Reference values of spirometry for Finnish adults

A. Kainu[1], K. L. Timonen[2,3], J. Toikka[4,5], B. Qaiser[6], J. Pitkäniemi[6], J. T. Kotaniemi[7], A. Lindqvist[8], E. Vanninen[3], E. Länsimies[3] and A. R. A. Sovijärvi[9]

[1]Department of Pulmonary Medicine, HUCH Heart and Lung Center, Peijas Hospital, University of Helsinki and Helsinki University Hospital, Helsinki, [2]Department of Clinical Physiology, Central Hospital of Central Finland, Jyväskylä, [3]Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital and University of Eastern Finland, Kuopio, [4]Department of Clinical Physiology, Turku University Hospital, Turku, [5]Department of Clinical Physiology, Tampere University Hospital, Tampere, [6]Department of Public Health, Hjelt –institute, University of Helsinki, Helsinki, [7]Department of Respiratory Diseases, Tampere University Hospital, Tampere, [8]Research Unit of Pulmonary Diseases, University of Helsinki and Helsinki University Hospital, Helsinki and [9]Department of Clinical Physiology and Nuclear Medicine, HUS Medical Imaging Center, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

Summary

Correspondence

Annette Kainu, Department of Pulmonary Medicine, HUCH Heart and Lung Center, Helsinki University Central Hospital, Peijas Hospital P.O. Box 900, 00029 HUS, Finland
E-mail: annette.kainu@helsinki.fi

Background Diagnostic assessment of lung function necessitates up-to-date reference values. The aim of this study was to estimate reference values for spirometry for the Finnish population between 18 and 80 years and to compare them with the existing Finnish, European and the recently published global GLI2012 reference values.

Methods Spirometry was performed for 1380 adults in the population-based FinEsS studies and for 662 healthy non-smoking volunteer adults. Detailed predefined questionnaire screening of diseases and symptoms, and quality control of spirometry yielded a sample of 1000 native Finns (387 men) healthy non-smokers aged 18–83 years. Sex-specific reference values, which are estimated using the GAMLSS method and adjusted for age and height, are provided.

Results The predicted values for lung volumes are larger than those obtained by GLI2012 prediction for the Caucasian subgroup for forced vital capacity (FVC) by an average 6·2% and 5·1% and forced expiratory volume in 1 s (FEV1) by an average 4·2% and 3·0% in men and women, respectively. GLI2012 slightly overestimated the ratio FEV1/FVC with an age-dependent trend. Most reference equations from other European countries, with the exception of the Swiss SAPALDIA study, showed an underestimation of FVC and FEV1 to varying degrees, and a slight overestimation of FEV1/FVC.

Conclusion This study offers up-to-date reference values of spirometry for native Finns with a wide age range. The GLI2012 predictions seem not to be suitable for clinical use for native Finns due to underestimation of lung volumes.

Introduction

Spirometry is the most commonly used method for assessment of lung function. Up-to-date reference values, which reflect the target population, are necessary. In an optimal situation, reference values for diagnostic purposes are estimated from randomly selected healthy subjects without any exposure to inhaled substances known to potentially affect lung function. Lung function depends on ethnicity due to differences in body composition and the proportional size of thoracic cavity in relation to height.

In Finland, the presently used reference values were derived from a selected occupational cohort of 296 males and 257 females with a relatively narrow age range between 18 and 65 years in the late 1970s; in the age group from 60 to 65 years, only 9 men and 13 women were included (Viljanen *et al.*, 1982). Thus, these reference values are not appropriate for elderly subjects and may not reflect a current healthy population due to changes in health-related factors in Finland. From 1970s, the average height of the Finnish population has increased on average almost 5 cm in both men and women, a change reflecting generally improved level of health and nutrition during childhood and adolescence (Silventoinen *et al.*, 1999). In addition, the reference values of Viljanen (Viljanen *et al.*, 1982) were compiled using a rolling-seal spirometer coupled to an analog x-y-writer. This system has not been used for over two decades in clinical work. The performance of modern flow transducer devices with digital data processing

differ in comparison with the system used by Viljanen in terms of internal resistance, friction, time constants and accuracy.

In 2012, Quanjer et al. (2012) presented new Global Lungs Initiative (GLI2012) reference equations using the LMS method. These all-age reference values provide a statistically more valid description of the evolution of lung function from childhood to adulthood (Cole & Green, 1992; Stanojevic et al., 2008; Cole et al., 2009). Different racial groups were included, but the majority of included studies were grouped to 'Caucasian' despite a very large geographical variation. No data from Finnish people were included in that material. Predicted values for lung volumes according to the ECSC are up to 10% smaller than the current Finnish reference values (Viljanen et al., 1982; Quanjer et al., 1993). In Australia and Poland, the GLI2012 values have recently been evaluated in hospital tertiary care patients undergoing routine spirometry. The GLI2012 prediction models were found to produce significantly larger lung volumes than the old ECSC reference values (Quanjer et al., 2013).

In healthy subjects from across Australia and New Zealand, the GLI2012 reference values for Caucasians have been found to reflect current values adequately although some statistically significant differences were found (Hall et al., 2012). The reported z scores of around +0·3 (SD 1·0) reflect slight underestimation of lung volumes with GLI2012. In Tunis, a Northern African country which is also categorized into the same Caucasian ethnic group as Finland, the GLI2012 reference values have recently been found to significantly overestimate lung volumes in healthy non-smokers with z scores in the magnitude of −0·6 (SD 0·9) (Ben Saad et al., 2013).

The aim of this study was to produce for clinical use new reference values of flow-volume spirometry for native Finns from healthy non-smoking adults with a wide age range recorded with modern flow transducers and compiled with the most recent statistical modelling. We also aimed to assess differences between the new reference data, existing Finnish reference values, other European reference values and the new global GLI2012 reference values.

Methods

Study subjects were recruited from four locations (Helsinki, Kuopio, Tampere and Kemi) in Finland representing geographically diverse populations. Thus, the subjects covered people from southern, eastern, middle and northern Finland. All study participants were native Finns. Uniform inclusion and exclusion criteria were applied to include healthy and non-smoking subjects with no morbid obesity, no occupational exposure to vapours, gases, dusts or fumes and accepting only short previous exposure to tobacco smoke. Approvals of the study protocols were obtained from Helsinki University Central Hospital Coordinating Ethics Committee (for Helsinki), Research Ethics Committee for the Hospital District of Northern Savo (for Kuopio and Tampere) and

Länsi-Pohja Central Hospital Ethics Committee (for Kemi). A written informed consent was obtained from all participants.

In Tampere and Kuopio, healthy non-smokers were specifically recruited for this study between 2005 and 2006 using newspaper announcements. To find also sufficient numbers of elderly subjects, announcements were targeted in Kuopio to newspapers read more frequently by senior citizens. Subjects were interviewed to determine eligibility for this study using predefined criteria, which were selected to correspond to the interview and anthropometric criteria applied to the population samples. From the FinEsS population study (Helsinki and Kemi), healthy non-smoking subjects were selected based on the FinEsS interview questionnaire. The study protocol for the FinEsS studies has been published previously (Kotaniemi et al., 2005; Pallasaho et al., 2006; Kainu et al., 2008). Of the 643 participants of the FinEsS-Helsinki population study undertaken between 2001 and 2003, 212 subjects were identified on predefined criteria healthy and non-smoking. In Helsinki, a separate repeatability study was conducted with 21 healthy non-smoking volunteers. Of the FinEsS-Kemi population study undertaken between 1996 and 1999, 695 participants yielded 233 healthy non-smoking subjects using identical screening protocol. The formation of the study sample from the different study centres is shown in Fig. 1.

Subjects were required not to have any diagnosed acute or chronic pulmonary disease, cardiac disease or neurological disability, morbid obesity, no prior chest surgery, radiotherapy or anomalies of the thorax, no systemic diseases known to affect respiratory function (e.g. connective tissue disease or muscular dystrophy), no prior use of lung medication (with the exception of temporary use of cough medication) or heart medication potentially affecting respiratory function. Prescribed use of medication was considered an indication of diagnosed disease and thus included in the predefined exclusion criteria. Reported use of medication for systemic hypertension was the only allowed exception, in otherwise healthy and asymptomatic eligible subjects. Subjects with abdominal surgery during the previous 6 months, pregnancy over 20 weeks of gestation or childbirth <3 months prior to the study visit, were excluded from the study sample. Smoking history of <10 pack years was allowed if the smoking cessation had taken place over 10 years previously in order to accommodate older age cohorts, where prior smoking in adolescence was known to be more prevalent. Height was measured without shoes using a calibrated stadiometer to the closest full cm and weight with only light clothing in kilograms with one decimal. The difference of study date and subject's date of birth was used to calculate age truncated to full years used for calculations.

Identical Vmax 22D spirometers (heated wire flow transducer spirometer) were used in each study centre (Sensor Medics Corporation, Yorba Linda, California, USA). Spirometers were calibrated on every day in the morning with a certified 3-L syringe for volume and flow; three different flow
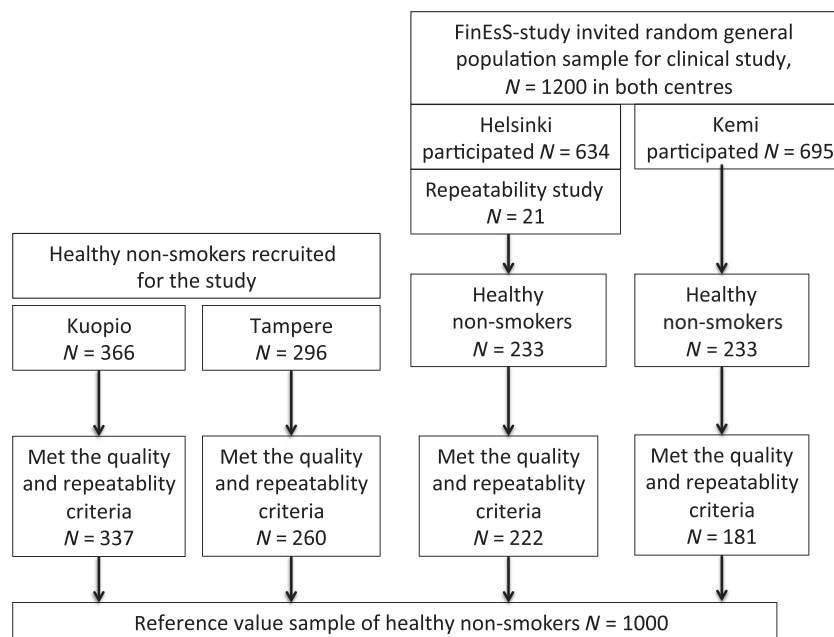
Figure 1 Formation of the study sample in the four study locations.

levels between 2 and 12 l s^{-1} were used. All the measurements were taken originally following the 1994 ATS Standard (American Thoracic Society, 1995). A minimum of three successive acceptable spirograms were recorded. All data included are taken from measurements before inhalation of eventual bronchodilator. During the study, the updated ERS/ATS standards were issued, which slightly changed, for example the repeatability criteria (Miller et al., 2005). In the ATS 1994 standard, both forced vital capacity (FVC) and forced expiratory volume in 1 s (FEV1) were acceptable, if the two best values differed no more than 200 ml (American Thoracic Society, 1995). The ATS/ERS criteria reduced this allowed variation to 150 ml in both volumes (Miller et al., 2005). Of the total of 1127 subjects eligible for the study sample based on questionnaire screening, all spirometries were reviewed for technical acceptability. Subjects not fulfilling the tighter repeatability criteria were excluded from the study sample. The final sample was 1000 subjects, 613 women and 387 men.

Table 1 shows descriptive characteristics of anthropometric measurements of the study sample. The age and height distribution of the study participants is shown in Table 2. A total of 149 subjects (14·9%) were classified as eversmokers having smoked between 0·1 and 9·5 pack years with an average of 3·6 (SD 3·3) pack years and 28·4 (SD 10·3) years from quitting smoking.

Flow-volume spirometry data were extracted electronically and back verified to the printed copies of spirometry reports. The largest values of FVC, forced expiratory volume in 6 s (FEV6) and FEV1 were taken, and their ratios (FEV1/FVC and FEV1/FEV6) were calculated using these largest values. All other flow values were taken from the curve with largest sum of FEV1 + FVC. Data on FEV6 were only available for

measurements taken in Helsinki and Kuopio. The forced expiratory flow-volume variables analysed include FVC, FEV6, FEV1, FEV1/FVC, FEV1/FEV6, maximum mid-expiratory flow (MMEF), maximum expiratory flow at 75%, 50% and 25% of FVC remaining (MEF75, MEF50, MEF25) and peak expiratory flow (PEF).

The measured and estimated predicted values were compared to selected other reference values available from Europe, namely Switzerland (Brändli et al., 1996), Norway (Langhammer et al., 2001), the United Kingdom (Falaschetti et al., 2004), Denmark (Løkke et al., 2013) and Sweden (Hedenström et al., 1985, 1986) in addition to the GLI2012 (Quanjer et al., 2012), Viljanen (Viljanen et al., 1982) and ECSC reference values (Quanjer et al., 1993).

A separate repeatability study was conducted in Helsinki with 21 healthy volunteers, 6 men and 15 women. The participants were on average 40·3 (SD 9·9) years old. Spirometry was repeated on two consecutive days on the same time of the day (±1 h) and undertaken by the same study nurse. Change in best FVC was on average 0·007 (SD 0·187) litres or 0·1% (SD 3·6%) from baseline. Change in best FEV1 was on average 0·013 (SD 0·098) litres or 0·5% (SD 3·9%) from baseline. FEV1/FVC changed in average 0·003 (SD 0·017) or 0·4% (SD 2·1%) relative to baseline. Only data from the spirometry completed on the first study day were included in the reference values study.

Statistical methods

All statistical analyses were performed with the R program (version 2.15.1, http://cran.r-project.org) and using the gamlss package (Rigby & Stasinopoulos, 2005; R Core Team, 2012). We applied the generalized additive model for location,

Table 1 Descriptive statistics of the study population. Values are presented as mean (SD).

	Men (n = 387)		Women (n = 613)	
	Mean (SD)	**Min–max**	**Mean (SD)**	**Min–max**
Age (years)	49·8 (16·3)	19–82	47·9 (16·0)	18–83
Height (m)	1·77 (0·07)	1·53–2·03	1·64 (0·06)	1·40–1·80
BMI (m/kg^2)	25·2 (2·6)	17·8–30·0	24·0 (3·0)	17·7–30·0
FVC (l)	5·15 (0·93)	2·26–7·34	3·65 (0·68)	1·52–6·20
FEV6 (l)	4·92 (1·01)	2·24–7·32	3·53 (0·75)	1·51–6·20
FEV1 (l)	4·01 (0·81)	1·92–6·09	2·91 (0·61)	1·20–5·18
FEV1/FVC (%)	77·7 (6·2)	52·6–95·4	79·7 (6·3)	52·9–99·5
FEV1/FEV6 (%)	80·1 (4·6)	67·1–93·5	81·3 (5·1)	64·2–94·3
PEF (l s^{-1})	10·3 (1·6)	5·3–18·1	7·2 (1·2)	3·5–11·5
MMEF (l s^{-1})	3·7 (1·3)	0·9–8·1	2·9 (1·0)	0·5–6·1
MEF75 (l s^{-1})	8·2 (1·9)	2·9–15·2	5·9 (1·3)	2·4–11·5
MEF50 (l s^{-1})	4·4 (1·4)	1·3–9·1	3·4 (1·1)	0·9–7·4
MEF25 (l s^{-1})	1·5 (0·7)	0·2–4·4	1·2 (0·6)	0·1–3·5

BMI, body mass index; FVC, forced vital capacity; FEVt, forced expiratory volume in t seconds, PEF, peak expiratory flow, MMEF, maximum mid-expiratory flow; MEFx, maximum expiratory flow at x% of FVC remaining.

Table 2 Height and age distribution of study participants.

	Men height category (cm)						Women height category (cm)					
Age (years)	<160	160–169	170–179	180–189	≥190	All	<150	150–159	160–169	170–179	≥180	All
<30	0	2	25	20	1	48	0	15	56	29	0	100
30–39	0	4	27	45	1	77	1	15	57	34	0	107
40–49	0	2	28	31	5	66	0	26	69	28	1	124
50–59	1	7	37	24	2	71	0	32	70	10	0	112
60–69	1	19	40	12	1	73	2	33	62	6	1	104
≥70	2	17	28	5	0	52	4	40	22	0	0	66
All	4	51	185	137	10	387	7	161	336	107	2	613

scale and shape (GAMLSS) used also in the modelling of the Global Lungs Initiative (GLI2012) (Quanjer et al., 2012). The original LMS method summarizes the changing distribution by three curves representing the median μ (M), coefficient of variation σ (S) and skewness, the latter expressed as after a Box–Cox power λ (L) transformation. The LMS method was developed by Cole and originally described by Cole & Green, (1992) (Cole et al., 2009). In our study material, the normal distribution (NO()) resulted in smaller prediction errors and the most parsimonious model in terms of the Schwarz Bayesian Criterion (SBC) than Box–Cox Cole and Green (BCCG) distribution. Adding a moment for skewness (L) did not significantly improve the fit of the model in any of the variables. Age and height were found to be the main determinants in all evaluated lung function parameters. To allow a flexible model for the relationship between age and the different spirometric variables, we used penalized b-splines for age in both the mean and the standard deviation model. With the normal distribution gamlss model (NO()), the moment for coefficient of variation is the standard deviation (Stasinopoulos et al., 2012). Weight and BMI were assessed separately, but neither provided any additional degree of explanation to the models.

We used the following prediction equations for the mean (M) and standard deviation (S) of the lung function variables:

$$M_i = \exp(a_0 + a_1 * \ln(height_i) + a_2 * \ln(age_i) + MSpline_i),$$

$$S_i = \exp(b_0 + b_1 * \ln(age_i) + SSpline_i),$$

where a_0, a_1, a_2 and b_0, b_1 are the regression coefficients estimated from the sample. The values for spline functions $MSpline_i$ and $SSpline_i$, for each predicted value, are presented in the lookup table in Annex S1. The predicted mean value (M_{pred}) of each lung function variable can also be expressed as: $M_{pred} = e^{a_0} * height^{a_1} * age^{a_2} * e^{MSpline}$

The lower limits of normal (LLN) (5th percentile) were obtained from formula:

$$LLN_i = M_i - 1\cdot645 * S_i$$

Z score for each individual measurement (z_i) can be calculated:

$$z_i = \frac{measured_i - M_i}{S_i}$$

Differences between study locations were assessed using a dummy variable for study location and assessing differences

between predictions from different locations. Similarly, a sensitivity analysis was conducted using dichotomous smoking (never and ever) and obesity (BMI > 30) as dummy variables in a full gamlss model. No statistically significant differences were found in the predicted values in terms of smoking or obesity when sex, age and height of the subjects were taken into consideration and when using a dummy variable for study location. Obese subjects were excluded from final models.

Bland–Altman plots for difference between predicted FVC and FEV1 from this study and those predicted from GLI2012 reference values were plotted as described by Bland and Altman (Bland & Altman, 1986).

Results

The estimated coefficients for the regression equations using the LMS method are listed separately for each lung function variable in Table 3. The mean predicted values and their respective LLN for FEV1, FEV6, FVC and FEV1/FVC and FEV1/FEV6 ratios are plotted in Fig. 2 for men and women using arithmetic mean of height of the sample. Bland–Altman plots for difference between predicted FVC and FEV1 from this data set and the Global Lungs Initiative 2012 (GLI2012) reference values in Fig. 3 show that the absolute difference in predicted values is greater in smaller predicted lung volumes and decreases with increasing predicted lung volume.

Model predicted mean values from this study compared to Viljanen reference values (Viljanen *et al.*, 1982) for MEF50 and compared to the GLI2012 reference values (Quanjer *et al.*,

2012) for MMEF with the respective LLN values are shown in Fig. 4 as a function of age for average height men and women. Viljanen reference values markedly overestimate the MEF50 values. GLI2012 values do not have MEF50 model, but the MMEF model is closely similar to the predicted model from our study. The coefficient of variation was larger in instantaneous flow values (MEF25, MEF50, MEF75) and in MMEF compared to main variables (FEV1, FVC and FEV1/FVC).

The difference between measured and predicted values calculated from this data set was in average FEV1 0·064 (SD 0·101) litres, FVC −0·014 (0·546) litres and FEV1/FVC 0·11% (5·24%). With GLI2012 reference values, the differences were 0·099 (0·442) litres, 0·205 (0·557) litres and −1·80% (5·34%), respectively. The difference in predicted mean and LLN values between the present study and GLI2012 reference values in Fig. 5 demonstrates that the GLI2012 reference values slightly underestimate both mean FEV1 and FVC in men and women with an age-dependent increasing trend. However, the GLI2012 reference values give better agreement in lung volumes than the ECSC reference values, which underestimate the FVC and FEV1 by around 0·4 l (Table 4). Particularly in women, the ECSC reference values underestimate FVC in this data set on average by 527 ml or 17%. Viljanen reference values on average performed better than the GLI2012 in both FVC and FEV1 in the applicable age range of 18–65 years. Of other comparable reference values, the prediction equations from Switzerland (Brändli *et al.*, 1996) were closest to the observed distribution, whereas values from

Table 3 Coefficients for predictive equations. Formula for mean (M) = exp(a_0 + a_1*ln(height) + a_2*ln(age) + MSpline) and standard deviation (S) = exp (b_0 + b_1*ln(age) + Sspline). For spline contributions to each model, see respective lookup tables[a].

Spirometry variable		a_o	a_1	a_2	b_o	b_1
FVC (l)	Men	−8·903	2·165	−0·173	−1·652	0·319
	Women	−9·278	2·214	−0·188	−0·219	−0·149
FEV6 (l)	Men	−8·377	2·101	−0·232	−1·352	0·258
	Women	−9·371	2·275	−0·254	0·052	−0·220
FEV1 (l)	Men	−7·437	1·908	−0·275	−1·547	0·239
	Women	−7·205	1·858	−0·316	−0·032	−0·266
FEV1/FVC	Men	1·673	−0·292	−0·108	−2·643	−0·065
	Women	1·869	−0·311	−0·135	−3·076	0·023
FEV1/FEV6	Men	1·128	−0·223	−0·051	−1·494	−0·424
	Women	1·556	−0·274	−0·096	−2·809	−0·091
PEF (l s^{-1})	Men	−3·402	1·172	−0·086	−0·596	0·248
	Women	−3·733	1·211	−0·123	0·462	−0·104
MMEF (l s^{-1})	Men	−1·741	0·984	−0·539	0·545	−0·139
	Women	−0·240	0·701	−0·607	0·804	−0·308
MEF75 (l s^{-1})	Men	−1·810	0·843	−0·117	0·233	0·090
	Women	−2·368	0·932	−0·158	0·343	−0·054
MEF50 (l s^{-1})	Men	−1·867	0·940	−0·397	0·350	−0·043
	Women	−1·495	0·852	−0·427	0·267	−0·113
MEF25 (l s^{-1})	Men	−4·195	1·512	−0·859	0·463	−0·303
	Women	0·123	0·761	−1·031	1·961	−0·779

[a]Height in cm, age in years.
FVC, forced vital capacity; FEVt, forced expiratory volume in t seconds; PEF, peak expiratory flow; MMEF, maximum mid-expiratory flow; MEFx, maximum expiratory flow at x% of FVC remaining.
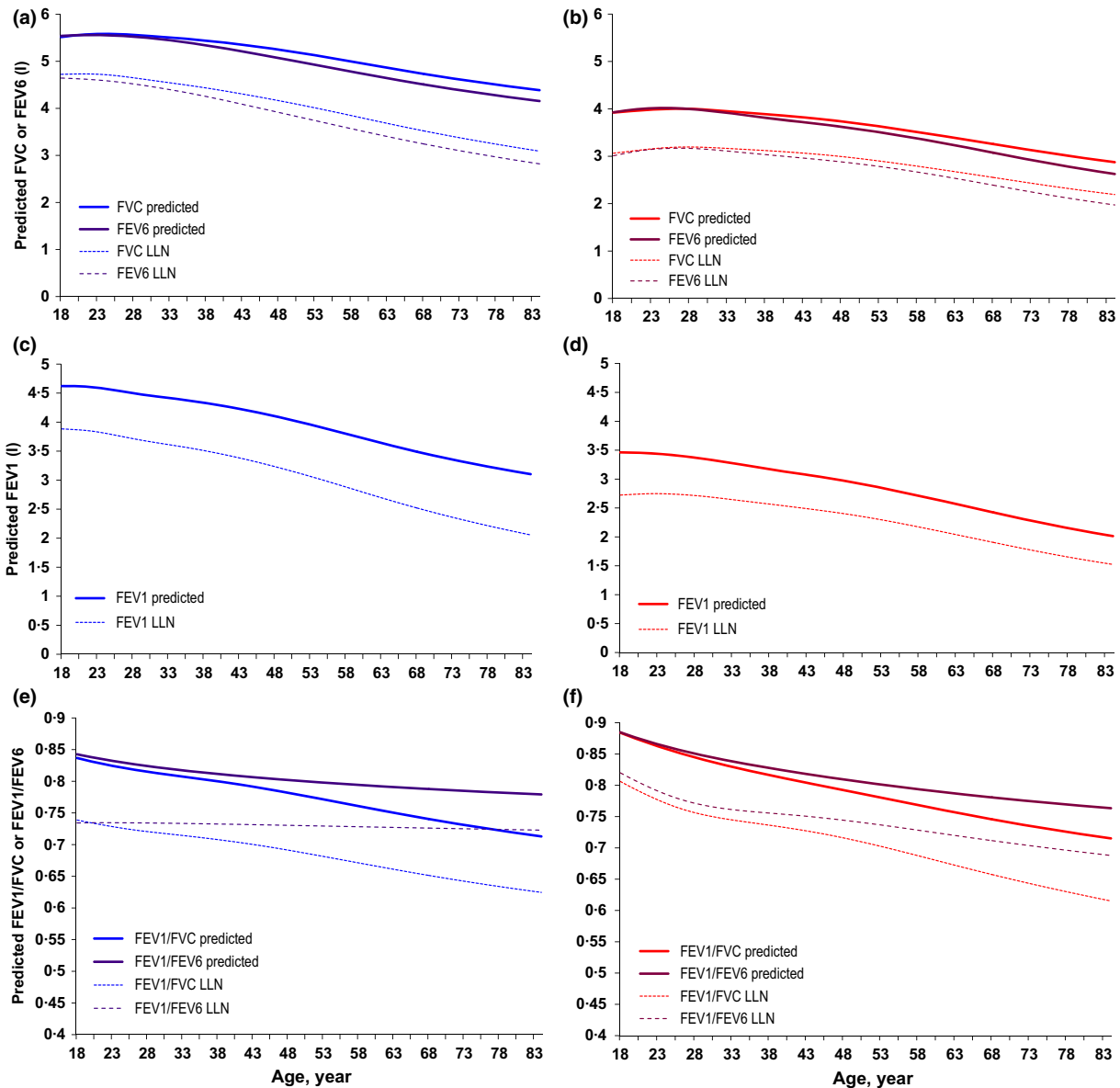
Figure 2 Predicted mean and lower limit of normal (LLN) for (a and b) forced vital capacity (FVC) and forced expiratory volume in 6 s (FEV6), (c and d) forced expiratory volume in 1 s (FEV1), and (e and f) FEV1/FVC and FEV1/FEV6. Graphs in a, c and e present values for average-age height men (177 cm) and in b, d and f for average height women (164 cm), respectively.

Nordic countries such as Denmark (Løkke et al., 2013) and Norway (Langhammer et al., 2001) underestimated FVC and overestimated FEV1/FVC. But the old Swedish values of FVC (Hedenström et al., 1985, 1986) were very close to the new Finnish values in their age range (Table 4).

The mean z score of FVC in men was 0·37 (SD 1·00) when calculated according the GLI2012 predictions and −0·03 (SD 1·00) according the predictions of the present study. For FEV1/FVC, the values were −0·23 (0·80) and 0·02 (1·00), respectively. The predicted model from the present study compared to the GLI2012 reference values (Quanjer et al., 2012) for FVC, FEV1 and FEV1/FVC in relation to different age categories are presented in Fig. 6. Although the mean difference between measured and predicted GLI2012 reference

value was small, the figure shows a slight trend of increasing underestimation of FEV1 by age when using GLI2012 predictions. Data calculated according to our new predicted model compared to those by GLI2012 reference model in different height categories are shown in Fig. 7; the underestimation of lung volumes is greater in shorter subjects and decreases with increasing height.

Discussion

The need for locally representative reference values is of utmost importance. The old Finnish reference values by Viljanen (Viljanen et al., 1982) are limited by the relatively narrow age scale of study subjects, potential cohort effect,
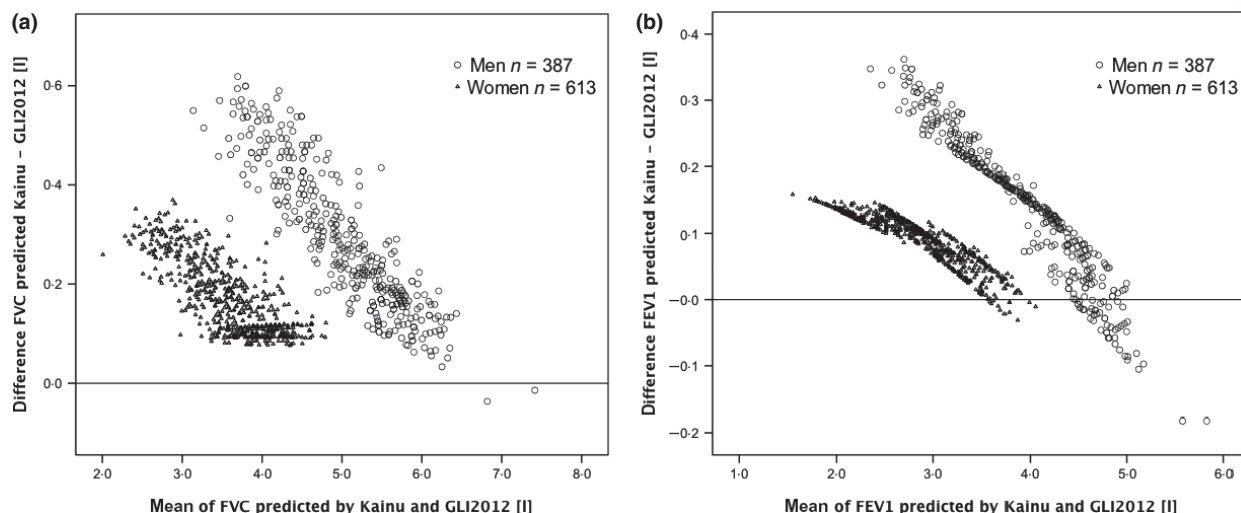
Figure 3 Bland–Altman plots of individual differences between predicted values according to present study (Kainu *et al.*) and predicted values according to GLI2012 (Quanjer *et al.*, 2012) for forced vital capacity (FVC) (a) and forced expiratory volume in 1 s (FEV1) (b).
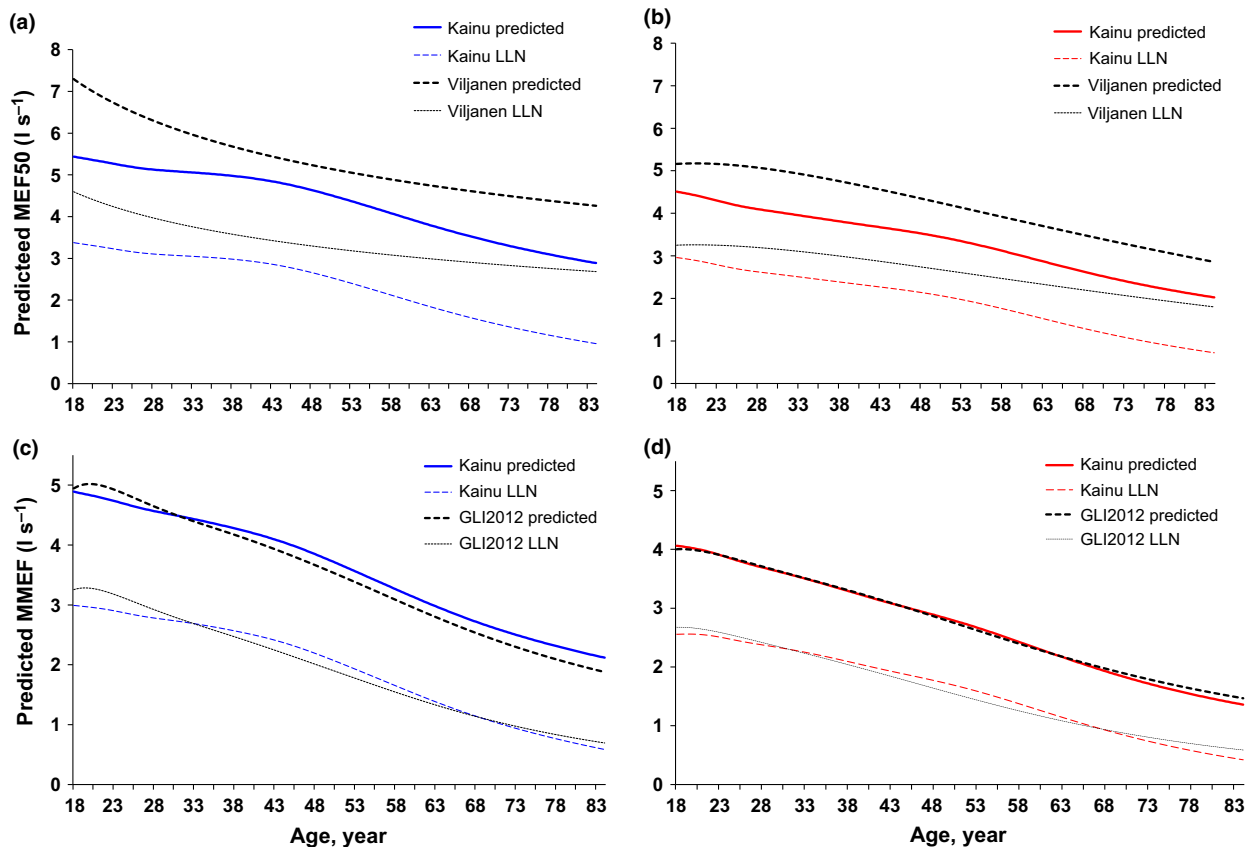

Figure 4 Predicted mean and lower limit of normal (LLN) of (a and b) maximum expiratory flow at 50% forced vital capacity remaining (MEF50) from this study and from Viljanen reference model (Viljanen *et al.*, 1982), and (c and d) maximum mid-expiratory flow (MMEF) from this study and from GLI2012 model (Quanjer *et al.*, 2012). Graphs a and c present values for average height men (177 cm), and b and d for average height women (164 cm), respectively.

substantial change in the technical equipment and improvement in statistical methodologies over the past 40 years. In this study, we evaluated 1000 healthy non-smoking native Finnish subjects for lung function and found that the recently published GLI2012 reference values on average provided with reasonably good fit, but underestimated lung volumes, especially FVC, with the degree of underestimation gradually increasing with advancing age.
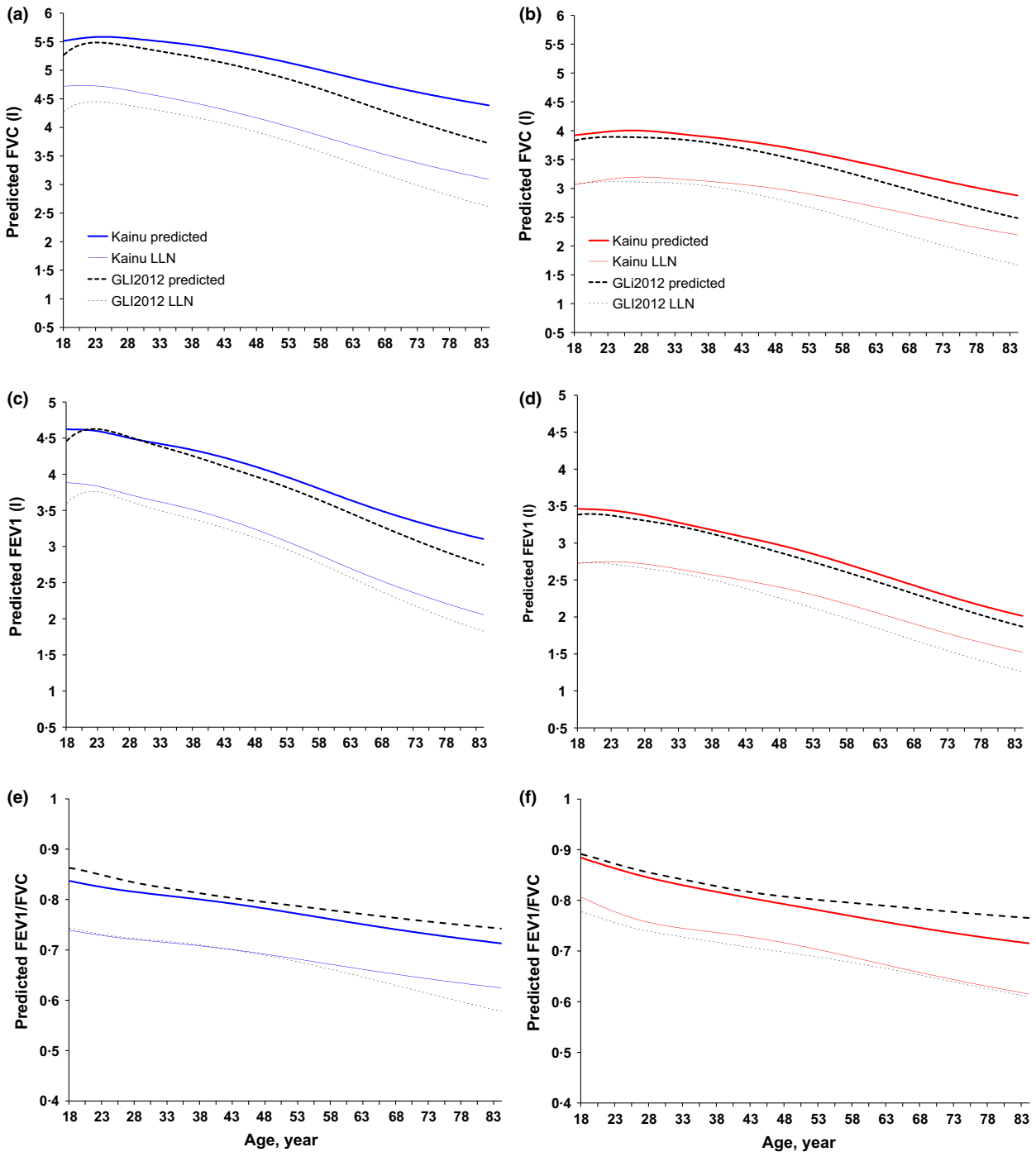
Figure 5 Comparison of predicted mean and lower limit of normal (LLN) of (a and b) forced vital capacity (FVC), (b and c) forced expiratory volume in 1 s (FEV1), and (d and e) FEV1/FVC between GLI2012 (Quanjer *et al.*, 2012) and predicted values from this study. Graphs a, c and e present values for average height men (177 cm), and graphs b, d and f for average height women (164 cm).

GLI2012 reference values are clearly more representative of current target population than the ECSC reference values in lung volumes, but the FEV1/FVC ratio was better predicted using the ECSC due to the fact that the magnitude of underestimation of FEV1 and FVC was similar in the ECSC predictions. Of other European reference values evaluated, the Swiss reference values of Brändli *et al.* showed best agreement both in terms of lung volumes and in the FEV1/FVC ratio. However, significant differences were found between the predicted GLI2012 values and new locally derived predicted values using a similar model. The Finnish adults had slightly larger lung volumes than predicted by the GLI2012 reference values and also other comparable reference values, with the exception of the Brändli *et al.* values. The found average underestimation of

Table 4 Mean predicted values and mean differences between observed and predicted values in the present study and predicted in some other European publications. Reference values are applied only to the indicated age ranges of study subjects.

		Men			Women		
	Age group	Subjects *n*	% Predicted mean (SD%)	Mean difference	Subjects *n*	% Predicted mean (SD%)	Mean difference
FVC (l)							
Kainu	18–83 years	387	99·7 (13·7)	−0·017	613	99·7 (12·7)	−0·011
GLI2012 (Quanjer et al., 2012)	3–95 years	387	106·2 (15·6)	0·271	613	105·1 (13·9)	0·163
Viljanen (Viljanen et al., 1982)	18–65 years	300	101·4 (12·1)	0·074	508	100·6 (13·0)	0·011
ECSC (Quanjer et al., 1993)	18–69 years	335	106·1 (13·5)	0·306	547	117·4 (16·6)	0·527
Brändli (Brändli et al. 1996)	18–60 years	267	100·6 (11·3)	0·032	456	100·0 (11·8)	0·002
Falaschetti (Falaschetti et al. 2004)	16–75 years	371	106·4 (15·4)	0·284	592	107·5 (14·7)	0·234
Hedenström (Hedenström et al. 1985, 1986)	20–70 years	337	100·0 (13·1)	0·035	548	98·3 (13·2)	−0·074
Langhammer (Langhammer et al. 2001)	20–80 years	384	104·9 (15·7)	0·202	602	103·8 (14·0)	0·113
Løkke (Løkke et al. 2013)	20–89 years	384	109·0 (16·2)	0·401	602	105·1 (13·8)	0·165
FEV1 (l)							
Kainu	18–83 years	387	100·0 (14·7)	−0·002	613	99·6 (12·4)	−0·012
GLI2012 (Quanjer et al., 2012)	3–95 years	387	104·2 (16·4)	0·135	613	103·0 (13·1)	0·077
Viljanen (Viljanen et al., 1982)	18–65 years	300	97·7 (12·1)	−0·098	508	97·3 (12·1)	−0·096
ECSC (Quanjer et al., 1993)	18–69 years	335	108·7 (14·6)	0·320	547	109·6 (14·5)	0·244
Brändli (Brändli et al. 1996)	18–60 years	267	101·5 (12·1)	0·061	456	101·1 (11·5)	0·033
Falaschetti (Falaschetti et al. 2004)	16–75 years	371	105·0 (16·1)	0·162	592	104·7 (13·6)	0·115
Hedenström (Hedenström et al. 1985, 1986)	20–70 years	337	101·1 (13·4)	0·063	548	97·3 (12·4)	−0·092
Langhammer (Langhammer et al. 2001)	20–80 years	384	100·2 (15·3)	−0·008	602	100·1 (12·7)	−0·004
Løkke (Løkke et al. 2013)	20–89 years	384	104·6 (16·4)	0·166	602	103·5 (13·1)	0·094
FEV1/FVC (%)							
Kainu	18–83 years	387	100·1 (7·2)	0·1%	613	100·2 (6·4)	0·1%
GLI2012 (Quanjer et al., 2012)	3–95 years	387	97·9 (7·0)	−1·7%	613	97·7 (6·4)	−1·9%
Viljanen (Viljanen et al. 1982)	18–65 years	300	96·5 (7·0)	−2·9%	508	97·0 (6·0)	−2·5%
ECSC (Quanjer et al., 1993)	18–69 years	335	99·3 (7·0)	−0·6%	547	99·7 (6·4)	−0·3%
Brändli (Brändli et al. 1996)	18–60 years	267	100·9 (7·2)	0·6%	456	100·8 (6·0)	0·7%
Falaschetti (Falaschetti et al. 2004)	16–75 years	371	97·1 (6·9)	−2·3%	592	97·7 (6·3)	−1·9%
Hedenström (Hedenström et al., 1985, 1986)	20–70 years	337	103·0 (7·2)	2·3%	548	100·9 (6·4)	0·6%
Langhammer (Langhammer et al. 2001)	20–80 years	384	92·0 (7·6)	−6·8%	602	97·2 (6·9)	−2·3%
Løkke (Løkke et al. 2013)	20–89 years	384	96·4 (7·0)	−2·9%	602	98·2 (6·6)	−1·4%

SD, standard deviation; FVC, forced vital capacity; ECSC, European Community for Steel and Coal; FEV1, forced expiratory volume in 1 s.

FEV1 by 3% with GLI2012 can be regarded as significant given the role of FEV1 in interpretation of spirometry in general. FEV1 is the main variable used for clinical decision-making and also grading of ventilatory defects in spirometry according to the ATS/ERS standard (Pellegrino et al., 2005). Even more importantly, when the two prediction models are compared on actual reference subjects, the GLI2012 reference model shows an age-dependent trend of increasing underestimation of FEV1 and particularly FVC especially in men. In FEV1/FVC this results in an overestimation that increases with age, which is quite unfortunate given the important role of FEV1/FVC reference value in the diagnosis of airflow limitation. The SAPALDIA study is one of the studies included in the GLI models, which showed one of the best concordances, but which is also limited to the age range 18–60 years. The number of Nordic subjects was very small in the GLI data set.

The observed difference between native Finns and the GLI2012 reference values poses several questions and avenues for further study. The GLI2012 model provided better agreement with younger and taller adults, but predicted smaller lung volumes with an age-dependent trend. The older adults, women in particular, were significantly shorter and also their height varied within the age category less, that is the older subjects were more homogenous with respect to size.

It has been proposed that the ratio of sitting height to standing height, the so-called Cormic index, would be larger in older age groups. Cormic index and sitting height correlate better with lung volumes than standing height. However, sitting height or the Cormic index is not commonly measured in epidemiological studies, not in our study either. The observed increase in lung volumes in relation to height with age could be explained by an increasing Cormic index with age. The maximum attainable height is influenced both from genetic and environmental effects, with changing environmental factors affecting heritability (Silventoinen et al., 2000). The genetic build of the native Finns cannot have changed markedly over decades, but there is a major change in the level of nutrition and health, which can explain why the GLI2012 values would seem to fit better in the younger age categories.
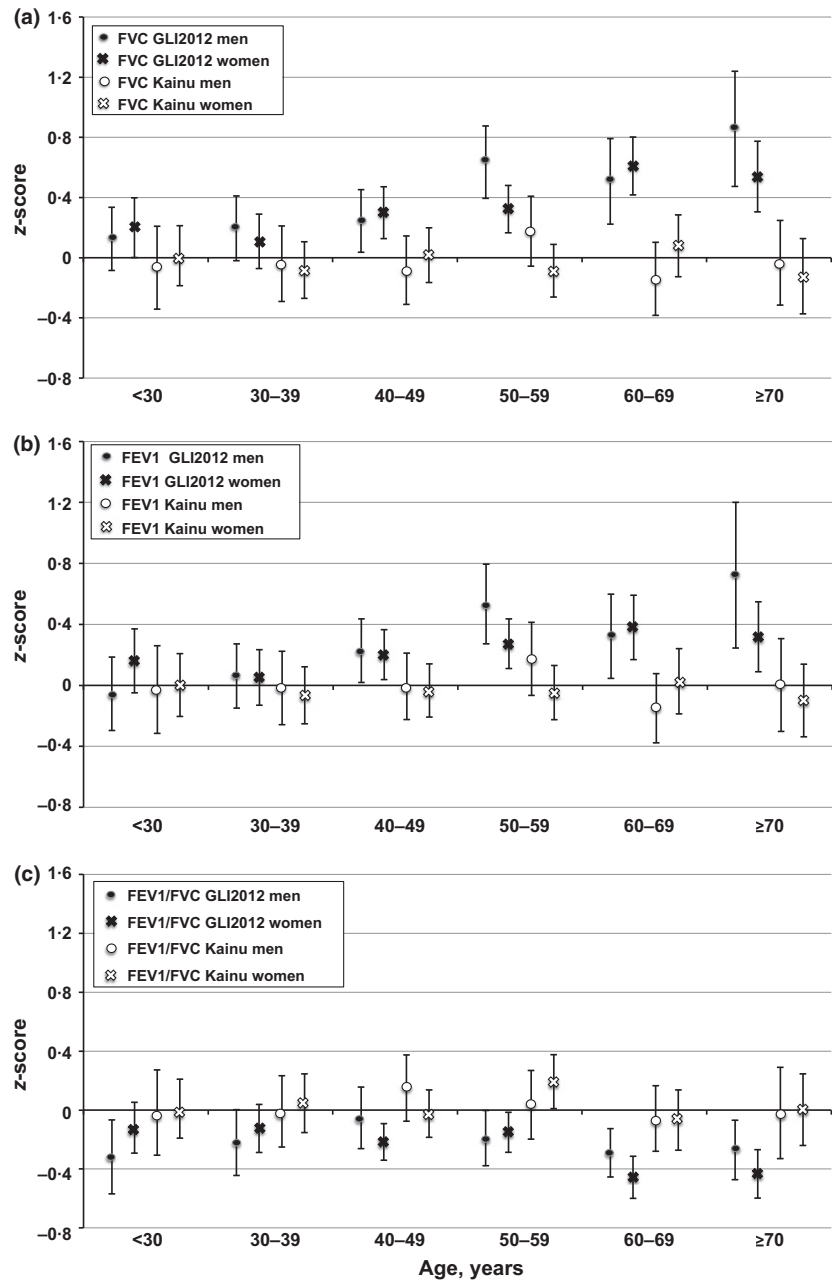
Figure 6 Comparison of z scores of predicted a) forced vital capacity (FVC), b) forced expiratory volume in 1 s (FEV1) and c) their ratio FEV1/FVC predicted from the present study and the GLI2012 (Quanjer *et al.*, 2012) reference models in different age categories for men and women in the reference data set.

This also suggests a significant secular trend in Finland, which can explain the continuing need for national reference values.

In average height men and women, the predicted mean FEV1/FVC ratio from this study was slightly lower than from the GLI2012 model, but the LLN was fairly similar except in the older age categories in men. However, the effect of height is not uniform in these models, which explains the observed difference in actual subjects and the models for average height adults. The mean height of the youngest age groups was approximately 9 cm taller in both men and women compared to the oldest age groups. The effect of height on the predicted FEV1/FVC was larger in our models than in the GLI2012 models. The FEV1/FVC ratio is pivotal in the diagnosis of

obstruction, but FEV1 is highly significant in the grading of reductions in lung function both in obstructive and restrictive disorders – also in international guidelines for diagnosis and management of major obstructive airways diseases particularly chronic obstructive pulmonary disease (Pellegrino *et al.*, 2005; Bousquet *et al.*, 2007; Vestbo *et al.*, 2012).

The proportion of male subjects in the study was relatively small (39%). This was partially due to relatively more frequent disqualification due to reported smoking exposure and respiratory symptoms. Also, younger men were less likely to participate in the study (Kotaniemi *et al.*, 2001). The absolute numbers and age distribution of the male subjects was judged to be adequate for the current analysis. In addition, to exclude
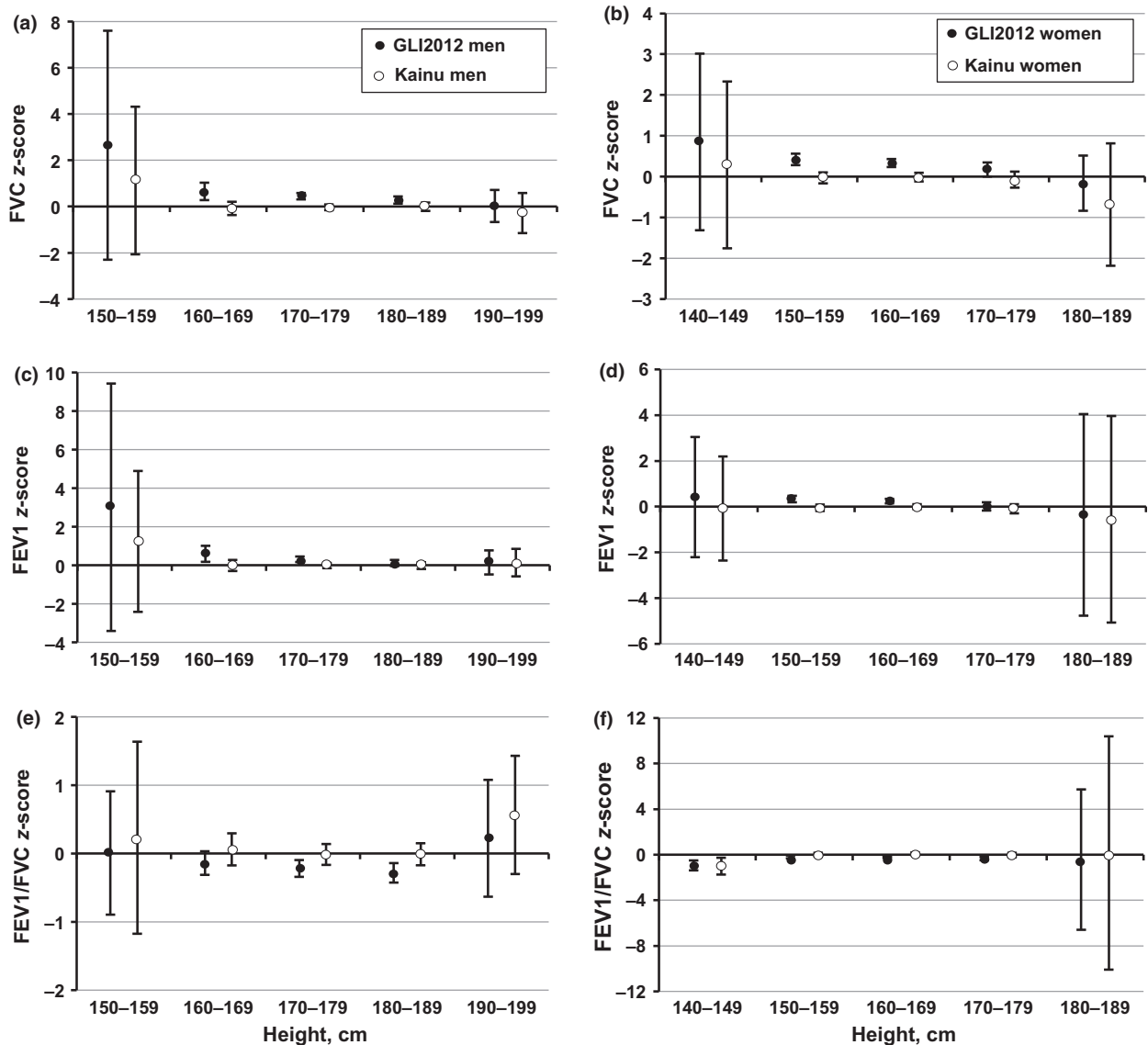
Figure 7 Comparison of z scores of predicted (a and b) forced vital capacity (FVC), (c and d) forced expiratory volume in 1 s (FEV1), and (e and f) their ratio FEV1/FVC predicted from the present study and the GLI2012 (Quanjer *et al.*, 2012) reference models in different height categories. Graphs a, c and e for men, and b, d and f for women in the reference data set.

the potential selection bias introduced by selection criteria, each spirometry variable was separately assessed for differences between study locations using a dummy variable for study location in the GAMLSS model. No statistically significant differences were found when gender, age and height of the subjects were taken into consideration. As the study centre in Kuopio recruited more older subjects than the other centres, all subjects over 60 years of age were excluded in a further effort to rule out selection bias affecting the found age-related increase in model discordance with the GLI2012 prediction model. The sampling differences between the study centres did not seem to significantly affect the prediction models.

To get more elderly subjects for the study, the newspaper announcements to recruit volunteers were targeted in Kuopio to papers frequently read by senior citizens. In the older age cohorts, it was recognized that short smoking exposure during, for example, war years by older men could make the recruitment of healthy older subjects unnecessarily difficult. Thus, it was decided *a priori* that <10 pack years of smoking was allowed, if smoking cessation took place at least 10 years previously. The smoking criteria were implemented stringently, thus also excluding a larger number of young men that had more often casual or short-term smoking, but with smoking cessation invariably <10 years previously.

The effect of including some eversmokers to the model was assessed in a separate sensitivity analysis which showed that the results obtained from eversmokers (n = 149) in this study did not differ from neversmokers with the exception of eversmokers being significantly older, given the required 10-year

interval from smoking cessation. The majority of eversmokers had smoked in their teenage years. The mean age of quitting smoking was 19 years with an average time 28·4 years from quitting smoking. This was the result of predefined criteria and not actually anticipated, but the exclusion of these subjects would have resulted in an additional selection bias.

Careful sensitivity analyses were conducted to rule out selection bias on grounds of smoking exposure, obesity and the different selection procedures in the four locations. The selection criteria from the population studies were predefined and the subjects from Kuopio and Tampere recruited using similar criteria. During the study, new simplified criteria were suggested in the literature for selection of subjects for lung function reference values (Johannessen *et al.*, 2007). As our criteria had been predefined for subject selection in the targeted recruitment, these new criteria could not be applied in this study. However, we find these criteria fairly similar.

Technical differences between the rolling-seal spirometer used by Viljanen and the pneumotachograph sensors used in our measurements and small number of subjects over 60 years of age in the Viljanen study offer an explanation to the observed overestimation of MEF50 in the old Finnish reference values (Viljanen *et al.*, 1982). For MMEF, the new GLI2012 values were very similar to our new values despite larger inherent variability of the instantaneous flow values. The instantaneous flow values provide additional information to complement main indexes and if used, should always be derived from the same data set and manoeuvres as the main variables.

GLI2012 reference values present one set of reference values for all Caucasians broadly defined as all European decent origins. Significant genetic differences exist between different European populations and from studies among twins for FEV1 additive genetic effects of 61% have been observed (Jakkula *et al.*, 2008; Ingebrigtsen *et al.*, 2011). Comparative studies on healthy subjects have not been undertaken, but it has been hypothesized that the northern populations would have slightly larger lung volumes. In Tunis, which is categorized into the same 'Caucasian' ethnic group with Finland, the GLI2012 values have been shown to overestimate lung volumes, with mean z scores of −0·62 (SD 0·86) for FVC and −0·55 (SD 0·87) for FEV1 (Ben Saad *et al.*, 2013). In our study, the same GLI2012 prediction model underestimated

FVC with mean z score of +0·37 (SD 1·00) and FEV1 + 0·25 (SD 1·04). In our study, the difference with the GLI2012 prediction was largest in the older age groups and in subjects with short stature. The new Finnish reference values present the current level of lung function for Finnish adults 18–80 years of age.

Conclusions

The present new reference values offer a closer prediction of spirometric values for native Finns than any other published predictions, including the GLI2012 reference values, which underestimated lung volumes with an age-dependent trend. Therefore, we recommend these new Finnish reference values of flow-volume spirometry for clinical use among native Finns.

Acknowledgments

Conflict of interest

The authors have no conflicts of interest pertaining to this manuscript.

References

American Thoracic Society. Standardization of spirometry. 1994 Update. *Am J Respir Crit Care Med* (1995); **152**: 1107–1136.

Ben Saad H, El Attar MN, Mabrouk KH, *et al.* The recent multi-ethnic global lung initiative 2012 (GLI2012) reference values don't reflect contemporary adult's North African spirometry. *Resp Med* (2013); **107**: 2000–2008.

Bland JM, Altman DG. Statistical methods for assessing agreement between two methods. *Lancet* (1986); **1**: 307–310.

Bousquet J, Clark TJ, Hurd S, *et al.* GINA guidelines on asthma and beyond. *Allergy* (2007); **62**: 102–112.

Brändli O, Schindler Ch, Künzli N, *et al.* Lung function in healthy never smoking adults: reference values and lower limits of normal of a Swiss population. *Thorax* (1996); **51**: 277–283.

Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. *Stat in Med* (1992); **11**: 1305–1319.

Cole TJ, Stanojevic S, Stocks J, *et al.* Age- and size-related reference ranges: a case study of spirometry through childhood and adulthood. *Statist Med* (2009); **28**: 880–898.

Falaschetti E, Laiho J, Primatesta P, *et al.* Prediction equations for normal and low lung function from the Health Survey for England. *Eur Respir J* (2004); **23**: 456–463.

Hall GL, Thompson BR, Stanojevic S, *et al.* The Global Lung Initiative 2012 reference

values reflect contemporary Australasian spirometry. *Respirology* (2012); **17**: 1150–1151.

Hedenström H, Malmberg P, Agarwal K. Reference values for lung function tests in females. Regression equations with smoking variables. *Bulletin Europeen de Physiopathologie Respiratoire* (1985); **21**: 551–557.

Hedenström H, Malmberg P, Fridriksson HV. Reference values for lung function tests in men: regression equations with smoking variables. *Upsala J Med Sci* (1986); **91**: 299–310.

Ingebrigtsen TS, Thomsen SF, var der Sluis S, *et al.* Genetic influences on pulmonary function: a large sample twin study. *Lung* (2011); **189**: 323–330.

Jakkula E, Rehnström K, Varilo T, *et al.* The genome-wide patterns of variation expose significant substructure in a founder population. *Am J Human Gen* (2008); **83**: 787–794.

Johannessen A, Omenaas ER, Eide GE, *et al.* Feasible and simple exclusion criteria for pulmonary reference populations. *Thorax* (2007); **62**: 792–798.

Kainu A, Lindqvist A, Sarna S, *et al.* FEV1 response to bronchodilation in an adult urban population and in healthy adults. *Chest* (2008); **134**: 387–393.

Kotaniemi JT, Hassi J, Kataja M, *et al.* Does non-responder bias have a significant effect on the results in a postal questionnaire study? *Eur J Epidemiol* (2001); **17**: 809–817.

Kotaniemi J-T, Sovijärvi A, Lundbäck B. Chronic obstructive pulmonary disease in Finland: prevalence and risk factors. *COPD* (2005); **3**: 331–339.

Langhammer A, Johnsen R, Gulsvik A, *et al.* Forced spirometry reference values for Norwegian adults: the Bronchial Obstruction in Nord-Trøndelag study. *Eur Respir J* (2001); **18**: 770–779.

Løkke A, Marott JL, Mortensen J, *et al.* New Danish reference values for spirometry. *Clin Respir J* (2013); **7**: 153–167.

Miller MR, Hankinson J, Brusasco V, *et al.* ATS/ERS Task Force: standardisation of lung function testing. Standardisation of spirometry. *Eur Respir J* (2005); **26**: 319–338.

Pallasaho P, Rönmark E, Haahtela T, *et al.* Degree and clinical relevance of sensitization to common allergens among adults: a population study in Helsinki, Finland. *Clin Exp Allergy* (2006); **36**: 503–509.

Pellegrino R, Viegi G, Brusasco V, *et al.* ATS/ERS Task Force: standardisation of lung function testing. Interpretative strategies for lung function tests. *Eur Respir J* (2005); **26**: 948–968.

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.

Quanjer PH, Stanojevic S, Cole TJ, *et al.* ERS Task Force. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J* (2012); **40**: 1324–1343.

Quanjer PH, Brazzale DJ, Boros PW, *et al.* Implications of adopting the Global Lungs Initiative 2012 all-age reference equations for spirometry. *Eur Respir J* (2013); **42**: 1046–1054.

R Core Team (2012). 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/.

Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. *Appl Statist* (2005); **54**: 507–554.

Silventoinen K, Lahelma E, Rahkonen O. Social background, adult body-weight and health. *Int J of Epid* (1999); **28**: 911–918.

Silventoinen K, Kaprio J, Lahelma E, *et al.* Relative effect of genetic and environmental factors on body height: differences across birth cohorts among Finnish men and women. *Am J Public Health* (2000); **90**: 627–630.

Stanojevic S, Wade A, Stocks J, *et al.* Reference ranges for spirometry across all ages. A new approach. *Am J Respir Crit Care Med* (2008); **177**: 253–260.

Stasinopoulos M, Rigby B, Akantziliotou C. Instructions on how to use the gamlss package in R. Second edition. 2012. URL http://www.gamlss.org/wp-content/uploads/2013/01/gamlss-manual.pdf

Vestbo J, Hurd SS, Rodriquez-Roisin R. The 2011 revision of the global strategy for the diagnosis, management and prevention of COPD (GOLD) – why and what? *Clin Resp J* (2012); **6**: 208–214.

Viljanen AA, Halttunen PK, Kreus K-E, *et al.* Spirometric studies in non-smoking, healthy adults. *Scand J Clin Lab Invest* (1982); **42**(Suppl. 159): 5–20.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Coefficients for predictive equations. Formula for mean $(M) = \exp(a_0 + a_1*\ln(\text{height}) + a_2*\ln(\text{age}) + \text{MSpline})$ and standard deviation $(S) = \exp(b_0 + b_1*\ln(\text{age}) + \text{Sspline})$. For spline contributions to each model, see respective lookup tables.