Effect of change of reference standard to NHANES III on interpretation of spirometric 'abnormality'

Akshay Sood
Beth K Dawson
Joseph Q Henkle
Patricia Hopkins-Price
Clifford Qualls

Department of Medicine, University of New Mexico Health Sciences Center School of Medicine, Albuquerque, NM, USA (AS, CQ), and the Division of Pulmonary and Critical Care Medicine (JQH, PHP) and the Division of Biostatistics and Research (BKD), Department of Medicine, Southern Illinois University School of Medicine, Springfield, IL, USA. 62794-9636

Correspondence: Akshay Sood
University of New Mexico Health Sciences Center School of Medicine Department of Medicine
1 University of New Mexico, MSC 10
5550 Albuquerque, NM 87131, USA
Tel + 505 272 4751
Fax + 505 272 8700
Email asood@salud.unm.edu

Abstract: The American Thoracic Society has recently recommended the use of NHANES III spirometric reference standard in the United States. The objective of this study was to better quantify the well-known 'problem' of the change in interpretation of spirometry, as a consequence of the change from the other commonly used reference standards (Morris, Kory, Crapo, Knudson 1976, and Knudson 1983) to NHANES III. This is a cross-sectional study of spirometries of 1,106 non-Hispanic Caucasian American adults, including 234 subjects with obstructive and 228 subjects with restrictive spirometric 'abnormalities'. A weighted Kappa statistic was used to evaluate the level of agreement between NHANES III and other commonly used reference standards. The level of agreement in assessing the presence of an 'abnormality' was poor to moderate – values of Kappa statistic ranged from 0.13 to 0.46. There was however, good to very good level of agreement in assessing the severity of the 'abnormality' – values of Kappa statistic ranged from 0.61 to 0.91. This study better quantifies the well-known differences in the interpretation of spirometric 'abnormalities' as a consequence of the recommended change of reference standard to NHANES III, which in turn may cause confusion among patients and their treating physicians.

Keywords: Reference standard, Classification of severity, Obstructive spirometric abnormality', Restrictive spirometric 'abnormality', Kappa statistic.

Introduction

Spirometry, the most frequently performed pulmonary function test (other than arterial blood gas study), plays an important role in diagnosing the presence and type of lung 'abnormality' and classifying its severity. It plays a key role in medical surveillance examinations for occupational lung diseases, in determining whether to institute preventive or therapeutic measures, and in granting benefits to individuals with lung impairment. Observed spirometric data are compared to reference data and are expressed as percent predicted values, based on age, gender, height and race (American Thoracic Society 1991). The purpose of such a comparison is to determine 'normality' vs. 'abnormality' and in cases of 'abnormality', to determine its 'severity'. The reference value is calculated from a regression equation derived from a population of 'normal' subjects.

In the United States, a variety of reference standards are available for commercial use by pulmonary function testing laboratories. The American Thoracic Society and European Respiratory Society (ATS-ERS) have recently recommended the use of the third National Health and Nutrition Examination Survey (NHANES III) reference standard for the interpretation of spirometry in the United States (Pellegrino et al 2005). This reference standard provides ethnically appropriate equations for Caucasian Americans, African Americans, and Mexican Americans, thus obviating the need for less suitable and arbitrary race/ethnic adjustment factors (Hankinson et al 1999). Over time, the NHANES III reference will likely become the standard in most pulmonary function testing laboratories around the country.

The purpose of this study was to characterize the well-known 'problem' of the changed interpretation of the presence and severity of a spirometric 'abnormality' in Caucasian Americans as a consequence of the change from other commonly used reference standards (Morris et al 1976; Knudson 1983) to NHANES III by pulmonary function testing laboratories in the United States. Although the concept of disagreement between the various reference standards is well-known (Morris et al 1971; Crapo et al 1981; Baur et al 1999; Hankinson et al 1999; Subbarao et al 2004), the objective of this study was to better quantify these differences with respect to spirometric interpretation.

Methods

This is a cross-sectional study of 1,106 non-Hispanic Caucasian American adult subjects who were referred by their physicians for spirometric testing at a single pulmonary function laboratory at a teaching hospital in Central Illinois. The spirometries were performed by trained technicians, using standard equipment and techniques that met the American Thoracic Society (ATS) criteria (1987, 1995). This study was limited to non-Hispanic Caucasian American subjects to minimize the effect of race and ethnicity on the study results. The subjects were weighed and height measured, in indoor clothing without shoes, using a calibrated scale and stadiometer respectively. Age was recorded to the nearest birthday. The research using human subjects was conducted in accordance with the Helsinki Declaration and was approved by the local Institutional Review Board.

A spirometry was considered 'abnormal' if it met either of the following obstructive or restrictive criteria. Obstructive spirometric 'abnormality' was defined as the ratio of the forced expiratory volume in one second to the forced vital capacity (FEV_1/FVC) below the lower limit of normal for the reference standard. Restrictive spirometric 'abnormality' was defined as the FEV_1/FVC ratio greater than or equal to the lower limit of normal and FVC below the lower limit of normal for the reference standard. The lower limit of normal for FEV_1, FVC, and FEV_1/FVC ratio were directly obtained from the literature for the NHANES III reference standard (Hankinson et al 1999). The lower limit of normal for the above parameters for the Morris (Morris et al 1971); Crapo (Crapo et al 1981), and Knudson (Knudson et al 1983) reference standards were calculated by using the following equation – Lower limit of normal for the parameter = Predicted value of the parameter –1.645*Standard error of the estimate of the parameter.

In order to further classify the severity of spirometric 'abnormality', regression equations from the NHANES III were first used, as described above, to define the population with a spirometric 'abnormality' (Hankinson et al 1999). Subsequently, the severity of spirometric 'abnormality' was calculated in this population for each of the various commonly used reference standards, based upon the percentage predicted FEV_1 values. The reference standards used for this part of the study included those by Morris (Morris 1976; Morris et al 1971, 1973); Kory (Kory et al 1961); Crapo (Crapo et al 1981); Knudson (Knudson et al 1976); Knudson (Knudson et al 1983); and NHANES III (Hankinson et al 1999). The use of percent predicted FEV_1 value in classifying severity of obstructive or restrictive spirometric 'abnormality' (as outlined in Table 1) has been recommended in the recent ATS/ERS guidelines (Pellegrino et al 2005). The use of a 'gold standard' reference standard (ie, NHANES III) in initially defining the presence of 'abnormality' allowed the use of a uniform population for this comparison.

Statistical analysis

A weighted Kappa (κ) statistic was used to evaluate the level of agreement between the various reference standards in classifying the presence and severity of obstructive and restrictive spirometric 'abnormalities'. κ is a measure of agreement compared to chance agreement. The κ statistic was interpreted using the guidelines suggested by Altman in Table 2 (Altman 1991). A generalized McNemar test was performed to evaluate for the presence of bias. A p value of < 0.05 was considered to be significant.

Results

The 1,106 non-Hispanic Caucasian American adult subjects studied were mixed rural and urban residents, smokers and non-smokers, diseased and healthy individuals, who underwent a spirometry for diagnostic purposes at a Midwestern teaching hospital. They included 706 women (63.87%) and 400 men (36.13%), ranging in age from 18 to 91 years, with a mean age of 49.19 years. The body mass index ranged from 10.22 to 62.76 kg/m2, with a mean body mass index of 27.41 ± 7.04 kg/m2. The 1,106 subjects included 644 with no 'abnormalities', 228 with restrictive, and 234 with obstructive 'abnormalities', as per the NHANES III reference values and lower limits of the normal range.

For comparing the presence of any spirometric 'abnormality', the level of agreement between the other reference standards and NHANES III varied from poor

Table 1 Classification of severity of any spirometric 'abnormality', based upon the forced expiratory volume in one second (FEV_1) (Pellegrino et al 2005)

Percent predicted FEV_1	Degree of severity
⩾70	Mild
60–69	Moderate
50–59	Moderately severe
35–49	Severe
<35	Very severe

(κ of 0.13 between Crapo and NHANES III) to moderate (κ of 0.44 between Knudson 1983 and NHANES III, and 0.46 between Morris and NHANES III), as shown in Table 3. As compared to the NHANES III reference standard, the use of Morris et al (1983) reference standards was associated with an increased likelihood of interpreting the presence of any 'abnormality' and of obstructive 'abnormality' and reduced likelihood of interpreting the presence of a restrictive spirometric 'abnormality'.

In order to compare the classification of severity of lung impairment, the study evaluated 234 subjects with obstructive and 228 subjects with restrictive spirometric 'abnormalities', as defined by the NHANES III reference standard (Hankinson et al 1999). For classifying the severity of obstructive spirometric 'abnormalities', the level of agreement between the other reference standards and NHANES III varied from good (κ of 0.77 between Morris and NHANES III) to very good (κ of 0.91 between Crapo and NHANES III), as shown in Table 4. As compared to the NHANES III reference standard, the 'abnormality' was classified as less severe by the Morris et al (1976) reference standards ($p < 0.001$), usually less severe by the Knudson (1983) reference standards (although the trend for the latter was not significant, $p = 0.06$), and of similar severity by the Crapo reference standard ($p = 0.77$).

A similar approach was used for comparing the severity classification of restrictive spirometric 'abnormalities' between the various reference standards. The level of agreement between the other reference standards and NHANES III for classifying the severity of restrictive 'abnormality' varied from good (κ of 0.61 between Morris and NHANES III) to very good (κ of 0.87 between Crapo and NHANES III), as shown in Table 5. As compared to the NHANES III reference standard, the 'abnormality' was classified as less severe by the Morris et al (1976) reference standards ($p < 0.001$), usually less severe by the Knudson (1983) reference standard (although the trend for the latter was not significant, $p = 0.06$), and of similar severity by the Crapo reference standard ($p = 1.0$). For both obstructive and restrictive spirometric 'abnormalities', the percent disagreement with NHANES III on severity rating was the largest when Morris was the original reference standard used.

Table 2 Guidelines for interpreting Kappa (κ) statistic (Altman 1991)

Kappa (κ) statistic	Level of agreement
0.81–1.00	Very good agreement
0.61–0.80	Good agreement
0.41–0.60	Moderate agreement
0.21–0.40	Fair agreement
≤0.20	Poor agreement

Discussion

As a US-based pulmonary function laboratory considers switching to the ATS/ERS-recommended NHANES III reference standard, it needs to be aware that the presence and severity of spirometric 'abnormality' may be classified differently, causing confusion in the minds of both the patients and their treating physicians. The extent of disagreement with the NHANES III-based interpretation will vary, depending upon the reference standard originally used. Particular caution must be exercised if the Crapo reference standard was originally used to classify the presence of spirometric 'abnormality' or if the Morris reference standard was originally used to classify the severity of spirometric 'abnormality' (Morris et al 1971). This is particularly important since Morris and Crapo were the two most common reference standards previously used in the United States, based upon a study in 1990 (Ghio et al 1990).

The 1991 ATS guidelines had recommended that the best way for each pulmonary function testing laboratory was to perform its own reference value study. The advantage of this approach was that it minimized biological variation (since the reference population was a sample of the population served by the laboratory), and analytical imprecision (since the same instruments, technical staff, and procedures were used for both the reference population and patients). However, the disadvantage of this approach was that it required a relatively large number of 'healthy' subjects to be tested by each laboratory (Pellegrino et al 2005). Unfortunately, reference standards in a laboratory were often chosen because they were available in the pulmonary function test equipment of the laboratory, rather than because they had been analyzed and found to be the best for the local population. The 2005

Table 3 Effect of change of reference standard to NHANES III (Hankinson et al 1999) on the presence of any spirometric 'abnormality' (n = 1,106, of which 644 had no 'abnormalities', 228 had restrictive and 234 had obstructive 'abnormalities' as per the NHANES III standard)

Reference equation	Percent disagreement	Weighted Kappa *	Generalized McNemar's test	Comment on comparison with NHANES III reference standard
Morris et al 1971	33.88%	0.46 (0.41, 0.50)	$p < 0.001$	Morris more likely to show any 'abnormality', or obstructive 'abnormality'; and less likely to show restrictive 'abnormality'. Morris is also more likely to interpret restrictive 'abnormality' as normal or obstructive
Crapo et al 1981	61.70%	0.13 (0.10, 0.15)	$p < 0.001$	Crapo more likely to show any 'abnormality', or obstructive 'abnormality'; and less likely to show restrictive 'abnormality'. Crapo is also more likely to interpret restrictive 'abnormality' as normal or obstructive
Knudson et al 1983	35.14%	0.44 (0.40, 0.49)	$p < 0.001$	Knudson more likely to show any 'abnormality', or obstructive 'abnormality'; and less likely to show restrictive 'abnormality'. Knudson is also more likely to interpret restrictive 'abnormality' as normal or obstructive

* Numbers in parenthesis reflect 95% confidence intervals

ATS/ERS guidelines therefore, recognized that the previous recommendation (American Thoracic Society 1991) that each laboratory perform its own reference value study was impractical for most laboratories (Pellegrino et al 2005).

However, the previous recommendation (1991) also resulted in a variety of reference standards being used by the laboratories in the United States, posing a challenge for the geographically mobile American population (Sharma, 1995) and their providers. For example, a 1990 questionnaire survey of 139 adult respiratory disease training programs in the United States and Canada (Ghio et al 1990) revealed the use of the Morris reference standard by 47% (Morris

Table 4 Effect of change of reference standard to NHANES III (Hankinson et al 1999) on the classification of severity of obstructive spirometric 'abnormality' (n = 234), using percent predicted FEV_1

Reference equation	Percent disagreement	Weighted Kappa *	Generalized McNemar's test	Comment on comparison with NHANES III reference standard
Morris et al 1971	28.20	0.77 (0.72, 0.82)	$p < 0.001$	Morris classifies a lower level of severity
Kory et al 1961	17.94	0.85 (0.81, 0.90)	$p < 0.001$	Kory classifies a lower level of severity
Crapo et al 1981	11.96	0.91 (0.88, 0.94)	$p = 0.77$	Crapo classifies a similar level of severity
Knudson et al 1976	18.80	0.85 (0.80, 0.89)	$p < 0.001$	Knudson usually classifies a lower level of severity
Knudson et al 1983	14.53	0.88 (0.85, 0.92)	$p = 0.06$	Knudson tends to classify a lower level of severity, although the trend is not statistically significant

* Numbers in parenthesis reflect 95% confidence intervals

Table 5 Effect of change of reference standard to NHANES III (Hankinson et al 1999) on the classification of severity of restrictive spirometric 'abnormality', using percent predicted FVC (n = 228)

Reference equation	Percent disagreement	Weighted Kappa	Generalized McNemar's test	Comment on comparison with NHANES III reference standard
Morris et al 1971	32.46	0.61 (0.54, 0.69)	p < 0.001	Morris classifies a lower level of severity
Kory et al 1961	21.49	0.75 (0.69, 0.81)	p < 0.001	Kory classifies a lower level of severity
Crapo et al 1981	12.28	0.87 (0.82, 0.92)	p = 1.00	Crapo classifies a similar level of severity
Knudson et al 1976	22.37	0.74 (0.67, 0.80)	p < 0.001	Knudson usually classifies a lower level of severity
Knudson et al 1983	14.49	0.84 (0.78, 0.89)	p = 0.06	Knudson tends to classify a lower level of severity, although the trend is not statistically significant

* Numbers in parenthesis reflect 95% confidence intervals

1976; Morris et al 1971, 1973), the Crapo standard by 19% (Crapo et al 1981), the Knudson 1983 standard by 17% (Knudson et al 1983), the Kory standard by 5%, (Kory et al 1961) and other reference standards (such as the Knudson 1976 standard; Knudson et al 1976) by 10% of the programs. Subsequently, data obtained from the NHANES III study, that included a random sampling of 7,429 healthy subjects from 81 counties across the US (Hankinson et al 1999), was used to develop regression equations as well. Characteristics of these various reference standards are summarized in Table 6.

The 2005 ATS/ERS guidelines also recommended the use of the NHANES III reference standard in the United States, although it suggested that other standards may be used if there were valid reasons for that choice (Pellegrino et al 2005). The advantages of the NHANES III standard include the large size and random selection of the population studied, across a large age range, with nearly equal numbers of Caucasian-Americans, African-Americans, and Mexican-Americans, rigorous quality control measures, and statistically sound coefficients for the lower limit of normal values. Further, the universal use of a nation-wide equation is likely to decrease the inter-laboratory variability in interpreting spirometric values, a challenge for the geographically mobile American population (Sharma 1995).

Use of alternative reference standards may however, be valid in certain clinical settings that may currently 'mandate' the use of a specific standard. For instance, the Knudson 1976 Standard (Knudson et al 1976) is currently used by cotton and other industries for spirometry testing done for medical surveillance of workers in the occupational setting (Occupational Safety and Health Administration 1978, 1985). The disagreement in classification of severity of spirometric 'abnormality' using NHANES III instead of the Knudson 1976 Standard (Knudson et al 1976) in the occupational setting may range as high as 21.5%, as shown in Table 5. Similarly, the Black Lung Benefits Program for coal miners (Employment Standards Administration 2000) currently employs only the Knudson 1983 standards (Knudson et al 1983) and the American Medical Guides to the Evaluation of Permanent Impairment (American Medical Association 2001) currently uses only the prediction equations by Crapo et al (Crapo et al 1981). Although it is likely that some of these programs may switch to the NHANES III reference standard in the future, health care providers need to be mindful that its current use may result in significant disagreements regarding the presence and severity of spirometric 'abnormality' under such clinical settings.

The reported disagreement between NHANES III and various other reference standards in interpretation of spirometry in this study may relate to either biological variation or analytical imprecision. Biological variation may have been introduced by the fact that the population studied by Morris (Morris et al 1971) were relatively unexposed to significant urban air pollution or cigarette smoke, were at low altitude, and were largely volunteers from the Church of Jesus Christ of the Latter-day Saints (Mormon) of diverse northern and middle European background but possibly not representative of white Europeans or Americans as a whole. The only other population involving the Mormon religious sect was studied by Crapo et al (1981) but at a higher altitude of 1,400 ms. and in urban areas. Thus, ancestral background, altitude, and rural residence may contribute to biological variation that may potentially cause lack of agreement with the NHANES III reference population. Further, prediction equations based on cross-sectional analyses may not be

Table 6 Summary of characteristics of various reference standards

Reference equation and date of publication	Type of Instrument and ability to meet ATS specifications	Time zero technique	Population characteristic	Method of screening
NHANES III-1999 Hankinson et al 1999	Dry rolling-seal spirometer with a digital shaft encoder, met ATS specifications	Back extrapolation	Nonsmoker Caucasians, African Americans, and Mexican Americans	History
Morris et al 1971	Stead-Wells, met ATS specifications	Kory technique	988 rural non-Hispanic white nonsmokers,largely of Mormon sect, in an area <150 ms. altitude, free of air pollution	History
Kory et al 1961	Collins 13.5 L metal bell, met ATS specifications	Kory technique	Hospital employees, patients, medical students, doctors in urban locale	History, physical examination, chest radiographs, and electrocardiogram
Crapo et al 1981	Collins 13.5 L metal bell, met ATS specifications	Back extrapolation	Urban, non-Hispanic white nonsmokers, largely of Mormon sect, at an altitude of 1,400 ms.	History, physical examination, and chest radiographs
Knudson et al 1976	Pneumotachygraph, met ATS specifications	Back extrapolation	746 nonsmoker, non-Mexican white population, without pregnant women, at 730 ms. altitude	History
Knudson et al 1983	Pneumotachygraph, met ATS specifications	Back extrapolation	697 nonsmoking non-Mexican white population	History

predictive of longitudinal changes (Glindmeyer et al 1982). The possibility of secular trends in improving lung function among the US birth cohorts may be a reason why the best reference standards of yesterday may differ with the more recent NHANES III reference standard.

Further, analytical imprecision may result from a difference in the technique used. Morris calculated FVC and FEV_1 using the Kory technique (Kory et al 1961) rather than the back extrapolation technique now recommended by the ATS/ERS (Miller et al 2005). The average FEV_1 calculated with the back extrapolation technique exceeds that calculated with the Kory technique by 179 ml (Smith and Gaensler 1975). This may explain why Morris classifies a lower level of severity for both obstructive and restrictive spirometric 'abnormalities', when compared to the NHANES III reference standard. Further, Crapo et al (1981) showed that their study produced predicted values for FVC and FEV_1 that were almost identical to those predicted by Morris et al when the data from the Morris study were modified to be compatible with the back extrapolation technique recommended by the ATS/ERS (Miller et al 2005). Crapo et al used the single curve with the largest sum of FVC and FEV_1 and not the ATS recommended largest values from separate curves, if needed. This may result in a reduction of about 50 ml in the predicted value of FVC by the Crapo reference standard. Knudson et al used an older pneumotachograph spirometer for his reference standards that may have terminated the maneuver prematurely, resulting in lower mean FVC values (Knudson et al 1983, 1976). The study with the highest predicted FVC and FEV_1 values, NHANES III, had extensive quality control and subjects performed at least 5 FVC maneuvers – likely explaining the slightly larger mean FVC and FEV_1 values (Hankinson et al 1999). Further, if a laboratory uses the NHANES III reference values and does not emphasize deep inhalations with sufficient expiratory times, a larger number of their patients may falsely appear to have a restrictive lung disease pattern.

The strength of this study is that it better quantifies the differences with respect to spirometric interpretation, between the NHANES III and other reference standards used in the United States. The results of this study are therefore, of practical significance to an American treating physician. This study however, has several limitations. The NHANES III reference standard was used to define 'abnormality' in this study, instead of an extensive clinical work-up of symptoms, pulmonary function, and radiographic testing. The study results depend upon the prevalence of disease in the study population, and will therefore change if the population studied differs in disease prevalence. Further, the study does not use the vital capacity measure instead of FVC, as is recommended by the 2005 ATS/ERS guidelines (Pellegrino et al 2005).

Summary

The change of spirometric reference standard to NHANES III, as is recommended by the ATS/ERS guidelines (Pellegrino et al 2005), by pulmonary function testing laboratories across the United States, may result in varying interpretations of the presence and level of severity of 'abnormality' (Rosenfeld et al 2001). This difference, in turn, may result in differences in clinical follow-up and prognosis, different conclusions in longitudinal studies, and influence eligibility criteria for clinical interventions and for research studies. Particular caution must be exercised if the Crapo reference standard (Crapo et al 1981) was originally used to rate the presence of spirometric 'abnormalities' or if the Morris reference standard was originally used to rate its' 'severity' (Morris et al 1971).

Acknowledgments

The authors would like to thank William S. Beckett, M.D., M.P.H., Professor of Medicine and Environmental Medicine, University of Rochester School of Medicine and Dentistry, Rochester, NY, and Mark Schuyler, M.D., Professor of Medicine, University of New Mexico Health Sciences Center, Albuquerque, NM, for their careful review and editorial changes.

Institution at which the work was performed: Southern Illinois University School of Medicine, Springfield, IL

Grant support: University of New Mexico General Clinical Research Center grant number M01-RR-00997 (AS, CQ)

Conflict of interest

None of the authors have any conflict of interest, including any financial conflict of interest (such as employment, consultancy, stock ownership, honoraria and paid expert testimony) as well as other forms of conflict of interest, including personal, academic and intellectual issues.

References

Altman DG. 1991. *Practical statistics for medical research*, Chapman and Hall.

American Thoracic Society. 1991. Lung function testing: Selection of reference values and interpretative strategies. *Am Rev Respir Dis*, 144:1202–18.

American Thoracic Society. 1987. Statement of the standardization of spirometry update. 1987. *Am Rev Respir Dis*, 136:1285–98.

American Thoracic Society. 1995. Standardization of spirometry update 1994. *Am J Respir Crit Care Med*, 152:1107–36.

Baur X, Isringhausen-bley S, Degens P. 1999. Comparison of lung-function reference values. *Int Arch Occup Environ Health*, 72:69–83.

Cocchiarell AL, Anderson GBJ (Eds.) 2001. Respiratory system. American Medical Association: Guides to the evaluation of permanent impairment. Chicago, IL, AMA Pr.

Crapo RO, Morris AH, Gardner RM. 1981. Reference spirometric values using techniques and equipment that meet ATS recommendations. *Am Rev Respir Dis*, 123:659–64.

Employment Standards Administration. 2000. Regulations implementing the federal coal mine and safety act of 1969, as amended. *20 C.F.R. Parts, 718, 722, 725, 726, 727*.

Ghio AJ, Crapo RO, Elliott CG. 1990. Reference equations used to predict pulmonary function. Survey at institutions with respiratory disease training programs in the United States and Canada. *Chest*, 97:400–3.

Glindmeyer HW, Diem JE, Jones RN, et al. 1982. Noncomparability of longitudinally and cross-sectionally determined annual change in spirometry. *Am Rev Respir Dis*, 125:544–8.

Hankinson JL, Odencrantz JR, Fedan KB. 1999. Spirometric reference values from a sample of the general US population. *Am J Respir Crit Care Med*, 159:179–187.

Knudson RJ, Lebowitz MD, Holberg CJ, et al. 1983. Changes in the normal maximal expiratory flow-volume curve with growth and aging. *Am Rev Respir Dis*, 127:725–34.

Knudson RJ, Slatin RC, Lebowitz MD, et al. 1976. The maximal expiratory flow-volume curve. Normal standards, variability, and effects of age. *Am Rev Respir Dis*, 113:587–600.

Kory RC, Callahan R, Boren HG, et al. 1961. The veterans administration-army cooperative study of pulmonary function. I. Clinical spirometry in normal men. *Am J Med*, 30:243–58.

Miller MR, Hankinson J, Brusasco V, et al. 2005. Standardisation of spirometry. *Eur Respir J*, 26:319–38.

Morris JF. 1976. Spirometry in the evaluation of pulmonary function. *West J Med*, 125:110–8.

Morris JF, Koski A, Johnson LC. 1971. Spirometric standards for healthy nonsmoking adults. *Am Rev Respir Dis*, 103:57–67.

Morris JF, Temple WP, Koski A. 1973. Normal values for the ratio of one-second forced expiratory volume to forced vital capacity. *Am Rev Respir Dis*, 108:1000–3.

Occupational Safety and Health Administration. 1978. Occupational exposures to cotton dust, final mandatory occupational health and safety standards. *27351*.

Occupational Safety and Health Administration. 1985. Occupational exposures to cotton dust, final rule. *50 Fed. Reg*.

Pellegrino R, Viegi G, Brusasco V, et al. 2005. Interpretative strategies for lung function tests. *Eur Respir J*, 26:948–68.

Rosenfeld M, Pepe MS, Longton G, et al. 2001. Effect of choice of reference equation on analysis of pulmonary function in cystic fibrosis patients. *Pediatr Pulmonol*, 31:227–37.

Sharma HL. 1995. Geographical mobility and mobility expectancy: trends in the United States of America, 1956–1987. *Genus*, 51:133–46.

Smith AA, Gaensler EA. 1975. Timing of forced expiratory volume in one second. *Am Rev Respir Dis*, 112:882–5.

Subbarao P, Lebecque P, Corey M, et al. 2004. Comparison of spirometric reference values. *Pediatr Pulmonol*, 37:515–22.