Differences in Pulmonary Function Based on Height Prediction Obtained by Using Alternative Measures

José N. Sancho-Chust Eusebi Chiner Ana Camarasa Cristina Senent

Pneumology Section, Hospital Universitari Sant Joan d'Alacant, Alicante, Spain

Key Words

Anthropometry · Prediction equations · Pulmonary function · Spirometry

Abstract

Background: Standing height is an independent variable used to predict pulmonary function; however, some patients are not able to stand. **Objectives:** The objectives of this study were to compare predicted pulmonary function values obtained by using various alternative measures to estimate height, analyse the values' reproducibility and evaluate their agreement. **Methods:** Standing height, knee height, ulna length, demi-span and arm span were measured in 100 subjects who were able to stand. Five groups of values were generated for predicted FVC and FEV₁ based on measured standing height and the height estimated using each alternative measure. **Results:** The differences found between the height estimated using the different measures and measured standing height were statistically significant. The reproducibility was excellent; however, agreement was poor and all the measures tended to overestimate the predicted values based on standing height although this tendency was less marked in the case of knee height. **Conclusions:** The use of alternative measures to predict height introduces a

certain degree of error in predicted pulmonary function and this error should be quantified. Knee height is the measure that shows the greatest agreement and, thus, could be used in patients who are unable to stand.

Copyright © 2009 S. Karger AG, Basel

Introduction

In spite of the fact that, in the last years, some techniques such as imaging [1] have been developed to improve our knowledge of respiratory physiology, pulmonary function testing [2], including exercise testing [3, 4], is necessary to evaluate the degree of respiratory incapacity. The interpretation of pulmonary function tests is based on the comparison of data measured in an individual and predicted values derived from data obtained in healthy subjects. Most prediction equations use age, sex and standing height (SH) as independent variables [5–7].

However, in many patients referred to pulmonary function laboratories, SH cannot be accurately measured due to the patients' advanced age, extreme debility or the presence of neuromuscular disease. Other patients may have axial skeletal deformities that make it impossible to

take accurate measurements. In such cases, SH must be estimated by alternative methods using anthropometric measurements that do not require the patient to stand.

Knee height (KH) may be easily and accurately measured [8]. It is increasingly used in nutrition [9] and the World Health Organization has recommended its use when SH cannot be obtained [10]. Ulna length (UL) is not affected by age; it may be measured more accurately than other anthropometric measures and is highly reproducible [11]. In addition, it has been studied in the field of pulmonary function testing [12]. Demi-span (DS) is measured on an accessible limb and correlates well with height [13]. Arm span (AS) is the measure most widely studied in the field of pulmonary function testing and may be used to substitute for height [14] directly or after adjustment by applying a fixed correction factor [15]. Prediction equations for pulmonary function based on AS have even been proposed [16]. Consequently, the European Respiratory Society (ERS) and American Thoracic Society (ATS) recommend using AS when SH cannot be measured [2, 17].

However, only few studies have been done comparing the different measures in the same population. Reproducibility was shown to be high but agreement was poor [18]. Even fewer studies investigated the influence of these alternative measures on pulmonary function and there are no data available comparing predicted pulmonary function values obtained by using the different measures to estimate height.

Our hypothesis is that, if the estimated height is close to SH, it may be used to substitute for the latter in pulmonary function prediction equations in patients who cannot stand.

Subjects and Methods

Study Design

This is a cross-sectional study aimed to compare predicted pulmonary function values derived from alternative measures with those obtained by using SH, to analyse the values' reproducibility and to evaluate their agreement.

The study was approved by the Good Clinical Practice and Ethics Committee of the University Hospital of Sant Joan d'Alacant.

To determine the sample size, an SD of 22 cm in the sample was previously estimated, based on a Caucasian healthy population from our laboratory, ranging between 18 and 80 years. The maximum acceptable difference between SH and height predicted by using each alternative measure was assumed at 10 cm. We based our sample size calculation on the 95% CI as the level of certainty that the sample mean does not differ from the true population mean by more than the maximum acceptable difference.

Thus, keeping the type I error rate at 0.05 and the type II error rate at 0.90 to detect a reasonable departure from the null hypothesis, a total number of 102 individuals was the minimum considered necessary to be included in our study.

Methods

Consecutive Caucasian subjects of both sexes and all ages that were referred to the pulmonary function laboratory were included in the study. Written informed consent was obtained from all subjects; in the case of minors, consent was obtained from their parents or legal representatives.

Excluded were subjects in whom SH could not be satisfactorily measured as a result of advanced age, extreme debility or neuromuscular disease. Subjects with axial deformities making measurement of SH inexact and those with deformed or amputated limbs were also excluded. The paediatric population was not systematically excluded since reference equations for the estimation methods were also used.

During the study, all the measurements were made in the Pulmonary Function Laboratory of the University Hospital of Sant Joan d'Alacant by the same observer, using the same spirometer and anthropometric apparatus.

The subjects were weighed, in light clothing and stockinged feet, on a mechanical column scale with sliding weights (model 711; SECA GmbH & Co., Hamburg, Germany). The unit was kilogram to the nearest 100 g. BMI ($\text{weight}/\text{SH}^2$; expressed in $\text{kg}\cdot\text{m}^{-2}$) and body surface area ($0.20247 \times \text{height}^{0.725} \times \text{weight}^{0.425}$; expressed in m^2) were calculated.

SH was measured using a stadiometer adapted to the scales (model 211; SECA GmbH & Co.). The subjects stood in their stockinged feet with their head in the Frankfurt horizontal plane, legs straight and together, and feet touching at the ankles. Measurement was made to the nearest millimetre, then repeated, and the mean of the 2 measurements calculated.

All the alternative measures were obtained in centimetres to the nearest 0.1 cm using a metal tape measure (Lufkin L516CME; Cooper Tools, Apex, N.C., USA).

KH was measured with the subject in the sitting position, legs hanging over the edge of the chair, and knees and hips bent at 90°. The tape measure was placed along the outside of the leg, parallel to the major axis of the tibia. The distance between the lower edge of the heel and the upper part of the knee (just above the kneecap) was measured (fig. 1a). The mean of 2 measurements was used in the following height estimation equations [8]:

$$\begin{aligned}\text{Females: height (cm)} &= \\ &84.88 - [0.24 \times \text{age (years)}] + [1.83 \times \text{KH (cm)}] \\ \text{Males: height (cm)} &= \\ &64.19 - [0.04 \times \text{age (years)}] + [2.02 \times \text{KH (cm)}]\end{aligned}$$

UL was measured in the sitting position on the left arm. The shoulder was placed in adduction and internal rotation, with the elbow bent at 45° and the palm of the hand placed on the chest with fingers extended. The distance between the proximal end of the ulna at the elbow and the point of the styloid apophysis at the wrist was measured (fig. 1b). The mean of 2 measurements was calculated to the nearest 0.5 cm and used to predict height using a standardised table [19].

DS was measured in the sitting position on the left arm. The arm was raised to the level of the shoulder and outstretched with the fingers extended. The distance between the centre of the su-

Fig. 1. Various alternative measures for estimating height. **a** Knee height (KH). **b** Ulna length (UL). **c** Demi-span (DS). **d** Arm span (AS).

prasternal hollow and the root of the middle finger was measured (fig. 1c). The mean of 2 measurements was used to estimate height in the following equations [13]:

Females: height (cm) = $60.1 + [1.35 \times \text{DS (cm)}]$

Males: height (cm) = $57.8 + [1.40 \times \text{DS (cm)}]$

AS was measured in the sitting position, both arms outstretched horizontally at 90° in the Frankfurt plane. The distance between the end of the middle finger of each arm was measured (fig. 1d). The mean of 2 measurements was used to directly estimate height [14].

Spirometry was performed with a flow spirometer (model Vmax 22; Sensormedics, Yorba Linda, Calif., USA) following the guidelines of the ATS/ERS Statement [2, 20]. Briefly, spirometry was performed with the patient in the sitting position wearing a nasal peg. The spirometer was calibrated daily using a 3-litre sy-

ringe (Hans Rudolph, Inc., Kansas City, Mo., USA), following the manufacturer's instructions. Barometric pressure, room temperature and humidity were recorded. The results were corrected for BTPS conditions. Spirometric manoeuvres were done following standard procedure. They were considered repeatable when the 2 largest values of FVC and FEV₁ were within 0.150 litres of each other. If the first 3 manoeuvres were not satisfactory, additional manoeuvres were done until the criteria were met or a total of 8 manoeuvres had been performed. The largest FVC and FEV₁ obtained in an acceptable test were recorded for each subject.

Predicted values of FVC (FVC pred) and FEV₁ (FEV₁ pred) were calculated using prediction equations for a healthy Spanish population proposed by the Sociedad Española de Neumología y Cirugía Torácica (SEPAR) [7]. In subjects ≥65 years old, prediction equations proposed for an elderly European population were used [21].

Five groups of values for FVC pred and FEV₁ pred were calculated for each subject: one based on SH and the others on the height estimated using each of the alternative measures (KH, UL, DS and AS).

Analysis

The data were analysed using the statistical packages SPSS for Windows (SPSS version 13; SPSS Inc., Chicago, Ill., USA) and MedCalc (version 7.2.1.0; MedCalc Software, Mariakerke, Belgium). The Kolmogorov-Smirnov test was used to evaluate the normality of the distribution of all the quantitative variables.

Predicted pulmonary function values obtained using each alternative measure were compared with those based on SH, using the Student t test for paired data or the Mann-Whitney test.

The reproducibility was analysed by means of the intraclass correlation coefficient (ICC) [22], using a random effect model and taking SH as the standard method with which to compare the others. Briefly, the ICC was designed to evaluate the agreement between 2 or more quantitative variables, including questions of reproducibility. It estimates the equivalence of repeated measurements taken in the same subjects.

The method described by Bland and Altman [23] was used to evaluate agreement, taking SH as the standard method with which to compare the rest. This method is used to investigate how close the agreement is between 2 methods of measuring the same parameter. The agreement is determined by plotting a graph in which the means of the results obtained by using each method are compared, together with the absolute differences between them for each individual. The limits of agreement are defined as the mean difference $\pm$ 1.96 SD and show to what extent the tested method varies compared to the standard method, with 95% of the differences falling within these limits. For all analyses, the significance level was set at $p \leq 0.05$.

Results

A total of 118 subjects were referred for pulmonary function evaluation. Sixteen subjects were excluded due to causes that made it impossible to measure their height accurately: advanced age (2), extreme debility (1), neuromuscular disease (9), deformity of the axial skeleton (3) and deformity or amputation of limbs (1). Two subjects revoked their informed consent after inclusion. One hundred subjects (53 males and 47 females) were finally included with a weight of 74 ± 18 kg, SH of 164 ± 11 cm, body surface area of 1.80 ± 0.25 m² and BMI of 27.51 ± 1.53 . The age of the included subjects showed a normal distribution in the Kolmogorov-Smirnov test (mean age 55 ± 19 years; range 5–81, lower quartile 43, median quartile 58, upper quartile 73).

When comparing the differences between the height estimated using each alternative measure and measured SH, statistically significant differences were found (table 1). When the values for FVC pred and FEV₁ pred de-

Table 1. Comparison of measured SH with height estimated using each alternative measure

	Mean, cm	SD, cm	p
SH	164	11	
KH	169	9	<0.001
UL	174	9	<0.001
DS	181	9	<0.001
AS	173	14	<0.001

Table 2. Comparison of FVC pred and FEV₁ pred values obtained using each alternative measure with the predicted values based on SH

		Mean, l	SD, l	p
FVC pred	SH	3.75	1.03	
	KH	4.02	0.93	<0.001
	UL	4.37	0.98	<0.001
	DS	4.83	0.92	<0.001
	AS	4.32	1.15	<0.001
FEV ₁ pred	SH	2.94	0.84	
	KH	3.13	0.78	<0.001
	UL	3.38	0.82	<0.001
	DS	3.68	0.77	<0.001
	AS	3.34	0.93	<0.001

Table 3. ICC and degree of reproducibility with its corresponding CI for FVC pred and FEV₁ pred for each alternative measure

		ICC	95% CI	Degree
FVC pred	KH-SH	0.93	0.90–0.95	excellent
	UL-SH	0.88	0.83–0.92	excellent
	DS-SH	0.78	0.69–0.84	excellent
	AS-SH	0.90	0.84–0.92	excellent
FEV ₁ pred	KH-SH	0.95	0.93–0.97	excellent
	UL-SH	0.93	0.89–0.95	excellent
	DS-SH	0.89	0.84–0.93	excellent
	AS-SH	0.93	0.90–0.95	excellent

rived from each alternative measure were compared with those calculated using SH, statistically significant differences were found in each case (table 2).

The ICC for FVC pred, depending on the measure used, were: 0.93 for KH, 0.88 for UL, 0.78 for DS and 0.90 for AS. In the case of FEV₁ pred, the ICC were: 0.95 for

Fig. 2. Analysis of agreement for FVC pred using the Bland-Altman plot. **a** Difference in FVC pred between KH and SH vs. mean. **b** Difference in FVC pred between UL and SH vs. mean. **c** Difference in FVC pred between DS and SH vs. mean. **d** Difference in FVC pred between AS and SH vs. mean.

KH, 0.93 for UL, 0.89 for DS and 0.93 for AS (table 3). All the results reflect an excellent degree of reproducibility.

Taking SH as the standard method with which to compare the rest, the Bland-Altman test was used to compare the FVC pred calculated using each alternative measure (fig. 2). The same procedure was used for each FEV₁ pred (fig. 3). Both figures show that wide limits of agreement are obtained with the 4 alternative measures, and they all overestimate pulmonary function values, especially DS (mean difference: 1.1 litres for FVC pred; 0.74 litres for

FEV₁ pred). Estimations based on KH were the closest (mean difference: 0.26 litres for FVC pred; 0.19 litres for FEV₁ pred).

Discussion

Pulmonary function laboratories use standardised reference values based on ethnic and anthropometric characteristics such as SH [20]. It is increasingly necessary to measure pulmonary function in patients who are

Fig. 3. Analysis of agreement for FEV_1 pred using the Bland-Altman plot. **a** Difference in FEV_1 pred between KH and SH vs. mean. **b** Difference in FEV_1 pred between UL and SH vs. mean. **c** Difference in FEV_1 pred between DS and SH vs. mean. **d** Difference in FEV_1 pred between AS and SH vs. mean.

not able to stand and so their SH cannot be accurately measured. Due to the progressive aging of the population, some authors have suggested using reference equations exclusively for the elderly population [21]. Others used AS as the parameter to estimate height [15] or to construct reference equations for pulmonary function directly [16]. The latter could be the ideal option but the need for various groups of reference equations increases the complexity of the process. Our approach uses anthropometric measurements to estimate SH since we believe that an application of this technique in clinical

practice is feasible and may thus become more widespread.

Various suggestions have been made for estimating height such as using KH, UL, DS and AS [8–16], but to date the impact of these alternative measures on the variability of pulmonary function, and thus their practical usefulness, has not been evaluated.

Our study shows that the differences between the FVC pred and FEV_1 pred values calculated on the basis of each alternative measure and those calculated on the basis of SH are statistically significant. The reproducibility is ex-

cellent, although agreement is poor and all the measures overestimate FVC pred and FEV₁ pred, especially DS (mean difference: 1.1 and 0.74 litres, respectively). Estimations based on KH are the closest (mean difference: 0.26 and 0.19 litres, respectively). In this respect, Chumlea et al. [8] proposed using KH to estimate height for the purpose of calculating nutritional indexes when it was not possible to measure stature directly.

Although the results obtained show that the predicted values based on height estimation are not similar to those based on SH, KH could be used in patients who are unable to stand since this has been shown to agree most closely with SH. However, the tendency to overestimate height and, consequently, predicted pulmonary function values should be borne in mind. It should be remembered that solving the equations does not give the actual height but simply provides a general approximation to stature. In fact, if the systematic error in the estimations is known, it may be corrected to give values very close to the real ones. Moreover, the measurement of pulmonary function has an eminently clinical purpose, as the aim is not only to determine it at a given moment, but also to show how it evolves in an individual over time. Thus, the error introduced by the chosen method of estimation is minimised when this method is used systematically in the same patient.

To the best of our knowledge, ours is the first study that evaluates various alternative measures for estimating height for the purpose of predicting pulmonary function. Hickson and Frost [18] compared 3 measures (KH, DS and AS) for nutritional purposes in a population of 484 patients and concluded that there was poor agreement between all 3 and SH. Our results show a greater mean difference. We hypothesise that the difference in agreement may partly be explained by the fact that the measures are used to predict a subject's maximum height rather than his or her actual height, which might not be the same since aging itself causes variations in height even in those in whom SH can be measured.

The reproducibility of the different measures is high and the results obtained using the different measures in the same population are extremely reliable. This suggests that most of the variability observed is due to the individuals and not the observer.

However, agreement is poor with wide limits of agreement. Taking the upper limits of agreement of the best measure (KH), FVC pred may be overestimated by 0.980 litres, and FEV₁ pred by 0.680 litres, which could lead to errors when interpreting spirometric data. These changes in FVC pred and FEV₁ pred could mean that the nor-

mality limits are exceeded, with the result that spirometry is erroneously classified as abnormal when, if it were possible to measure SH, it would be interpreted as normal. In the case of a patient with an obstructive disorder, an overestimation of 0.680 litres in FEV₁ pred could lead to reclassification into a higher severity category. Since this phenomenon is more marked in the case of FVC pred (0.980 litres), there might be a tendency to overdiagnose restrictive disorders or overvalue those already diagnosed. All these considerations should be borne in mind when interpreting pulmonary function tests in these patients.

A potential limitation of this study is that the mean difference in predicted pulmonary function values obtained with the best measure (0.260 litres for FVC; 0.190 litres for FEV₁) is greater than the current repeatability limit established by the main international guidelines (0.150 litres for both) [20]. We agree that this is not desirable but, when faced with this situation (inevitable and ever more frequent) in the laboratory, we take the best option available, especially since, at present, there is no alternative for this type of patient.

In contrast to other studies, we did not use callipers to measure KH for practical reasons. Since metal tape measures are universally available, their use simplifies the procedure. However, using a metal tape measure to take anthropometric measurements has been shown to result in an overestimation of 1.03 cm [24]. This may also account for part of the overestimation found; however, the degree of error introduced is similar to that obtained in the normal practice of measuring SH to the nearest centimetre.

In summary, the alternative measures used to estimate height introduce a certain degree of error in predicted pulmonary function and this error should be quantified in each case since it results in overestimation of the predicted values. In patients in whom SH cannot be measured, the alternative measures may be the best option to obtain a predicted value, as close as possible to the real one, which may be used in the prediction equations. In this way, approximate predicted pulmonary function values are obtained which should be evaluated in the clinical setting. Of the measures analysed, KH was found to be the most recommendable. Further studies are necessary to determine its true clinical significance.

References

- 1 Handa T, Nagai S, Hirai T, Chin K, Kubo T, Oga T, Niimi A, Matsumoto H, Ito Y, Takahashi K, Watanabe K, Izumi T, Mishima M: Computed tomography analysis of airway dimensions and lung density in patients with sarcoidosis. *Respiration* 2009;77:273–281.
- 2 Ferrazza AM, Martolini D, Valli G, Palange P: Cardiopulmonary exercise testing in the functional and prognostic evaluation of patients with pulmonary diseases. *Respiration* 2009;77:3–17.
- 3 Miller MR, Crapo R, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Enright P, van der Grinten CPM, Gustafsson P, Jensen R, Johnson DC, MacIntyre N, McKay R, Navajas D, Pedersen OF, Pellegrino R, Viegi G, Wanger J: General considerations for lung function testing. *Eur Respir J* 2005;26:153–161.
- 4 Koegelenberg CF, Diacon AH, Irani S, Bolliger CT: Stair climbing in the functional assessment of lung resection candidates. *Respiration* 2008;75:374–379.
- 5 Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC: Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official statement of the European Respiratory Society. *Eur Respir J Suppl* 1993;16:5–40.
- 6 American Thoracic Society: Standardization of spirometry: 1994 update. *Am J Respir Crit Care Med* 1995;152:1107–1136.
- 7 Roca J, Sanchis J, Agustí-Vidal A, Segarra F, Navajas D, Rodriguez-Roisin R, Casan P, Sans S: Spirometric reference values for a Mediterranean population. *Bull Eur Physio-pathol Respir* 1986;22:217–224.
- 8 Chumlea WC, Roche AF, Steinbaugh ML: Estimation of stature from knee height for persons 60 to 90 years of age. *J Am Geriatr Soc* 1985;33:116–120.
- 9 Chumlea WC, Guo SS, Wholihan K, Cockram D, Kuczmarski RJ, Johnson CL: Stature prediction equations for elderly non-Hispanic white, non-Hispanic black, and Mexican American persons developed from NHANES-III data. *J Am Diet Assoc* 1998;98:137–142.
- 10 Physical status: the use and interpretation of anthropometry. Report of a WHO expert committee. World Health Organization Report Series No. 854. Geneva, World Health Organization, 1995.
- 11 Gauld LM, Kappers J, Carlin JB, Robertson CF: Height prediction from ulna length. *Dev Med Child Neurol* 2004;46:475–480.
- 12 Gauld LM, Kappers J, Carlin JB, Robertson CF: Prediction of childhood pulmonary function using ulna length. *Am J Respir Crit Care Med* 2003;168:804–809.
- 13 Bassey EJ: Demi-span as a measure of skeletal size. *Ann Hum Biol* 1986;13:499–502.
- 14 Hibbert ME, Lanigan A, Raven J, Phelan PD: Relation of arm span to height and the prediction of lung function. *Thorax* 1988;43:657–659.
- 15 Aggarwal AN, Gupta D, Jindal SK: Impact of substitution of arm span for standing height in adults from North India. *Chest* 1999;115:557–562.
- 16 Golshan M, Crapo RO, Amra B, Jensen RL, Golshan R: Arm span as an independent predictor of pulmonary function parameters: validation and reference values. *Respirology* 2007;12:361–366.
- 17 Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, Coates A, van der Grinten CPM, Gustafsson P, Hankinson J, Jensen R, Johnson DC, MacIntyre N, McKay R, Miller MR, Navajas D, Pedersen OF, Wanger J: Interpretative strategies for lung function tests. *Eur Respir J* 2005;26:948–968.
- 18 Hickson M, Frost G: A comparison of three alternative methods for estimating height in the acute ill elderly population. *J Hum Nutr Diet* 2003;16:13–20.
- 19 Todorovic V, Russell C, Stratton R, Ward J, Elia M: The 'MUST' Explanatory Booklet: A Guide to the 'Malnutrition Universal Screening Tool' ('MUST') for Adults. Malnutrition Advisory Group: a Standing Committee of BAPEN. Redditch, BAPEN, 2003, pp 10–12.
- 20 Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P, van der Grinten CPM, Gustafsson P, Jensen R, Johnson DC, MacIntyre N, McKay R, Navajas D, Pedersen OF, Pellegrino R, Viegi G, Wanger J: Standardisation of spirometry. *Eur Respir J* 2005;26:319–338.
- 21 García-Río F, Pino JM, Dorgham A, Alonso A, Villamor J: Spirometric reference equations for European females and males aged 65–85 years. *Eur Respir J* 2004;24:397–405.
- 22 Müller R, Büttner P: A critical discussion of intraclass correlation coefficients. *Stat Med* 1994;13:2465–2476.
- 23 Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–310.
- 24 Spender QW, Cronk CE, Charney EB, Stallings VA: Assessment of linear growth of children with cerebral palsy: use of alternative measures to height or length. *Dev Med Child Neurol* 1989;31:206–214.