10	0.17	-0.18	2003-2004	201	169	4.7-5.6	4.6-6.1
17	-0.07	-0.15	0.00	0.08	-0.08	0.05	1990	249	165	18.1-78.5	19.0-78.4
18	-0.50	-0.60	0.63	-0.18	-0.38	0.69	2002	133	149	3.1-6.3	3.1-6.3
19	0.08	0.24	-0.49	0.00	0.20	-0.53	1998	72	117	3.1-6.9	3.0-7.0
20	-0.28	-0.35	-0.02	-0.54	-0.25	-0.40	1993	61	100	4.6-8.0	4.4-7.9
21	0.49	0.24	0.66	0.33	0.06	0.81	2006-2007	59	81	4.8-7.9	4.2-7.9
22	-0.17	0.10	-0.62	0.01	0.36	-0.48	2003-2006	49	68	2.5-6.5	2.7-6.4
23	0.49	0.67	-0.32	0.32	0.43	-0.12	2002-2008	64	68	20-66	22-67
24	0.72	0.63	0.65	0.72	0.40	0.24	2002	58	53	3.3-5.7	3.3-5.7
25	-0.22	0.17	-0.03	-0.11	-0.09	-0.18	2000-2007	46	52	2.6-6.7	3.4-7.0
26	0.11	0.45	-0.47	0.09	0.54	-0.73	2008	72	51	8.9-88.7	7.8-95.9
27	0.04	0.61	-1.03	0.29	0.47	-0.34	2005-2008	42	75	5.9-7.6	5.9-7.7
28	-0.41	-0.42	-0.25	-0.42	-0.34	-0.20	2009	55	31	9-58	6.0-71
29	0.69	0.82	-0.67	0.11	0.23	-0.32	2007-2008	20	18	3.2-5.1	3.5-5.0
30	0.08	-0.34	0.52	-0.58	-0.82	0.71	2001-2002	33	17	6.5-8.0	6.6-8.0
							Total	23741	19291		

Legend: The mean standardised residuals (z scores) from the predicted values (derived from all data) are depicted for each centre, together with the number of subjects, the age ranges and the years during which data were collected.

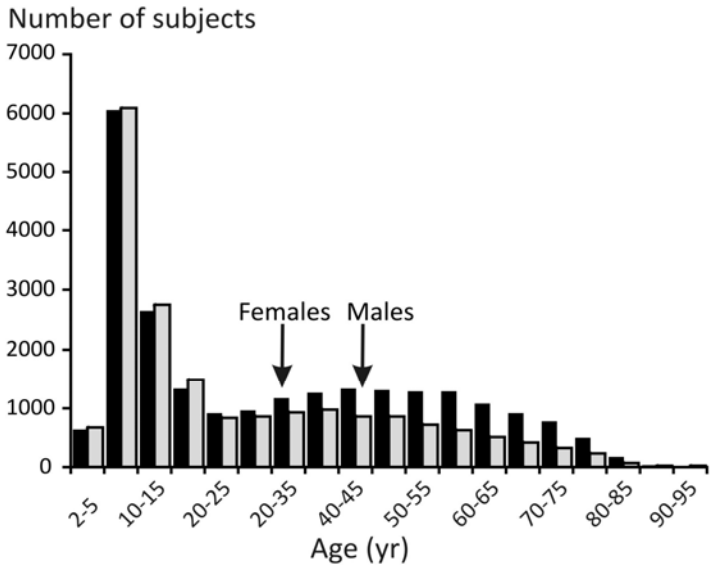

Figure E1 – Distribution of healthy white females (N=23,741, black columns) and males (N=19,291, gray columns) by age.

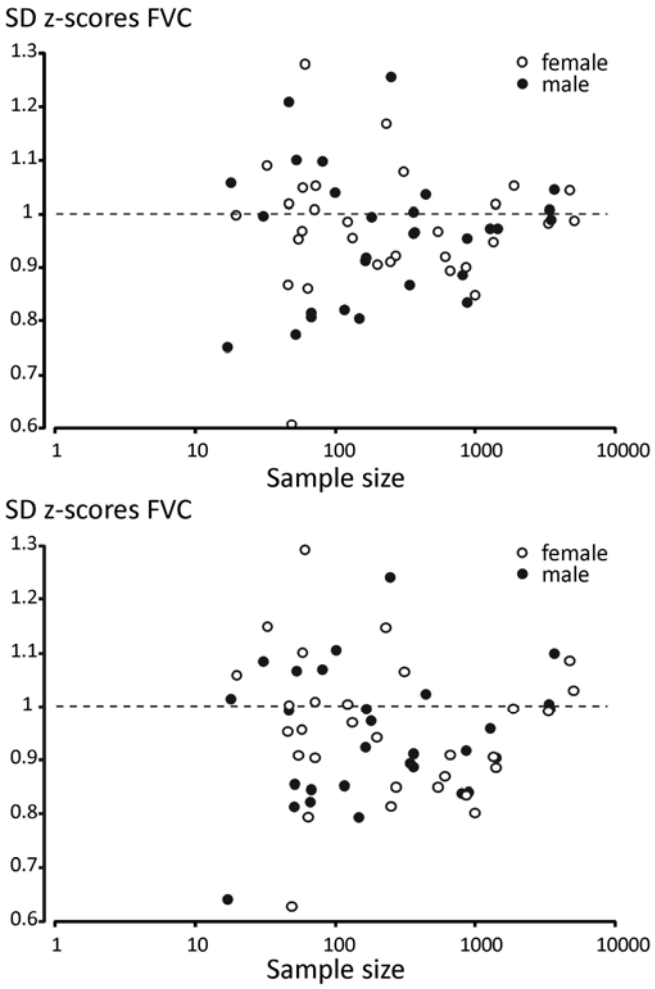

Figure E2 - Standard deviations of z-scores as a function of sample size in white males and females in 30 centres.

References

1. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape (with discussion). *Applied Statistics* 2005; 54: 507–544.
2. Mikis Stasinopoulos and Bob Rigby with contributions from Calliope Akantziliotou. (2008). GAMLSS: Generalized Additive Models for Location Scale and Shape. R package version 1.9-4. http://www.gamlss.com/.
3. R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.
4. 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.
5. Stanojevic S, Wade A, Cole TJ, *et al*. on behalf of the Asthma UK collaborative group. Spirometry centile charts for young Caucasian children: The Asthma UK Collaborative Initiative. *Am J Respir Crit Care Med*, 2009; 180: 547-552.

6. Quanjer PH, Stanojevic S, Stocks J, *et al*. Changes in the FEV1/FVC ratio during childhood and adolescence: an intercontinental study. *Eur Respir J*, published online March 29, 2010 as doi: 10.1183/09031936.00164109.
7. van Buuren S, Fredriks M. Worm plot: a simple diagnostic device for modelling growth reference curves. *Stat Med* 2001; 20(8): 1259-1277.