Volume 2 • Number 3 • 2007

Brendan G Cooper

Consultant Clinical Scientist, Honorary Chair, ERS Group 9.01; Lung Investigation Unit, 1st Floor Nuffield House, Queen Elizabeth Site, Metchley Park Road, University Hospital Birmingham NHS Foundation Trust, B15 2TH, UK

Reference values in lung function testing: All for one and one for all?

Lung function tests are used in clinical practice to detect "abnormality" against a background of "normal" function. "Abnormality" may be determined either by (i) detection of some discrete abnormality (eg, the shape of the flow volume loop) or (ii) by the use of reference values to show a relative change in values (eg, forced expiratory volume in one second [FEV_1] = 30% of predicted) or (iii) relative differences in several values (eg, obstructive spirometry but relatively raised lung volumes). Lung function tests are also used to longitudinally monitor therapy or disease progression.

Using reference values to produce "percent of predicted" measures can lead to serious misinterpretation of results and subsequent inappropriate diagnosis and treatment. It has been recognized in Europe and North America for over 20 years (Quanjer et al 1983, 1993; Pellegrino et al 2005) that using "percent of predicted" is scientifically unsound and should therefore be replaced by or used alongside "standardized residuals" which are a more scientifically robust way to express results. Knowing that a patient's FEV_1 value is below 1.64 standardized residuals gives more confidence when interpreting a respiratory disorder. Taken together with the "percent of predicted" values the two figures can help to indicate the "severity" of abnormality without falling into the trap of the reference equations producing "false positives", particularly in elderly patients (Miller 2007).

Traditionally the AARC/ATS guidelines have recommended that utilization of reference values from a variety of populations should be verified in local healthy nonsmoking subjects to determine their "local best fit" reference values. In Europe, the adoption of a generic and averaged "euro-lung" set of reference values (Quanjer et al 1983. 1993) has been recommended with some ethnic correction factors for non-Caucasian subjects. Clearly neither of these methods is ideal or without significant errors. Whilst these ethnic correction factors were probably financially expedient in the 1980s, it is clearly unacceptable to continue this practice today, and more reference value studies on most of the major races populating Europe needs undertaking.

The greatest problem for introducing new reference values is deciding that they are "better" than previously used values. "Better" usually means that the normal distribution of the values is "tightened" by the improved selection of control subjects criteria, the use of improved quality spirometry (performance or equipment) or a greater number of subjects being included. The dilemma is deciding which set of reference values is "the gold standard" and whether older values should be replaced.

In this issue Sood and colleagues (2007) compare new reference values (NHANES III) (Hankinson et al 1999) from a population of non-Hispanic Caucasians living in the USA with the traditional values used by AARC/ATS for over 30 years (Kory et al 1961; Knudson et al 1976, 1983; Morris et al 1971, 1973; Morris 1976; Crapo et al 1982). Data from two populations of restrictive and obstructive patients are used to highlight the lack of agreement between sets of reference values in detecting "abnormality", but good agreement when it comes to "severity" of disease. The authors cite the new ATS/ERS recommendation (Pellegrino et al 2005) as the driver to adopt NHANES III values, and highlight the consequences which may become costly in terms of confusion, misdiagnosis, and the practicalities of respiratory specialists and equipment manufacturers adopting the new reference values.

The study begs the question whether the European (ECSC) reference values (Quanjer et al 1983, 1993) should themselves be updated for European populations.

There needs to be an examination of whether the costs of developing what may amount to perhaps a subtle "intellectual" change will actually make any significant clinical difference when there are far more fundamental questions to answer with limited financial resources. The effect on the burden of lung disease must be economically beneficial to make such a fundamental change worthwhile.

Whilst reference values are an important issue, they need to be considered with respect to other potential errors in lung function testing. For example, in primary care spirometry the largest errors are the poor quality of the tests performed which are often a direct result of poor training and support (Kaminsky et al 2005; Raghunath et al 2006; Tinkelman et al 2006; Miravitiles et al 2007). Growing evidence indicates that equivalence in spirometry only occurs between "lab and office" when an adequate training and mentoring programme is in place in primary care (Eaton et al 1999; Upton et al 2000; Johansen 2007).

Arguably the best way to monitor lung function for both the individual and the population as a whole is to adopt the "lung health" approach to respiratory healthcare. All adults at 25 years of age (ie, when lung growth is complete) (Sherill et al 1989), should have their spirometry measured accurately by trained and experienced healthcare professionals with quality-assured spirometers. This benchmark (similar to blood pressure monitoring) can help to establish "abnormality" from "normality" in an individual against reference values. It may then be possible to distinguish which individuals have a relative reduction in lung function. These susceptible individuals can then be monitored regularly (probably each decade) to see how much capacity they have left before respiratory disease causes major disability or morbidity. The concept of "loss of functional units on the road to death" has been described eloquently by Miller and colleagues (2005a, 2005b).

The data from such a lung health study like NHANES III can also be collected nationally to provide better reference values and monitor the lung health of the nation and evaluate the intervention of therapies and behavioral changes (new guidelines, smoking cessation, etc.). The savings from detecting early disease and decreasing the burden of lung disease in the elderly should offset any financial pressures from the implementation of a lung health study. However, this has to be seen as a long term strategy in health economics and not a short term "quick fix".

Sood and colleagues (2007) have boldly addressed the issue of differences in reference values and their implications for healthcare. This article will undoubtedly generate more discussion, further consideration, and careful implementation before the NHANES III data (or any other new reference values) are widely adopted.

References

Crapo RO, Morris AH, Gardner RM. 1982. Reference values for pulmonary tissue volume, membrane diffusing capacity, and pulmonary capillary blood volume. *Bull Eur Physiopathol Respir,* 18:893–9.

Eaton T, Withy S, Garrett JE, et al. 1999. Spirometry in primary care practice: the importance of quality assurance and the impact of workshops. *Chest*, 116:416–23.

Hankinson JL, Odencrantz JR, Fedan KB. 1999. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med*, 159:179–87.

Johansen S. 2007. Spirometry in primary care in a region of Northern Norway – before and after a brief training course. *Prim Care Respir*, 16:112–14

Kaminsky DA, Marcy TW, Bachand N, et al. 2005. Knowledge and use of office spirometry for the detection of chronic obstructive pulmonary disease in primary care physicians. *Respir Care*, 50:1639–48.

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. 2007. Controversies with GOLD [online]. Accessed on September 3rd, 2007. URL: http://www.millermr.com/Info.html.

Miller MR, Pedersen OF, Lange P, et al. 2005a. Prediction of all cause mortality in a population sample is improved by using FEV_1 quotient. *Eur Respir J*, 26(Suppl. 49):469s.

Miller MR, Pedersen OF, Lange P, et al. 2005b. Prediction of respiratory disease mortality in a population sample is improved by expressing lung function impairment as FEV_1 quotient *Eur Respir J*, 26 (Suppl. 49):217s.

Miravitiles M, de la Roza C, Naberan K, et al. 2007. Use of spirometry and patterns of prescribing in COPD in primary care. *Respir Med*, 101:1753–60.

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.

Morris JF. 1976. Spirometry in the evaluation of pulmonary function. *West J Med*, 125:110–18.

Pellegrino R, Viegi G, Brusasco V, 2005 et al. ATS/ERS Task Force: Standardisation of lung function testing. Interpretative strategies for lung function tests. *Eur Respir J*, 26:948–68.

Quanjer P, Dalhuijsen A, Van Zoramen B. 1983. Standardised Lung Function Testing. Report of the working party for the European Community for Coal and Steel. *Bull Eur Physiopathol Respir*, 19(suppl. 5):1–95.

Quanjer PH, Tammeling GJ, Cotes JE, et al. 1993. 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*, 16:5–40.

Raghunath AS, Innes A, Norfolk L, et al. 2006. Difficulties in the interpretation of lung function tests in the diagnosis of asthma and chronic obstructive pulmonary disease. *J Asthma*, 43:657–60.

Sherrill DL, Camilli A, Lebowitz MD. 1989. On the temporal relationships between lung function and somatic growth. *Am Rev Respir Dis*, 140:638–44.

Sood A, Dawson BK, Henkle JQ, et al. 2007. Effect of change of reference standard to nhames iii on interpretation of spirometric abnormality. *Int J COPD*, 2:1–8

Tinkelman DG, Price DB, Nordyke RJ, et al. 2006. Misdiagnosis of COPD and asthma in primary care patients 40 years of age and over. *J Asthma*, 43:75–80.

Upton MN, Ferrell C, Bidwell C, et al. 2000. Improving the quality of spirometry in am epidememiological study: The Renfrew-Paisley (Midspan) family study. *Public Health*, 114:353–60.