Respiratory Medicine (2007) **101**, 814–825

ELSEVIER

respiratoryMEDICINE

Reference equations for spirometry from a general population sample in central Italy☆

Francesco Pistelli[a,b,*], Matteo Bottai[c], Laura Carrozzi[a,b], Sandra Baldacci[b], Marzia Simoni[b], Francesco Di Pede[b], Giovanni Viegi[b]

[a]*U.O. Pneumologia e Fisiopatologia Respiratoria Universitaria, Dipartimento Cardio-Toracico, Azienda Ospedaliero-Universitaria Pisana, Pisa, Italy*
[b]*Unità di Epidemiologia Ambientale Polmonare, Istituto di Fisiologia Clinica (IFC), Consiglio Nazionale delle Ricerche (CNR), Pisa, Italy*
[c]*Arnold School of Public Health, University of South Carolina, Columbia, SC, USA*

Received 31 March 2006; accepted 30 June 2006

KEYWORDS
Spirometry;
Reference equations;
Forced expiratory volume in 1 s;
Vital capacity;
General population;
Airway obstruction

Summary
Aim of this study was to derive new lung function reference equations and compare the predicted values with those from three sets of existing reference equations: one derived from a Northern Italy population and the two others widely used in European (ECCS) and American (NHANES III) clinical practice.
Reference equations for flow-volume curve indexes and VC were derived on 497 normal subjects, aged 8–74, from the epidemiological survey in Pisa, Central Italy (1991–1993). By applying natural cubic splines, one single smooth and continuous equation for the entire age range was provided for each index, separately by gender.
Along with age and height, reference values also depended on BMI. Differences among the four reference equations for FEV_1, FVC, VC were quantified for average subjects. The magnitude largely varied over the age range in both genders, reaching up to half litre of air volume at specific ages. Age-gender-specific prevalence rates of airway obstruction, as defined by the ERS criterion, largely varied by applying the considered equations, the differences ranging from −3% to 28%.
The observed discrepancies confirm that reference equations should be derived from a population most similar to that for which the equations are to be used and based on measurements obtained by the same instrument and testing procedures, in order to

☆This work was supported in part by the National Research Council, Targeted Project "Prevention and Control Disease Factors—SP2—Contract no. 91.00171.PF41", the Italian Electric Power Authority (ENEL)—CNR Project "Interactions of energy system with human health and environment", Rome, Italy, and by an unrestricted research Grant from GlaxoSmithKline, Greenford, Middlesex, UK.
*Corresponding author. Unità di Epidemiologia Ambientale Polmonare, Istituto di Fisiologia Clinica CNR, Via Trieste, 41, 56126 Pisa, Italy. Tel.: +39 050 502031; fax: +39 050 503596.
E-mail address: francesco.pistelli@ifc.cnr.it.

doi:10.1016/j.rmed.2006.06.032

minimize technical variability in lung function both for clinical and epidemiological purposes.

Introduction

It is known that the variation in reference equations between studies can be quite considerable. In 1983, the European Community for Coal and Steel (ECCS)[1] derived summary reference equations by pooling existing equations from several different populations published in the previous three decades. The authors stated that no reference equation was superior to the other, and envisaged the use of the summary equations in all laboratories, as they described an overall mean of data in the literature. Conversely, in an official document of the American Thoracic Society (ATS) published in 1991,[2] it was stated that when selecting reference values for lung function testing "neither is as important as the choice of a reference population that (1) provides an appropriate comparison for the subjects to be evaluated, and (2) is based on measurements made with instruments and methods comparable to those used in the laboratory for which reference values are being selected". The same standpoint has been stated in the official document published in 2005 by the ATS/European Respiratory Society (ERS) Task Force on the standardization of lung function testing.[3]

In the last 15 yr, publication of spirometric reference values published from a variety of ethnic/race groups and age ranges has continuously increased, demonstrating the interest of world-wide researchers in deriving and using spirometry reference equations.[4–13] We previously derived smooth continuous reference equations from the Po river delta study in Northern Italy[14] ('Pistelli 2000').

In the USA, the National Health and Nutrition Examination Survey (NHANES) III reference equations ('Hankinson 1999')[7] are recommended for individuals aged 8–80 yr.[3] In Europe, the ECCS summary equations ('ECCS 1983')[1] are often used for individuals 18–70 yr but they are not currently recommended.[3] Several reports have demonstrated that 'ECCS 1983' provide lower predicted values than more recent equations from different European countries.[6,8,11,13,15]

Prevalence rates of airway obstruction vary largely in different studies, ranging from 0.2% to 18.3%.[16] The use of different criteria of airway obstruction may partly explain the variability of prevalence rates, as originally reported by us[17] and confirmed by Celli et al.[18] on the NHANES III data. An additional, not yet explored, source of variability is related to the use of different sets of reference equations, when they are applied for diagnosing airway obstruction. This is the case for the criterion based on the reduction of forced expiratory volume in 1 s (FEV_1)/slow vital capacity (VC) ratio (FEV_1/VC) with respect to its "lower normal limit". This criterion was first recommended by the ATS in 1991[2] and by the ERS in 1995,[19] and confirmed in 2005 by the ATS/ERS Task Force on the standardisation of lung function testing.[3]

The aim of this paper was to derive new lung function reference equations from the sample of the second cross-sectional survey of the Pisa study in Central Italy, and to compare them with those of 'Pistelli 2000'[14]—from a different Italian population sample—and the two most world-wide used 'ECCS 1983'[1] and 'Hankinson 1999'.[7] The implications of using different reference equations to compute prevalence rates of airway obstruction were also exemplified.

Material and methods

Reference equations were derived on 497 "normal" subjects from the second cross-sectional survey in Pisa, Central Italy[20] (1991–1993, urban area, $n = 2841$, males 45%, age range 8–97 yr).

The "normal" subjects were selected by excluding those who reported: past or current smoking history, including occasional ($n = 1397$ remaining subjects); occupational exposure to noxious agents >5 yr ($n = 1308$ remaining subjects); respiratory symptoms, i.e. cough, phlegm, wheeze, effort dyspnea, attacks of breathlessness with wheeze or whistles ($n = 779$ remaining subjects); medical diagnosis of emphysema, asthma or chronic bronchitis ($n = 772$ remaining subjects); chest injuries ($n = 754$ remaining subjects); heart troubles ($n = 663$ remaining subjects). Further, were excluded those who did not meet measurement protocol ($n = 497$ remaining subjects). Data were obtained by the interviewer-administered CNR-questionnaire.[20]

A water-sealed spirometer (Baires, Biomedin, Padua, Italy) was used for measuring flow and volume. The same instrument was utilized in the European Community Respiratory Health Survey.[6] This instrument has a frequency response of 30 Hz and the linearity of the transducer of position is 0.25%. The day-by-day variation in calibration observed over time for this instrument ranges between 0% and 1.8% with a mean error of 0.4%. Volume calibration was performed daily by using a 3-l standard syringe. A team of six nurses/technicians used two instruments of the same type in two different locations. All the nurses/technicians received the same training for performing lung function tests. Instrument characteristics and measurement protocol fulfilled the 1987 American Thoracic Society recommendations.[21] At least two trials were repeated to obtain a satisfactory VC value. The highest VC was used for analyses. Up to eight-forced vital capacity (FVC) maneuvers were performed to obtain at least three acceptable trials. The two largest FVC and FEV_1 values should not vary by more than 5%. The largest FVC and FEV_1 values were selected, regardless of the maneuver. From the maneuver with the largest sum FVC+FEV_1 the following flows were selected: peak expiratory flow (PEF); forced expiratory flow between 25% and 75% of FVC ($FEF_{25-75\%}$); maximal expiratory flow at 50% (MEF50%) and 75% (MEF75%) of FVC. As a further quality criterion, the largest FVC and the largest VC should not vary by more than 5%. In the present study, VC_{max} was also computed as the largest value of VC and FVC for each individual.

Height (cm) and weight (kg) were measured in standing position without shoes in subjects wearing clothes. Age at last birthday was recorded. Body mass index (BMI) was computed as weight/height-squared, expressed in kg/m-squared.

Statistical analysis

For each index, one single smooth continuous equation was derived for each gender.[14] Independent variables were age, height and BMI. The non-linear relationship between age and lung function indexes was modelled by natural cubic splines for all indexes, except for FEV_1/FVC and FEV_1/VC whose dependence on age appeared to be linear.

Natural cubic splines provide smooth curves with continuous first and second derivatives over the whole age range. In the extreme intervals, where fewer observations are available, natural cubic splines are constrained to be linear. Details about computation of the predicted values are reported in the Appendix. The two breakpoints used for the different indexes were chosen, based on a grid search, as those which maximized the goodness-of-fit as measured by R^2. BMI and height showed a curvilinear relationship with all indexes that was adequately fit by second-order polynomials. For all the indexes the lower normal limit was expressed by the "normal fifth percentile", which was defined as the percent predicted value below which fell 5% of the sample of normal subjects.[2,14] Exact binomial confidence intervals about the point estimate for the "normal fifth percentile" were also provided. All pair-wise interactions were not significant at 0.05-level and kept out of the models.

Prediction plots of FEV_1 versus age are shown for "average" subjects. Since, at any given age, height and BMI vary across subjects, height and BMI were averaged by using Lowess smoothing (robust locally weighted regressions) of height and BMI on age.[22] Prediction plots of FEV_1 versus height and BMI were analogously computed.

We performed an internal cross-validation of the predicted values, separately in males and females. The sample of "normal" subjects was split in 10 equally sized sub-samples at random and new regression models were estimated by using data from nine of the ten sub-samples (90% of the entire dataset), leaving one sub-sample out (10% of the entire dataset). The differences and the ratios between the observed values and the predicted values were calculated for the sub-sample that was left out ('validation sample'). Such a procedure was repeated ten times, one for each sub-sample in turn.

Bland Altman plots of FEV_1 and FVC predicted differences by means, for the new equations and 'Pistelli 2000', 'ECCS 1983', and 'Hankinson 1999' were included.[23]

All statistical analyses were performed in Stata software (Stata Corporation, College Station, TX, USA).

Results

Derivation of new reference equations

Table 1 shows the characteristics of the sample of "normal" subjects used to derive the reference equations. Males were on the average younger, taller, heavier and had larger respiratory volumes and flows than females: differences were statistically significant. On the contrary, the average ratios FEV_1/VC and FEV_1/FVC were similar as was the case for BMI. In both genders, FVC mean values were slightly

Table 1 Descriptive statistics of the normal subjects used to derive the new reference equations.

	Males ($n = 196$)				Females ($n = 301$)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Age (yr)[a]	26.89	15.49	8	70	36.94	19.25	8	74
BMI (kg/m^2)	23.19	3.74	14.57	34.26	23.62	4.14	14.72	39.74
Height (cm)[a]	171.63	11.85	125	195	160.56	7.46	125	181
Weight (kg)[a]	69.10	15.32	25	100	60.99	11.42	25	103
VC (l)[a]	4.78	1.24	1.66	7.77	3.48	0.66	1.43	5.40
FVC (l)[a]	4.92	1.20	1.65	7.92	3.54	0.65	1.49	5.41
VC_{max} (l)[a,b]	4.95	1.20	1.66	7.92	3.56	0.65	1.49	5.41
FEV_1 (l)[a]	4.16	0.99	1.46	6.27	2.98	0.62	1.36	4.85
FEV_1/VC	0.86	0.07	0.65	0.98	0.85	0.08	0.62	0.98
FEV1/FVC	0.85	0.07	0.67	0.98	0.84	0.07	0.62	0.98
FEF25–75 (l/s)[a]	4.68	1.43	1.40	8.74	3.37	1.09	0.70	5.93
PEF (l/s)[a]	8.33	2.03	2.65	12.79	5.94	1.23	3.09	8.99
MEF50 (l/s)[a]	5.43	1.56	1.96	10.00	3.98	1.06	1.38	6.56
MEF75 (l/s)[a]	2.42	1.07	0.50	6.32	1.69	0.87	0.18	3.99

[a] $P<0.05$ by Wilcoxon rank-sum test.
[b] VC_{max} is defined for each individual as the maximum value between VC and FVC.

higher than VC mean values, as were VC_{max} mean values with respect to FVC mean values.

Table 2 reports the estimated coefficients for VC, FEV_1, FVC, FEV_1/FVC, FEV_1/VC from the new equations. In both sexes, predicted values of VC, FEV_1 and FVC showed a parabolic nonlinear relationship with BMI and height and a more complex nonlinear association with age. For the two ratios, age was a significant linear predictor. Nearly all the estimated coefficients were statistically significant. The "normal fifth percentile" values ranged from 82% to 90% predicted.

The nonlinear relationships between predicted values and predictors are shown in Fig. 1 for FEV_1. Predicted values steeply increased with age until about 25 yr in males (Fig. 1a) and 20 yr in females (Fig. 1d), and steadily decreased afterwards in a continuous and smooth fashion. Furthermore, FEV_1 predicted values increased according to a slightly convex parabolic trend as height increased both in males (Fig. 1b) and females (Fig. 1e), while they slowly increased up to a BMI value of about 24 kg/m^2 in males (Fig. 1c) and about 20 kg/m^2 in females (Fig. 1f) and decreased afterwards.

The estimated coefficients for the remaining spirometric indexes are shown in Table 3. Except for age, the associations with the independent variables were seldom statistically significant, largely due to the high variability of respiratory flows measurements. However, BMI and height showed trends that were consistent with those shown in Table 2. The "normal fifth percentile" was 76% and 72% predicted for PEF in males and females, respectively, and within 46–63% for the rest of the indexes.

In the cross-validation of the prediction for FEV_1 and FVC, we calculated the differences and the ratios between predicted values and observed measurements on the ten 'validation samples.' Across the ten 'validation samples,' differences (ratios) for FEV_1 varied between −0.13 (−5%) and 0.14 l (6%) in females, and between −0.19 (−3%) and 0.19l (7%) in males; for FVC they varied between −0.16 (−3%) and 0.13l (4%) in females and between −0.24 (−4%) and 0.24l (7%) in males.

Comparison with existing reference equations

Figure 2 shows the differences as age varies between predicted values for FEV_1, FVC and VC from the new equations, and those from 'Pistelli 2000', 'ECCS 1983' and 'Hankinson 1999', separately for normal males and females of the Pisa study.

Differences were calculated as predicted values for average subjects derived by the new equations minus

Table 2 Estimated coefficients associated with the independent variables for different volumes and ratios indexes, separately for the two genders.

	VC	FEV_1	FVC	FEV_1/FVC	FEV_1/VC
Males					
Constant	11.1943	5.43862^{ns}	13.18461	−0.0213471^{ns}	−0.0553481^{ns}
BMI	0.2310701	0.1918785	0.2285536	0.0034452^{ns}	0.0013997^{ns}
BMI^2	−0.0037232^{ns}	−0.0034633	−0.0036842^{ns}	−0.0001494^{ns}	−0.0000977^{ns}
Height	−0.2087963	−0.1241893	−0.2355706	0.0116165	0.0120524^{ns}
Height^2	0.0008009	0.0005119	0.0008819	−0.0000359	−0.0000362^{ns}
Age	0.1621762	0.1603364	0.1834322	−0.00211	−0.0023676
$Spline_1$(age)	0.0010401	0.0011688	0.001246	—	—
$Spline_2$(age)	−0.0007796	−0.0008873	−0.0009158	—	—
(breakpoints)	(21, 25)	(21, 25)	(20, 24)	—	—
Normal fifth percentile					
(% of predicted)	82	83	82	87	88
[95% confidence interval]	[80,84]	[79,85]	[79,84]	[84,91]	[84,90]
Females					
Constant	4.702777^{ns}	2.714397^{ns}	4.472363^{ns}	−0.1297842^{ns}	−0.0454675^{ns}
BMI	0.1784679	0.0852553	0.1894842	−0.0214414	−0.0080805^{ns}
BMI^2	−0.0032607	−0.0015578	−0.0033201	0.0003665	0.0000928^{ns}
Height	−0.1116252^{ns}	−0.0710514^{ns}	−0.1074871^{ns}	0.0175565	0.0148438^{ns}
Height^2	0.0004762	0.0003158^{ns}	0.0004603	−0.000057	−0.0000489^{ns}
Age	0.1752335	0.2189869	0.155336	−0.0022295	−0.0024913
$Spline_1$(age)	0.0018615	0.0025242	0.0016162	—	—
$Spline_2$(age)	−0.0009931	−0.0011336	−0.0010608	—	—
(breakpoints)	(14, 19)	(13, 19)	(16, 20)	—	—
Normal fifth percentile					
(% of predicted)	83	83	83	90	89
[95% confidence interval]	[81,85]	[81,86]	[81,85]	[87,91]	[87,91]

For a numerical example of the computation of predicted values, see Appendix.
^{ns}Non-significant: $P > 0.10$ by Wald test.

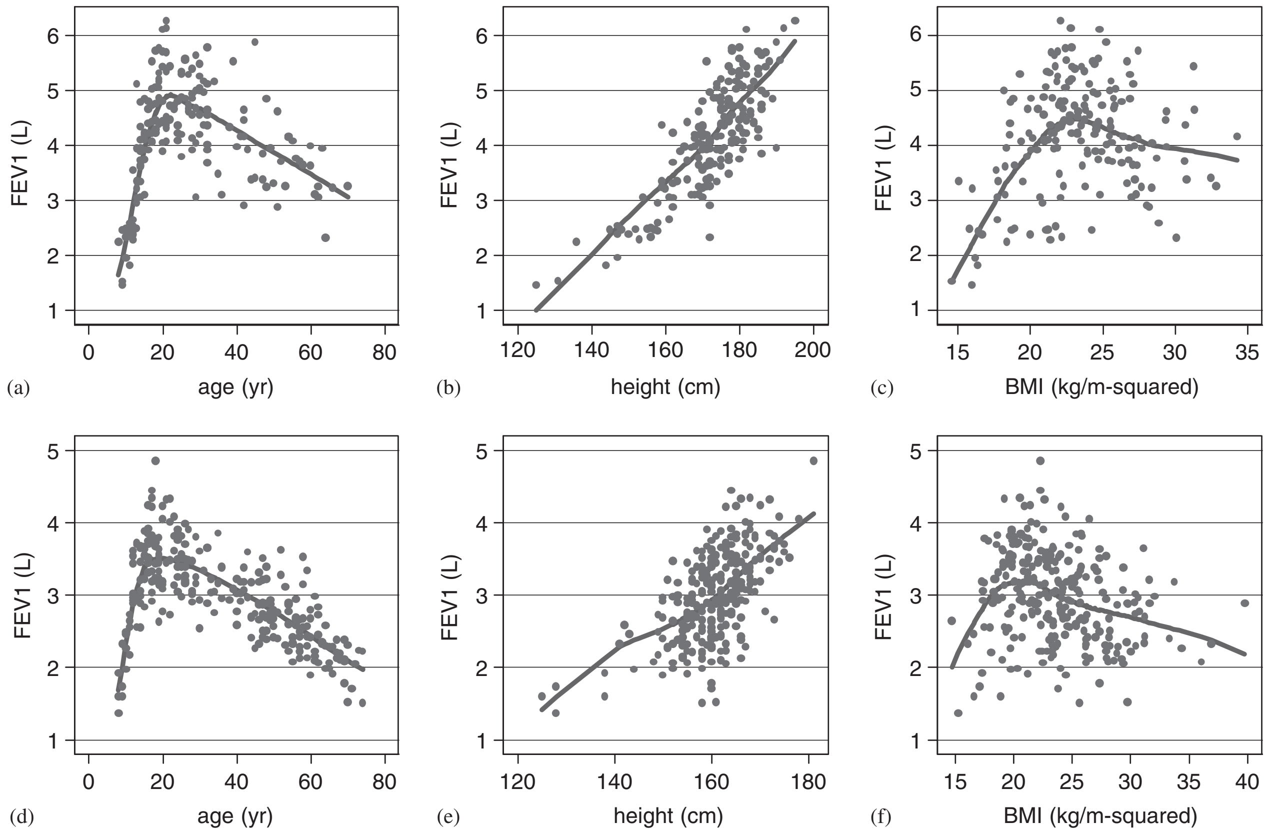

Figure 1 Predicted values for average subjects (continuous curves) and observed values (dots) for FEV_1 in males (a–c) and in females (d–f) as age (a, d), height (b, e) and BMI (c, f) vary.

predicted values derived by each of the existing equations. The horizontal line at zero indicates no difference. Curves above the ''no difference'' line indicate larger values of the new equations than those of the comparing equation, while curves below indicate smaller values.

The new FEV_1 predicted values were higher than 'Pistelli 2000' until the age of 60 in males and 65 in females, and lower afterwards. The highest difference was for FEV_1 around the age of 20 (about +400 ml in males, and +300 ml in females), and for VC over the age of 60 (about–500 ml in males, and −300 ml in females).

In both genders, the new predicted values for FEV_1, FVC and VC were nearly always higher than 'ECCS 1983' across the whole age span. The highest difference was for FVC around the age of 50 (about +500 ml in both genders).

Differences with 'Hankinson 1999' showed sharp, sudden changes, since 'Hankinson 1999' predictions were obtained from different equations for different age ranges (stratified regressions). At age 18, the difference in FEV_1 with the new equations was about +200 ml in males and +300 ml in females.

Differences in predictions for FEV_1/FVC and FEV_1/VC are shown in Fig. 3. ECCS[1] and Hankinson et al.[7] did not provide equations for FEV_1/FVC and for FEV_1/VC, respectively. Differences between predicted values calculated according to the different sets of reference equations also varied along with different values of BMI and height, for all the indexes (results not shown).

Figures 4–7 present the Bland-Altman plots of the differences by means, separately by sex, for FEV_1, FVC, FEV_1/FVC, and FEV_1/VC. Generally, the predicted values with the new equations tended to be larger than the others, especially for larger mean values of FEV_1 and FVC (Figs. 4 and 5). For FVC, the new equations were more similar to 'Pistelli 2000' and 'Hankinson 1999', than to 'ECCS 1983' (Fig. 5). The latter were steadily lower than the new equations over the whole range of mean values of both FEV_1 and FVC (Figs. 4 and 5). The predictions for FEV_1/FVC with the new equations were more similar to 'Hankinson 1999' than to 'Pistelli 2000' (Fig. 6). The predictions for FEV_1/VC were closer to 'ECCS 1983' than to 'Pistelli 2000' (Fig. 7).

Effects on airway obstruction prevalence

To exemplify the relevance of the choice of the reference equations, we calculated prevalence rates of airway obstruction on the whole general population samples (i.e. not only on the ''normal'' individuals) of Pisa and Po river delta (Fig. 8). Only adult subjects aged over 24 yr were included, for they are those who are expected to present sizable obstruction rates.[17] Airway obstruction was defined according to ERS criterion[19]: $FEV_1/VC < 0.88$ of predicted value in males and $FEV_1/VC < 0.89$ of predicted value in females. FEV_1/VC predicted values were obtained by

Table 3 Estimated coefficients associated with the independent variables for different flows indexes, separately for the two genders.

	PEF	$FEF_{25-75\%}$	$MEF_{50\%}$	$MEF_{75\%}$
Males				
Constant	−6.123093^{ns}	−4.289687^{ns}	1.733825^{ns}	−9.499382^{ns}
BMI	0.4512205	0.3887155	0.5291521	0.0882047^{ns}
BMI-squared	−0.0080618^{ns}	−0.0077062^{ns}	−0.010235	−0.0024357^{ns}
Height	−0.0179494^{ns}	−0.0133726^{ns}	−0.1029152^{ns}	0.0924414^{ns}
Height-squared	0.0001547^{ns}	0.0001343^{ns}	0.0003808^{ns}	−0.0001753^{ns}
Age	0.4269676	0.1779775	0.2273209	0.0606975^{ns}
$Spline_1$(age)	0.0010371	0.0005379	0.0007703	0.0003964
$Spline_2$(age)	−0.0004952	−0.0002835	−0.0004507	−0.0002528
(breakpoints)	(20, 33)	(21, 34)	(21, 30)	(21, 30)
Normal fifth percentile				
(% of predicted)	76	60	61	53
[95% confidence interval]	[69,80]	[56,68]	[57,71]	[42,58]
Females				
Constant	−4.061646^{ns}	−1.452058^{ns}	3.194874^{ns}	−4.400701^{ns}
BMI	0.0575598^{ns}	−0.034093^{ns}	0.0459279^{ns}	−0.1503959
BMI^2	−0.0012472^{ns}	0.0004008^{ns}	−0.0009887^{ns}	0.0024881^{ns}
Height	0.0078681^{ns}	−0.0178674^{ns}	−0.0865969^{ns}	0.0583972^{ns}
$Height^2$	0.0001057^{ns}	0.0001086^{ns}	0.0003046^{ns}	−0.0001485^{ns}
Age	0.4885604	0.5601075	0.6128726	0.2876721
$Spline_1$(age)	0.008834	0.0159724	0.0255259	0.0054201
$Spline_2$(age)	−0.0029044	−0.0039649	−0.0169722	−0.0024136
(breakpoints)	(11, 17)	(10, 16)	(12, 14)	(12, 17)
Normal fifth percentile				
(% of predicted)	72	62	63	46
[95% confidence interval]	[69,76]	[57,67]	[60,67]	[44,50]

For a numerical example of the computation of predicted values, see Appendix.
^{ns}Non-significant: $P > 0.10$ by Wald test.

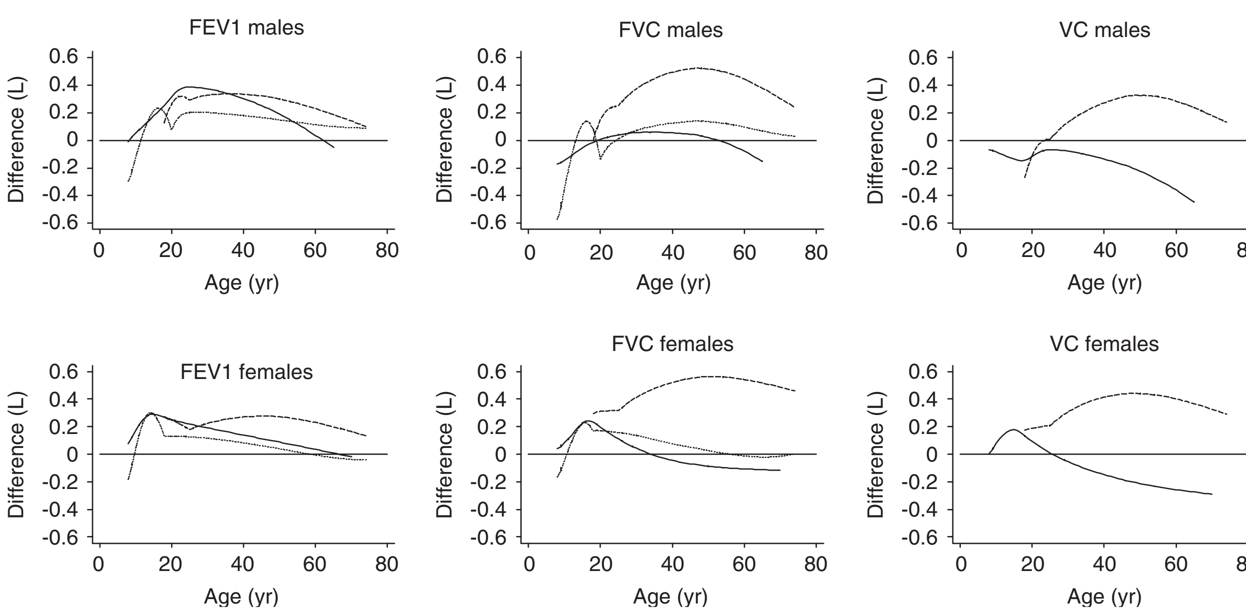

Figure 2 Differences in predicted values as age varies for FEV_1, FVC, VC, separately for males (top panels) and females (bottom panels), between the new equations and those of 'Pistelli 2000' (solid curve), 'ECCS 1983' (dashed curve) and 'Hankinson 1999' (dotted curve). (Hankinson et al.[7] did not provide equations for VC).

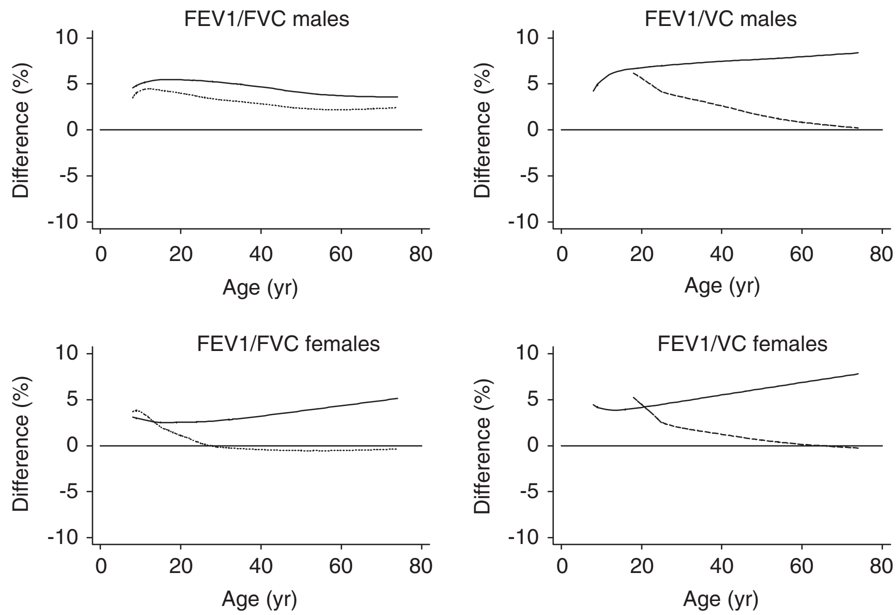

Figure 3 Differences in predicted values as age varies for FEV_1/FVC and FEV_1/VC, separately for males (top panels) and females (bottom panels), between the new equations and those of 'Pistelli 2000' (solid curve), by 'ECCS 1983' (dashed curve) and of 'Hankinson 1999' (dotted curve). (ECCS[1] and Hankinson et al.[7] did not provide equations for FEV_1/FVC and for FEV_1/VC, respectively.).

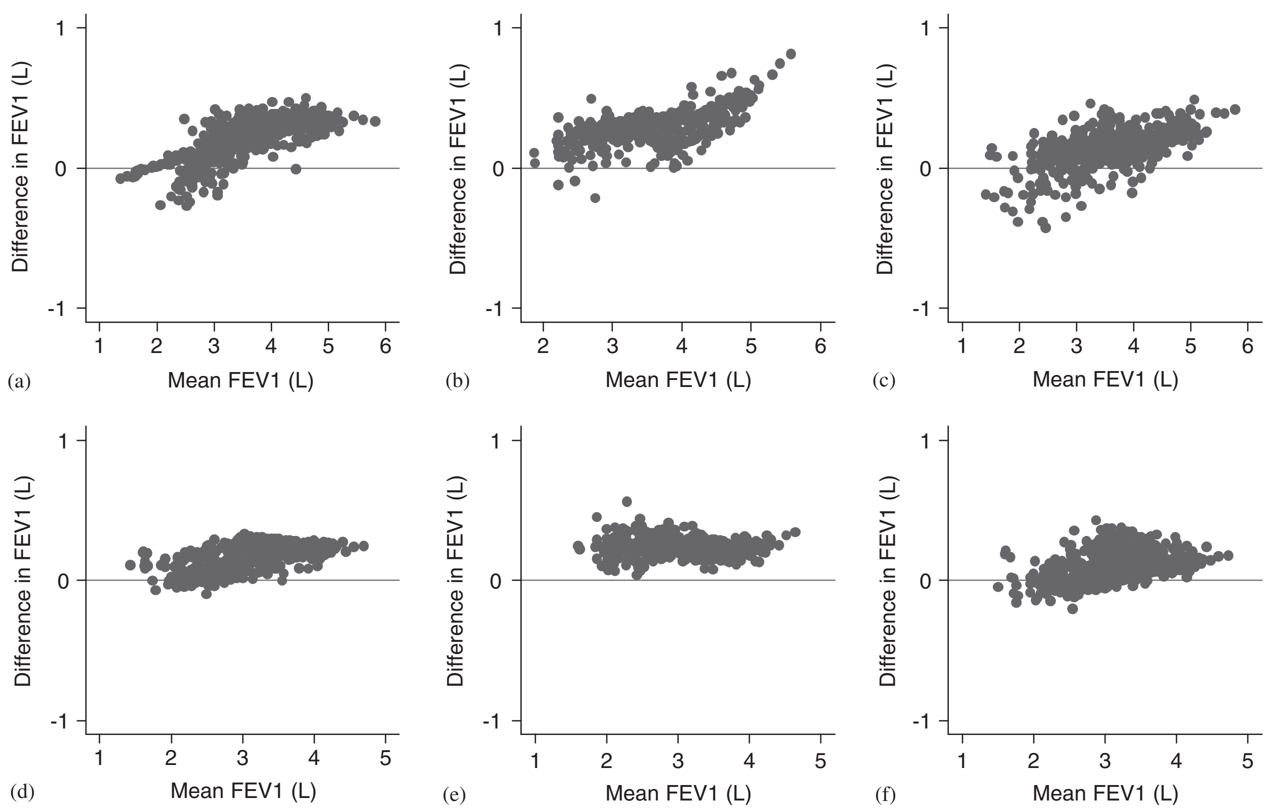

Figure 4 Bland-Altman plots for predictions of FEV_1 based on the new equations and 'Pistelli 2000' (a, d), 'ECCS 1983' (b, e), and 'Hankinson 1999' (c, f), for males (a–c) and females (d–f).

the three European sets of reference equations only, i.e. the new ones, 'Pistelli 2000' and 'ECCS 1983'. Reference equations of 'Hankinson 1999' were not considered, as the authors did not provide reference equations for FEV_1/VC.

Prevalence rates of airway obstruction largely varied depending on which reference equation was applied. Differences between the age-gender-specific prevalence rates obtained from the new and the existing equations ranged from −3% to 28%.

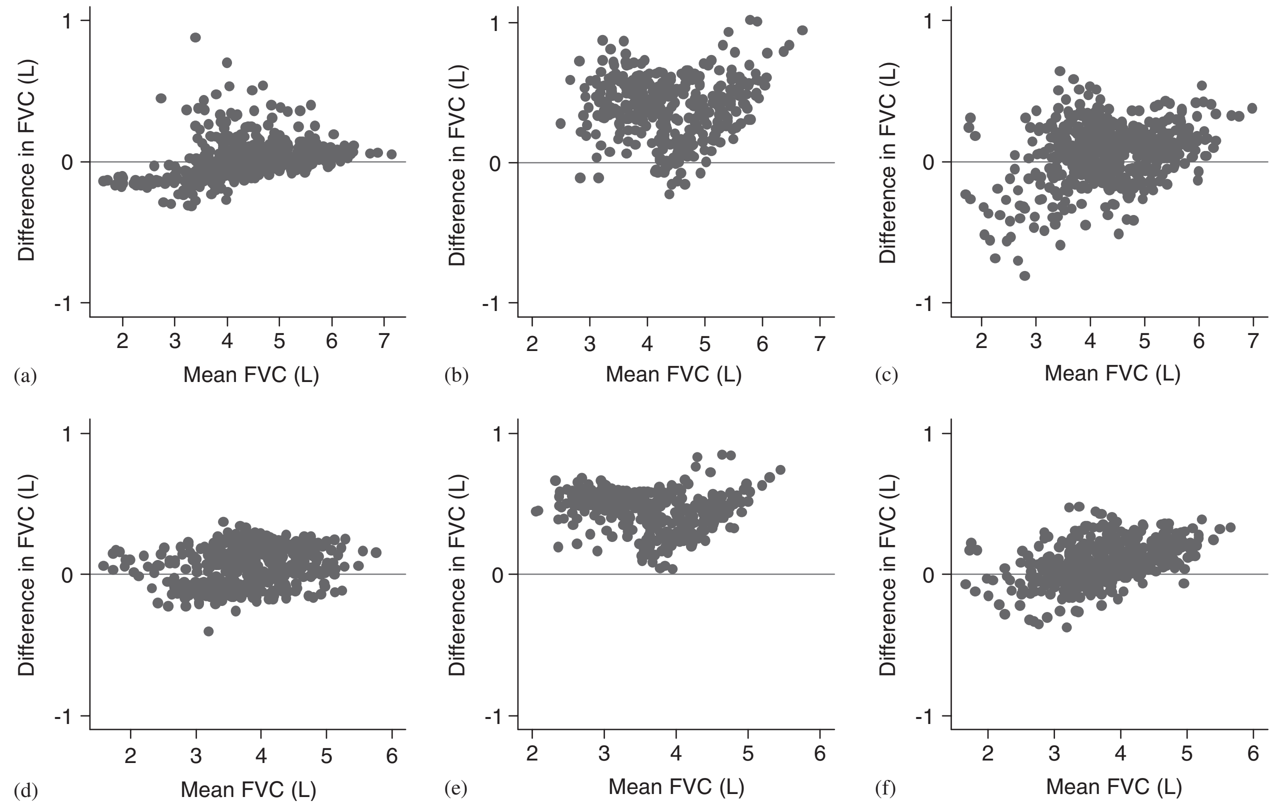

Figure 5 Bland-Altman plots for predictions of FVC based on the new equations and 'Pistelli 2000' (a, d), 'ECCS 1983' (b, e), and 'Hankinson 1999' (c, f), for males (a–c) and females (d–f).

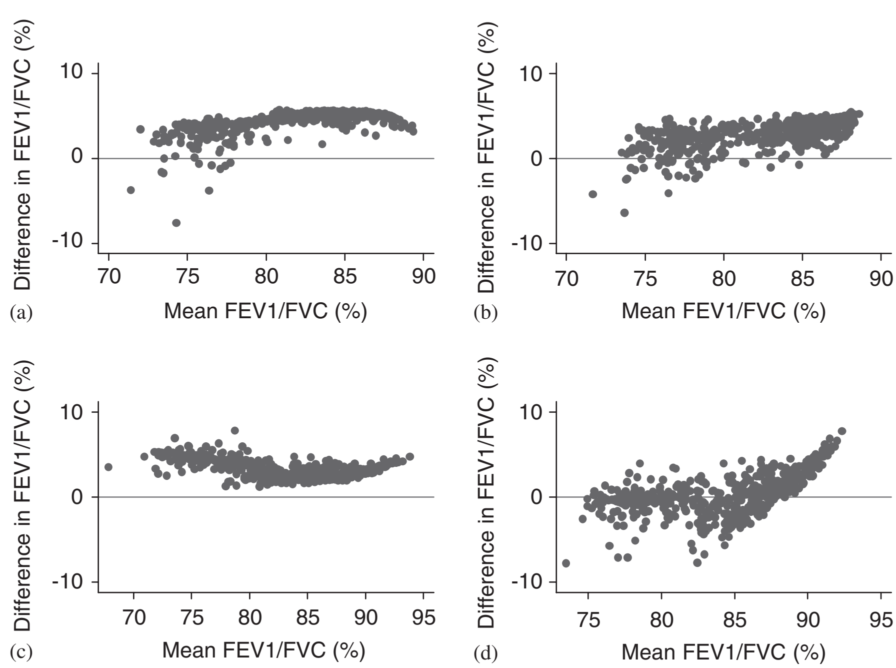

Figure 6 Bland-Altman plots for predictions of FEV_1/FVC % based on the new equations and 'Pistelli 2000' (a,c) and 'Hankinson 1999' (b,d), for males (a,b) and females (c,d). (ECCS[1] did not provide equations for FEV_1/FVC.)

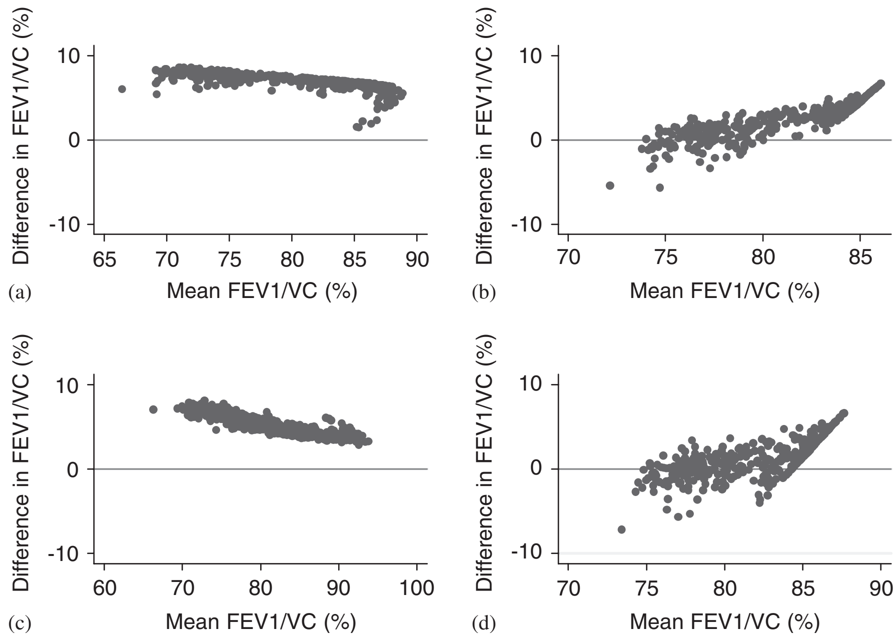

Figure 7 Bland-Altman plots for predictions of FEV_1/VC % based on the new equations and 'Pistelli 2000' (a, c) and 'ECCS 1983' (b, d), for males (a, b) and females (c, d). (Hankinson et al.[7] did not provide equations for FEV_1/VC.)

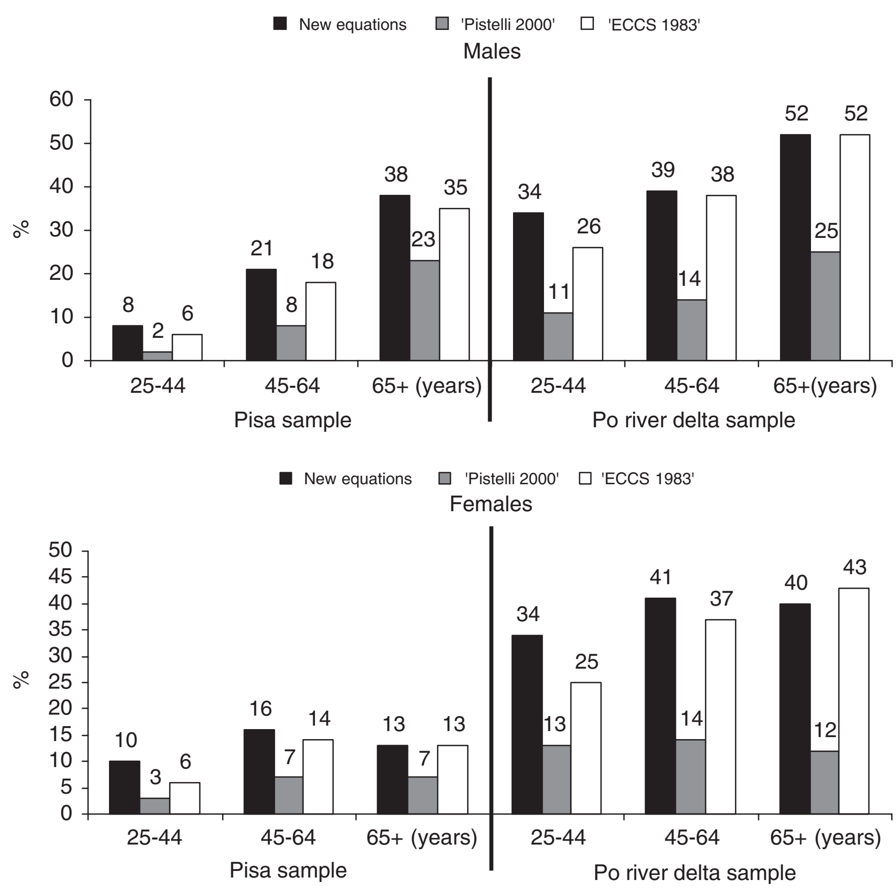

Figure 8 Prevalence rates of airway obstruction (%), as defined by using the ERS criterion[19] (FEV_1/VC <88% of predicted value in males and <89% of predicted value in females) with different reference equations, in two Italian general population samples, by age groups and gender.

Discussion

Derivation of new reference equations

We derived reference equations for flow-volume curve indexes and VC from the 497 ''normal'' subjects participating in the second cross-sectional survey of Pisa study in Central Italy, carried out in 1991–1993. Natural cubic splines were used to model the nonlinear relationship of the lung function with age. They provide one single equation for the entire age range. Differently from multiple equations derived for age intervals, they provide predicted values smoothly varying with age, which is without discontinuities or angled points. As also previously demonstrated,[14] reference equations that are continuous and smooth over the age range may avoid biologically implausible jumps in predictions associated with consecutive age intervals. On the contrary, discontinuous equations may yield different diagnoses for a patient who is examined twice within a short period of time, who reports the same exact measurements twice, whose predicted values jump from an age interval to the next. For sake of completeness, for VC, FVC, and FEV_1, separately in males and females, we estimated two separate regressions, above and below age 20. Their overall mean squared error for prediction was almost as good as, never better than, that of the continuous curves.

Lacking of a validation sample, we performed an internal cross-validation of the predicted values, separately in males and females. Overall, the predictions on new, independent subjects were within −5% to 7%, which appeared reasonably small. In addition, the means of the ratio of the validation samples were symmetrically distributed about 100%. Such variability in each sample was expected, since the prediction, conditional on sex, age, height and BMI, is an estimate of the population conditional mean. Each individual may differ from the population mean. Even if the model we estimated was perfectly true, we would see a residual error in the individual prediction, which is often known as variability of the individual forecast.

Along with age and height, BMI was also introduced as independent variable in the reference equations. BMI is also included among independent variables in reference equations for arterial oxygen tension[24] and the 6-min walking test in healthy adults.[25] BMI proved to be a significant predictor that improved prediction capabilities, as also noted for the equations we derived on a rather different population.[14] Contrarily, BMI was considered but not included among independent variables in other reference equations.[7] From our findings, BMI appears to be an important predictor of lung function both in youth, where it provides a proxy to body growth, not fully explained by age and height alone, and in adults, where it accounts for differences in physical complexion and obesity.[26] In our data, there is no evidence that, on average, the effect of BMI on lung function is not smooth over time; there is no clear cut-off age at which this effect may change. Indeed, the age of transition from childhood to adulthood varies naturally across different individuals. Setting a cut-off value to split the age range in childhood/adolescence and adulthood would be necessarily arbitrary and detrimental to the accuracy of the prediction. Also, as shown by Bottai et al.[27] variations of lung function in adults over time are influenced by variations of BMI, independently of age, smoking habits, respiratory symptoms and diseases, and other individual characteristics.

As also noted elsewhere,[15] the lung vital capacity may differ whether measured by the slow (VC) or forced expiration (FVC) maneuvers, VC being usually higher than FVC. By using a pneumotachograph, we found FVC higher than VC in 35% and 36%, for males and females respectively, among 1185 longitudinal observations provided by 1039 ''normal'' subjects on which the 'Pistelli 2000' equations were derived.[14] In the present sample, by using a water sealed spirometer, we observed a not negligible proportion of individuals with FVC greater than VC (58% in males, 50% in females). On average, FVC was 3% higher than VC in males and 2% in females, i.e. values within the threshold of acceptable variability (5%),[2] as described in the section 'Material and methods'. These results may be accounted for by instrument characteristics, measurement procedures and inter- or intra-individual variability. For sake of completeness, we also reported the summary statistics of the individual maximum value between VC and FVC (VC_{max}).

It is to point out that, the lower confidence limits for volumes and ratios, below which lung function is considered abnormal, were almost identical for the Pisa study and the Po river delta study and the same as those indicated by ERS to define airway obstruction.[19]

Comparison with existing reference equations

We compared the new equations with those of 'Pistelli 2000', 'ECCS 1983' and 'Hankinson 1999', on the whole sample of ''normal'' subjects from Pisa study. Differences among the equations for FEV_1, FVC, VC and their ratios were quantified for average subjects. Bland-Altman plots were produced as additional visual comparison. These equations were chosen, among several others, because the former were derived from another Italian population sample by using the same statistical model as the one used for the new equations, while the latter two equations are the most commonly used equations in lung function laboratories in Europe and United States.

Different predicted values between the new and 'Pistelli 2000' equations can be imputed to the joint effects of several factors: different areas (urban vs. rural), different cohorts (cross-sectional sample—survey time 1991–1993 vs. pooled cross sectional and longitudinal samples—survey time 1980–1982 and 1988–1991), different technicians and different instruments used (water sealed spirometer vs. pneumotachograph).[28,29]

It is known[30,31] that the spirometer generally tends to underestimate lung volumes when compared to the pneumotachograph. A small study, comparing three different instruments on the same four subjects,[32] reported that the water sealed spirometer used in the present study provides smaller measurements of FEV_1, FVC and VC and higher values of FEV_1/FVC and FEV_1/VC ratios than the pneumotacograph used in the previous study.[14] The same instrument used in the present study was utilized in many centres participating in the European Community Respiratory Health Survey.[6] Within the study conducted to derive spirometric reference equations from that population, a specific comparison among instruments used in the various participating centres was not performed.[6]

Consistently with other reference equations derived from younger cohorts of European populations,[6,8,13,15] the new equations appeared to provide higher predictive values than 'ECCS 1983'. The new equations differ from 'ECCS 1983' mostly for the materials and methods used, since the latter were based on published regression equations—instead of the original data—from older surveys (published between 1950s and 1980s), which employed different instruments. Besides, they did not include BMI.

'Hankinson 1999' equations[7] for Caucasians were based on an age range similar to the new equations. The instrument used was a dry rolling-seal spirometer. Regression models stratified by age groups were used instead of smooth equations. Differences between the predictions can be primarily, although not exclusively, attributed to the different methods applied (stratified regressions vs. smooth equations) and the populations studied.

Overall, differences among the predicted values of different equations remarkably varied over the range of values of age, height and BMI. Although the reference equations considered were derived from populations that differed with respect to age, height and BMI, these characteristics were accounted for in the models, and cannot explain the differences in the predictions.

Effects on airway obstruction prevalence

It seems that using equations that are not population-specific and instrument-specific (i.e. the new ones for the Po river delta sample and 'Pistelli 2000' for the Pisa sample) may lead to biased estimates of prevalence. It also appears that the use of summary equations, such as the 'ECCS 1983', may not be appropriate for all possible populations. Indeed, the 'ECCS 1983' provided rates that were closer to the new ones but very different from 'Pistelli 2000'.

It should be noted that the lower limit of normal vary across different reference equations. The use of one unique lower limit of normal should be discouraged in the clinical settings, as recently recommended by ATS/ERS Task Force on the standardisation of lung function testing,[3] which defines airways obstruction as a reduced FEV_1/VC ratio below the fifth percentile of the predicted value.

Conclusions

According to current recommendations from ATS/ERS Task Force on the standardisation of lung function testing,[3] as well as other recent reports,[11–13] our results reinforce the need of applying reference equations derived from samples of "normals" who are as similar as possible to the study subject/population on which such equations are applied, and based on measurements obtained by the same instruments and protocols. Variations in time, space, instruments or population characteristics may result in biased predictions. Therefore it would be recommendable that such information be reported in detail whenever new reference equations are provided. Finally, our results confirm the need of periodically updating reference equations for lung function.

Acknowledgments

The authors wish to thank Dr. Paolo Paoletti and Prof. Carlo Giuntini (University of Pisa) for their contribution in planning the epidemiological studies of Pisa and Po river delta. The authors also thank all the persons who have concurred to the realization of the epidemiological studies of Pisa and Po river delta, and who have collaborated to the collection and elaboration of the data. Finally, they thank the residents of Pisa and Po river delta who participated in the studies.

Appendix A. Natural cubic splines regression models

The age span was split into three phases of lung function (growth, plateau, and decline) by means of two breakpoints. The regression model for the expected value of any lung function index was:

$$\beta_0 + \beta_1\text{BMI} + \beta_2\text{BMI}^2 + \beta_3\text{height} + \beta_4\text{height}^2 + \beta_5\text{age} + \beta_6\text{Spline1(age)} + \beta_7\,\text{Spline 2(age)},$$

where βs are parameters to be estimated, and $\text{Spline}_1(\cdot)$ and $\text{Spline}_2(\cdot)$ are natural cubic splines computed as follows:

$$\text{Spline}_j(\text{age}) = \frac{-(\text{age}-8)^3 - (74-\text{breakpoint}_j)}{(74-8)} + (\text{age}-\text{breakpoint}_j)^3 \cdot I(\text{age} \geqslant \text{breakpoint}_j),$$

where $j = 1, 2$, and $I(A)$ denotes the indicator function that equals one if A is true and zero otherwise; $\text{Spline}_1(\cdot)$ and $\text{Spline}_2(\cdot)$ were computed considering that the age range in the sample was 8–74 yr in females. The two breakpoints were different for each regression model and were estimated to maximize the goodness-of-fit in terms of R^2.

Example: calculation of predicted values for FEV_1

To exemplify the use of the estimated coefficients shown in Table 2, suppose we want to compute the predicted FEV_1 value for a man aged 32 yr whose height and weight are 168 cm and 57 kg, respectively. Then, we compute

$$\text{BMI} = 57/1.68^2 = 20.195578\,\text{kg/m}^2,$$

$$\text{Spline}_1(32) = -(32-8)^3(70-21)/(70-8) + (32-21)^3 = -9770.0909,$$

$$\text{Spline}_2(32) = -(32-8)^3(70-25)/(70-8) + (32-25)^3 = -9920.2727.$$

From the coefficients reported in Table 2,

$$\begin{aligned} FEV_1 \text{ predicted} = {} & 5.43862 + 0.1918785\,(20.195578) \\ & - 0.0034633\,(20.195578)^2 \\ & - 0.1241893\,(168) + 0.0005119\,(168)^2 \\ & + 0.1603364\,(32) + 0.0011688\,(-9770.0909) \\ & - 0.0008873\,(-9920.2727) = 4.00\,\text{l}. \end{aligned}$$

Note. The present reference equations should be used within the ranges observed in our samples as shown in the following table (also shown in Table 1).

	Age (yr)	BMI (kg/m^2)	Height (cm)
Males	8–70	14.57–34.26	125–195
Females	8–74	14.72–39.74	125–181

References

1. Quanjer PhH, Dalhuijsen A, Van Zomeren BH. Summary equations of reference values. *Bull Europ Physiopathol Respir* 1983;**19**(Suppl. 5):45–51.
2. American Thoracic Society. Lung Function Testing: Selection of Reference Values and Interpretative Strategies. *Am Rev Respir Dis* 1991;**144**:1202–18.
3. Pellegrino R, Viegi G, Enright P, et al. ATS/ERS task force: standardisation of lung function testing. interpretative strategies for lung function tests. *Eur Respir J* 2005;**26**:948–68.
4. Enright PL, Kronmal RA, Higgins M, et al. Spirometry reference values for women and men 65 to 85 years of age. Cardiovascular health study. *Am Rev Respir Dis* 1993;**147**:125–33.
5. Sharp DS, Enright PL, Chiu D, et al. Reference values for pulmonary function tests of Japanese-American men aged 71 to 90 years. *Am J Respir Crit Care Med* 1996;**153**:805–11.
6. Roca J, Burgos F, Sunyer J, et al. References values for forced spirometry. Group of the European Community Respiratory Health Survey. *Eur Respir J* 1998;**11**:1354–62.
7. 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**:179–87.
8. Langhammer A, Johnsen R, Gulsvik A, et al. Forced spirometry reference values for Norwegian adults: the Bronchial Obstruction in Nord-Trondelag Study. *Eur Respir J* 2001;**18**:770–9.
9. Golshan M, Nematbakhsh M, Amra B, et al. Spirometric reference values in a large Middle Eastern population. *Eur Respir J* 2003;**22**:529–34 [Erratum in: Eur Respir J 2004; 23: 356].
10. Pistelli R, Bellia V, Catalano F, et al. Spirometry reference values for women and men aged 65–85 living in southern Europe: the effect of health outcomes. *Respiration* 2003;**70**:484–9.
11. Garcia-Rio F, Pino JM, Dorgham A, et al. Spirometric reference equations for European females and males aged 65–85 yr. *Eur Respir J* 2004;**24**:397–405.
12. Fulambarker A, Copur AS, Javeri A, Jere S, Cohen ME. Reference values for pulmonary function in Asian Indians living in the United States. *Chest* 2004;**126**:1225–33.
13. 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**:456–63.
14. Pistelli F, Bottai M, Viegi G, et al. Smooth reference equations for slow vital capacity and flow-volume curve indexes. *Am J Respir Crit Care Med* 2000;**161**:899–905 [Erratum in: Am J Respir Crit Care Med 2001;164:1740].
15. Paoletti P, Pistelli G, Fazzi P, et al. Reference values for vital capacity and flow-volume curves from a general population study. *Bull Eur Physiopathol Respir* 1986;**22**:451–9.
16. Halbert RJ, Isonaka S, George D, Iqbal A. Interpreting COPD prevalence estimates: what is the true burden of disease. *Chest* 2003;**123**:1684–92.
17. Viegi G, Pedreschi M, Pistelli F, et al. Prevalence of airways obstruction in a general population: European Respiratory Society vs American Thoracic Society definition. *Chest* 2000;**117**:339S–45S.
18. Celli BR, Halbert RJ, Isonaka S, Schau B. Population impact of different definitions of airway obstruction. *Eur Respir J* 2003;**22**:268–73.
19. Siafakas NM, Vermeire P, Pride NB, et al. Optimal assessment and management of chronic obstructive pulmonary disease (COPD). The European Respiratory Society Task Force. *Eur Respir J* 1995;**8**:1398–420.
20. Viegi G, Pedreschi M, Baldacci S, et al. Prevalence rates of respiratory symptoms and diseases in general population samples of North and Central Italy. *Int J Tuberc Lung Dis* 1999;**3**(11):1034–42.
21. American Thoracic Society. Standardization of Spirometry—1987 Update. *Am Rev Respir Dis* 1987;**136**:1285–98.
22. Cleveland WS. Robust locally weighted regression and smoothing scatterplots. *J Am Stat Assoc* 1979;**74**:829–36.
23. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**1**:307–10.
24. Cerveri I, Zoia MC, Fanfulla F, et al. Reference values of arterial oxygen tension in the middle-aged and elderly. *Am J Respir Crit Care Med* 1995;**152**:934–41.
25. Enright PL, Sherrill DL. Reference equations for the six-minute walk in healthy adults. *Am J Respir Crit Care Med* 1998;**158**(5 Pt 1):1384–7.
26. Lazarus R, Sparrow D, Weiss ST. Effects of obesity and fat distribution on ventilatory function. *Chest* 1997;**111**:891–8.
27. Bottai M, Pistelli F, Di Pede F, et al. Longitudinal changes of body mass index, spirometry and diffusion in a general population. *Eur Respir J* 2002;**20**:665–73.
28. Mushtaq M, Hayton R, Watts T, et al. An audit of pulmonary function laboratories in the West Midlands. *Respir Med* 1995;**89**:263–70.
29. Kunzli N, Ackermann-Liebrich U, Keller R, et al. Variability of FVC and FEV1 due to technician, team, device and subject in an eight centre study: three quality control studies in SAPALDIA. Swiss Study on Air Pollution and Lung Disease in Adults. *Eur Respir J* 1995;**8**:371–6.
30. Paggiaro PL, Viegi G, Filieri M, et al. Confronto tra uno spirometro a pistone e uno pneumotacografo con sistema computerizzato. *Med Lavoro* 1981;**6**:494–504.
31. Viegi G, Simoni M, Pistelli F, et al. Inter-laboratory comparison of flow-volume curve measurements as quality control procedure in the framework of an international epidemiological study (PEACE project). *Respir Med* 2000;**94**:194–203.
32. Scalera AR, Di Pede F, Pistelli F, et al. Variabilità dei parametri di funzione respiratoria misurati con tre diversi strumenti. *Med Toracica* 1997;**19**:13–20.