Differences in spirometry reference values: a statistical comparison of a Mongolian and a Caucasian study

R.O. Crapo*, R.L. Jensen*, M. Oyunchimeg**, T. Tsh**, C. DuWayne Schmidt*

Differences in spirometry reference values: a statistical comparison of a Mongolian and a Caucasian study. R.O. Crapo, R.L. Jensen, M. Oyunchimeg, T. Tsh, C. DuWayne Schmidt. ©ERS Journals Ltd 1999.

ABSTRACT: New reference value studies for spirometry are commonly compared to existing reference value studies using average data derived from existing reference equations. Such comparisons are inherently flawed because they do not account for differences in distributions of the independent and dependent variables and they do not have identical methodologies. This study was undertaken 1) to derive reference equations for forced vital capacity (FVC) and forced expiratory volume in one second (FEV1) for natives of Mongolia and 2) to compare the Mongolian data with data from a 1981 reference study of Caucasians in Salt Lake City, UT, USA.

Spirometry was performed on 344 (176 male, 168 female) healthy, nonsmoking urban natives of Mongolia to generate reference equations for FVC, FEV1, and FEV1/ FVC. These data were compared with data from a 1981 reference study of Caucasians in Salt Lake City, using both an analysis of covariance of the raw data and parametric and nonparametric comparisons of a matched pair subset.

Average measured forced vital capacity and forced expiratory volume in one second in native Mongolians were within 1–2% of the Caucasian predicted values. These small differences are not statistically significant in any of the multiple methods of comparisons. Power analysis suggests that, if real differences exist, the differences in forced vital capacity are <155 mL for males, <105 mL for females, and the differences in forced expiratory volume in one second are <107 mL for males and <76 mL for females.
Eur Respir J 1999; 13: 606–609.

*Pulmonary Division, LDS Hospital and University of Utah School of Medicine, Salt Lake City, Utah, USA. **The National Medical University of Mongolia, Ulaanbaatar, Mongolia.

Correspondence: R.O. Crapo
Pulmonary Division
LDS Hospital
8th Ave and C Street
Salt Lake City
UT 84143
USA
Fax: 1 8014081671

Differences in ethnic background have been shown to be an important determinant of pulmonary function measurements [1–4]. Spirometry reference values have not previously been established for Mongolians living in the People's Republic of Mongolia. The purposes of this study were: 1) to establish reference data based on healthy nonsmoking Mongolian subjects and 2) to compare the Mongolian data with data collected in 1981 on a Caucasian population in the USA at a similar altitude using similar testing procedures.

Methods

Healthy Mongolian subjects were recruited in Ulaanbaatar, the capital city of Mongolia (1,350 m elevation). Volunteers associated with the National Medical University of Mongolia were recruited. This convenience sample included medical students, physicians, other health workers, military personnel and relatives of hospitalized patients. Each volunteer completed a health questionnaire and submitted to physical examination of the chest. The criteria for inclusion were: 1) no smoking in the 6 months before testing and a cumulative smoking history of <0.5 pack-yr; 2) no history or symptoms of past or present heart, lung, or thoracic cage diseases; 3) no known exposure to substances known to cause lung injury; and 4) normal physical examination of the chest.

Height and weight were measured with subjects wearing indoor clothing without shoes. Forced vital capacity (FVC), forced expiratory volume in one second (FEV1), and FEV1/FVC were measured from kymograph tracings obtained with a Collins Survey II water-seal spirometer (Warren E. Collins, Braintree, MA, USA). The spirometer calibration was checked daily with a 3.0-L syringe, and barometric pressure was recorded daily. The spirometer temperature, recorded at each study, was used to adjust FVC and FEV1 to body temperature and ambient pressure, and saturated with water vapour (BTPS) conditions. The average machine temperature was 22.4°C and ranged 16–32°C. Studies were performed using the acceptability standards outlined by the American Thoracic Society (ATS) [5, 6] with subjects sitting upright, wearing nose clips.

Spirogram tracings were measured manually by two of the authors. Individual elements of a spirogram were considered to be acceptable even if the entire test did not meet ATS standards. For example, if a tracing met acceptable start of test criteria and had an acceptable FEV1 but failed end-of-test criteria, FEV1 was accepted but FVC and FEV1/FVC were excluded. The largest acceptable FVC and FEV1 for each individual were used in the analysis.

Data were analysed using Statistica software (StatSoft, Tulsa, OK, USA). Backward, stepwise multiple linear regressions were performed with FVC, FEV1 and FEV1/

Table 1. – Demographic information

	Males	Females
Mongolian data		
n	176	168
Age yrs	33±11 (32) (18–54)	34±11 (34) (18–54)
Height cm	168.4±6.1 (168.0) (154–186)	158.0±5.0 (158.0) (146–170)
BMI kg·m^{-2}	23.7±5.2 (22.9) (18.7–31.3)	23.6±3.9 (22.9) (16.6–38.1)
Caucasian data		
n	125	126
Age yrs	49±20 (49.0) (15–91)	49±20 (50.0) (17–84)
Height cm	175.1±6.7 (175.0) (157–194)	161.4±6.7 (161.0) (146–178)
BMI kg·m^{-2}	26.4±4.7 (25.9) (18.6–55.8)	25.4±4.5 (24.8) (16.9–38.4)

Age, height, and BMI presented as mean±SD (median) (range). BMI: body mass index.

FVC as dependent variables, and height and age as independent variables. Independent variables were included in the final regression if they were significant at p<0.05. Per cent of predicted values were used to estimate the lower 5th percentiles.

Mongolian data were compared with 1981 reference data generated on Caucasian subjects of northern European descent in Salt Lake City [7]. These comparisons differ from most other such comparisons because raw data were available for both studies. Both studies recruited volunteers, and the testing conditions and altitudes were similar (the average barometric pressure in Salt Lake City is 86.1 kPa (647 mmHg) and in Ulaanbaatar is 86.3 kPa (649 mmHg)). The two water-sealed spirometers differed slightly. The Salt Lake City study used a Collins 13.5-L metal bell spirometer; the spirometer used in Mongolia was the Collins Survey II spirometer with a plastic bell.

FVC and FEV1 were compared in several ways. Using an analysis of covariance, tests were made to determine whether the age and height coefficients were significantly different from one another. If these coefficients were different, an average plane was created and used to test for treatment effects of the covariate. Numerically, this was a test for significant differences in the intercept terms of the equations, performed after forcing the two planes into one by using all data to calculate one age and one height coefficient.

The per cent of predicted value for each Mongolian subject was calculated using the Salt Lake City equations. These per cent of predicted values were averaged, and the number of individuals below the lower limits of the Salt Lake City reference study was used to give an overall perspective of how the Salt Lake City equations would predict spirometric values for Mongolian subjects. The percentage of Mongolian subjects classified as "low" using the Salt Lake City equations was compared to the traditional 5% criterion using a ratio test.

As differences in the age and height distributions between the Salt Lake City and Mongolian data were statistically significant, matched subgroups were created. For each Mongolian subject, a search for a Salt Lake City subject within 2 yrs of age and 2 cm in height was made. The distributions of FVC and FEV1 for the matched data sets were first examined using both Kolmogorov–Smirnov and Chi-squared tests to determine whether the distributions were Gaussian and then they were compared using both parametric (independent t-test) and nonparametric (Mann–Whitney U-Test) tests.

Results

A total of 565 Mongolian subjects performed spirometry. Of these, 221 (39.1%) were excluded because acceptable spirograms were not obtained even with multiple attempts. Reproducibility criteria were not used to exclude data from analysis. Short expiratory times and glottic closures were common; in 32 of the 168 females and 31 of the 176 males, only FEV1 met ATS acceptability criteria. The demographic data for this study and the Salt Lake City reference set are summarized in table 1. The average age of Mongolian subjects was about 15 yrs below that of Salt Lake City subjects. This reflects the method of selection in the Salt Lake City study (equal numbers of subjects for each decade). Mongolian males and females were, on average, 6.7 and 3.4 cm shorter, respectively, than Salt Lake City subjects. The body mass index (kg·m^{-2}) was higher in Salt Lake City subjects. Reference equations and lower 95th percentiles for Mongolian subjects are shown in table 2.

The per cent of predicted values and their 95% confidence limits calculated for Mongolian subjects using Salt Lake City equations are shown in table 3. None were significantly different from 100%. The number of Mongolian subjects classified as "low" for FVC or FEV1 ranged 1.4–2.4% and did not differ statistically from 5% (p>0.09 for all comparisons), the traditional threshold

Table 2. – Prediction equations for spirometric parameters in Mongolian males and females

	Coefficients						Normal 95th percentile (% pred)
	n	r^2	Constant	Height cm	Age yrs	SEE	
Males							
FVC L	145	0.386	-3.5922	0.05207	-0.01532	0.496	82.7
FEV1 L	176	0.545	-2.0956	0.04052	-0.02411	0.385	84.5
FEV1/FVC %	145	0.209	90.577		-0.2080	4.44	91.2
Females							
FVC L	136	0.371	-2.1706	0.03889	-0.01495	0.363	81.9
FEV1 L	168	0.576	-1.1576	0.03053	-0.02234	0.278	85.3
FEV1/FVC %	136	0.349	92.726		-0.2627	3.93	91.1

95th percentile is the lower 95th percentile (% pred). SEE: standard error of estimate; FVC: forced vital capacity; FEV1: forced expiratory volume in one second.

Table 3. – Measured Mongolian spirometry values compared to Salt Lake City prediction equations

	n	Mongolian measured values	Predicted values using Caucasian equations	Mongolian data as a per cent of Caucasian predicted values*	Percentage of Mongolian subjects below the reference range**
Males					
FVC L	145	4.622±0.632	4.683±0.487	98.9±10.70 (97.2–100.7)	1.4
FEV1 L	176	3.920±0.571	3.963±0.431	98.9±9.68 (97.6–100.3)	1.7
FEV1/FVC %	145	83.20±4.99	83.50±1.63	99.9±5.48 (88.8–100.3)	
Females					
FVC L	136	3.431±0.457	3.384±0.380	101.7±10.95 (98.8–103.6)	5.9
FEV1 L	168	2.908±0.427	2.959±0.370	98.4±9.38 (97.0–99.8)	1.8
FEV1/FVC %	136	83.34±4.87	86.11±2.67	97.2±4.58 (83.9–109.9)	

Data presented as mean±SD with 95% confidence interval in parentheses. FVC: forced vital capacity; FEV1: forced expiratory volume in one second. *: predicted Caucasian values from CRAPO *et al.* [7], 95% confidence interval for the mean value calculated using per cent of predicted; **: per cent of Mongolian subjects below the lower limits of the Salt Lake City reference range [7], none were statistically different from the traditional threshold of 5% (p>0.09).

used to determine the lower limits of "normal". These findings suggest that the Salt Lake City equations perform reasonably well in predicting reference values for Mongolian subjects.

Comparison of the raw data sets by analysis of covariance showed no significant differences between the planes or the intercepts described by the equations for FVC or FEV1 in the Mongolian and Salt Lake City datasets. Therefore, the relationship between age and height and either FVC or FEV1 was not different for the two studies.

In the matched pair comparisons, all datasets were found to be consistent with a Gaussian distribution using the Kolmogorov–Smirnov test (p>0.05). However, Chi-squared tests showed the distributions of Salt Lake City male FVC and FEV1 to be non-Gaussian (p≤0.007). The results of the independent t-tests and Mann–Whitney U-tests are summarized in table 4. No significant differences were found between the Mongolian and Salt Lake City Caucasian paired data subsets for FVC or FEV1 for either sex using these parametric and nonparametric comparisons.

An analysis of the power of the statistical comparisons revealed that for males, the analysis of covariance would allow the detection of a difference of 155 mL in FVC and 107 mL in FEV1 with 90% power at p=0.05. For females, a difference of 105 mL in FVC and 76 mL in FEV1 could be detected with the same level of power.

Discussion

Spirometric reference equations were derived from a sample of urban natives of Mongolia. No significant dif-

ferences in the relationships between height and age and spirometric variables were found between this sample of Mongolians and a 1981 study of residents of Salt Lake City using multiple methods of comparison.

The comparison of results from a new study with existing reference studies is difficult when authors do not have access to the raw data from previous studies and are limited to comparing their data with earlier equations. Such comparisons are limited because they cannot account for differences in the distributions of the independent and dependent variables. Distributions are more likely to differ when sample sizes are small, when convenience samples are used, and when real differences exist. Differences in the distributions of independent and dependent variables may lead to significant differences in reference equations that may be falsely interpreted as reflecting real differences in the population. Conversely, true differences could be missed.

Several potential sources of bias exist in the current data. These sources and the authors' attempts to address them are discussed below.

Exclusion rate

Thirty-nine per cent of the Mongolian subjects were excluded from the analysis because acceptable quality spirograms were not achieved. Although the number of exclusions may bias the study, no significant differences were found between excluded and included subjects for age, height, weight, medical history, or physical examination findings. The only pattern identified was that early

Table 4. – Comparison of matched groups

	Females						Males					
	Mongolian		Caucasian		p-values		Mongolian		Caucasian		p-values	
	n	Mean±SD	n	Mean±SD	t*	U**	n	Mean±SD	n	Mean±SD	t*	U**
Age yrs	86	33.3±11.2	86	33.4±11.3	0.96	0.27	41	26.2±10.7	41	25.5±10.7	0.79	0.10
Height cm	86	158.9±4.0	86	159.8±4.0	0.12	0.01	41	172.3±4.6	41	173.1±4.8	0.41	0.42
FVC L	70	3.48±0.43	86	3.48±0.47	0.96	0.83	31	4.93±0.62	41	5.14±0.60	0.15	0.10
FEV1 L	86	2.95±0.40	86	3.00±0.41	0.43	0.31	86	4.26±0.52	41	4.39±0.52	0.27	0.11

For each group, Mongolian subjects were matched to Salt Lake City subjects (match criteria: within 2 yrs of age and 2 cm in height). Sample size varied as a function of finding matches. *: independent t-test; **: Mann–Whitney U-test. FVC: forced vital capacity; FEV1: forced expiratory volume in one second.

glottic closures were more common in the younger subjects. It was also not required that each individual have a full set of acceptable data. The presence of early glottic closures or inadequate expiratory times led to the exclusion of some FVC values. The high rate of technical problems could be attributed to a language barrier and to the unfamiliarity of the Mongolian medical community with administering spirometry. Spirometry was introduced to the National Medical University of Mongolia by one of the authors (C. Duwayne Schmidt) at the time of the present study.

Technical difference

The technical details for the two studies were not entirely identical although a strong effort was made to match the equipment and procedural standards. The ideal way to compare the two groups would be to study both groups simultaneously in the same laboratory but this was not an option in the present study.

Sample selection

A convenience sample was used in both studies because of financial constraints. The nonrandom sample of the populations is another potential source of bias in both studies. However, VAN GANSE et al. [8] reported that the method of subject selection does not appear to influence either the mean or range of lung function values unless hospital patients are included.

Timing

The two studies were conducted about 13 yrs apart. This raises the possibility that the relationship between spirometric variables and height or age had changed over time. There is no direct evidence regarding this issue but it is doubtful that any consequential changes in these relationships would have occurred in such a short time interval.

Demographic differences

The Mongolian subjects were younger and shorter than the Salt Lake City subjects. The formal statistical comparisons showed no statistical differences between the two data sets either by analysis of variance or by the paired comparison of age and height-matched subject groups.

Conclusions

This was a negative study. No differences were found in spirometric reference values between a sample of Mongolians and a sample of Caucasians of northern European descent. The power analysis suggests that if real differences are present, they are <155 mL for male FVC, <105 mL for female FVC, <107 mL for male FEV1 and <76 mL for female FEV1. These differences are within the known <200 mL intra-individual variability of spirometric tests [6].

The problems encountered in this study illustrate some of the issues in attempting to determine ethnic differences using comparisons between separate studies. Even in the present study, where many of the interstudy variables were controlled, there were enough design and procedural differences to compromise the conclusions. The ideal way to address ethnic differences is to make ethnic comparisons part of a single study in which the study design is constant and procedures can be controlled and uniformly applied to both groups. When this is not possible, comparisons should be made using the raw data from both groups, and the results of such comparisons should be interpreted with caution.

References

1. American Thoracic Society. Lung function testing and selection of reference values and interpretive strategies. Official Statement of the American Thoracic Society. *Am Rev Respir Dis* 1991; 144: 1202–1218.
2. Schoenberg JB, Beck GJ, Bouhuys A. Growth and decay of pulmonary function in healthy blacks and whites. *Respir Physiol* 1978; 33: 367–393.
3. Schwartz JD, Katz SA, Fegly RW, Tockman MS. Sex and race differences in the development of lung function. *Am Rev Respir Dis* 1988; 138: 1415–1421.
4. Oscherwitz M, Edlavitch SA, Baker TR, *et al*. Differences in pulmonary functions in various racial groups. *Am J Epidemiol* 1972; 96: 319–327.
5. American Thoracic Society. Standardization of spirometry – 1987 Update. *Am Rev Respir Dis* 1987; 136: 1285–1298.
6. American Thoracic Society. Standardization of spirometry: 1994 Update. Official Statetement of the American Thoracic Society. *Am J Respir Crit Med* 1995; 152: 1107–1136.
7. Crapo RO, Morris AH, Gardner RM. Reference spirometric values using techniques and equipment that meet the ATS recommendations. *Am Rev Respir Dis* 1981; 123: 659–664.
8. Van Ganse W, Billiet L, Ferris BG Jr. Medical criteria for the selection of normal subjects. *In*: Arcangeli P, Cotes JE, Cournand A, Denolin H, Maria GD, Sadoul P, Scherrer M, Scarpa GL, eds. Introduction to the Definition on Normal Values for Respiratory Function in Man. Alghero, Italy, Panminerva Medica, 1969; pp. 15–27.