

Interpretation of Spirometry: We Can Do Better Than the GOLD Standard

Spirometric testing is basic to our understanding of lung diseases and is the primary tool for the diagnosis of airflow obstruction. Doing spirometry well requires proper equipment and testing technique, along with comparison of the results to appropriate reference data and the statistically determined normal range of those data. These issues have been thoroughly discussed previously in *RESPIRATORY CARE*.^{1,2} The diagnosis of airflow obstruction is based on an abnormally low value of the ratio of forced expiratory volume in the first second (FEV_1) to vital capacity (VC) or forced vital capacity (FVC). At one time it was common to consider any FEV_1/FVC value below 70% or 75% to be abnormal, but as more sophisticated reference data were obtained,^{3–5} it became apparent that this ratio decreases progressively with age in adults, typically from average normal values in the mid-80s at age 20 to the mid- or low-70s above age 70, so a “one size fits all” criterion for airflow obstruction was not appropriate. The lower limit of normal (LLN) for the FEV_1/FVC ratio (ie, the value exceeded by 95% of normal individuals of the same gender, age, and height) is 8–10 percentage units below the predicted value in most reference data sets. There is no benefit, only confusion, engendered by reporting the observed ratio as a percent of the predicted ratio, and, to avoid further confusion with the FEV_1 itself as a percent of its predicted value, it is helpful to communicate the ratio as a decimal value (eg, 0.70).

Thus, the 1991 American Thoracic Society statement on lung-function testing⁶ included the following comments: “The FEV_1/VC ratio is the most important measurement for distinguishing an obstructive impairment. . . . Normal ranges should be based on calculated 5th percentiles. . . . In adults it is not acceptable to use a fixed FEV_1/FVC ratio as a lower limit of normal.”

This stance was confirmed in the recently updated American Thoracic Society/European Respiratory Society (ATS/ERS) Task Force statement: “An obstructive ventilatory defect. . . is defined by a reduced FEV_1/FVC ratio below the 5th percentile of the predicted value.”*⁷

This consensus has been somewhat clouded by the recent promulgation of guidelines for the diagnosis and management of chronic obstructive pulmonary disease (COPD), such as the Global Initiative for Chronic Obstructive Lung

Disease (GOLD).^{8,9} The GOLD panel defined COPD as a post-bronchodilator FEV_1/FVC ratio < 0.70 , and the same criterion has been adopted in the ATS/ERS statement on COPD¹⁰ and the British National Institute for Clinical Excellence (NICE) COPD guidelines.¹¹ A review and comparison of these COPD guidelines recently appeared in *RESPIRATORY CARE*.¹² The experts involved in developing these guidelines are certainly aware of the oversimplification involved in using this fixed value in place of a statistically based LLN, but they made a pragmatic decision in an effort to overcome the too-often-delayed recognition of this common disease and to encourage spirometry in primary-care settings worldwide.¹ However, as respiratory clinicians, particularly those involved in diagnostic laboratories, we can and should do better.

SEE THE ORIGINAL STUDY ON PAGE 737

The potential for misclassifying young individuals as false negatives and older individuals as false positives for airflow obstruction by the use of 0.70 as the LLN for FEV_1/FVC has been demonstrated previously in this journal. Figures 1 and 2 compare this fixed-percentage criterion to the age-adjusted LLN determined from the 3rd National Health and Nutrition Examination Survey (NHANES III), which is a very large reference data set derived from a United States population sample.^{1,2,13} In this issue of *RESPIRATORY CARE*, Aggarwal and colleagues¹⁴ compare the interpretations of spirometry results from over 18,000 adults in India, using an $FEV_1/FVC < 0.70$ criterion versus an LLN expressed as the lower confidence limit, equivalent to the 5th percentile, derived from an earlier local reference population. They found that 28.2% of these individuals who presented to their hospital pulmonary-function laboratory had airflow obstruction by the statistical LLN, but one quarter of these were missed by

* The updated American Thoracic Society/European Respiratory Society guidelines for “Standardization of Lung Function Testing” were published as a 5-part series in the *European Respiratory Journal*, in 2005, volume 26, issues 1–5, and are available at the American Thoracic Society Web site (<http://www.thoracic.org>).

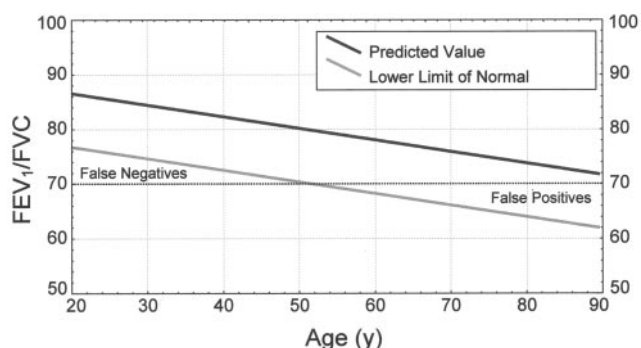


Fig. 1. The lower limit of normal for the ratio of forced expiratory volume in the first second to forced vital capacity (FEV_1/FVC) is about 10 units below the predicted value, and both decrease progressively with age in adult women, based on data from the 3rd National Health and Nutrition Examination Survey (NHANES III) prediction equations.¹³ Use of a constant value of 0.70 causes failure to recognize early airflow obstruction in young adults and falsely identifies healthy older adults as having chronic obstructive pulmonary disease. (From Reference 1, with permission.)

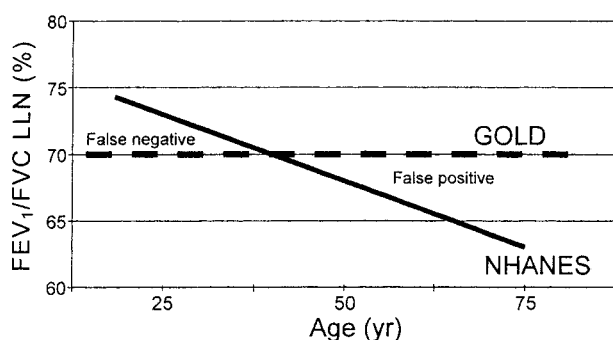


Fig. 2. The lower limit of normal for the ratio of forced expiratory volume in the first second to forced vital capacity (FEV_1/FVC) in adult men decreases progressively with age, based on the prediction equations from the 3rd National Health and Nutrition Examination Survey (NHANES III),¹³ and is slightly lower than for women of the same age. Thus, use of the criterion promulgated by the Global Initiative for Chronic Obstructive Lung Disease (GOLD) has an even greater risk of false positives for airflow obstruction in older men. (From Reference 2, with permission.)

the < 0.70 criterion. Conversely, over 10% of those deemed to have an obstructive airflow pattern by $FEV_1/FVC < 0.70$ fell within the normal confidence limits. Their Figure 3 indicates that nearly all of these false-positive diagnoses of airflow obstruction would have been in men, making the false-positive rate for men nearly 20%. The predicted LLN of the FEV_1/FVC ratio for this group of Indian men matches quite well with that for white men in the NHANES III reference data (shown in Figure 2), with values above 0.70 for young men, equal to 0.70 about age 40, and falling to ≤ 0.65 above age 70. However, the data presented for Indian women is notably different, with values 6–7% higher than those from young men and 10% higher at the older

end, whereas data from United States and European studies typically show women to have ratios no more than 2–3% higher than men, and in some studies this advantage is lost with age.^{4,5,13} Thus, the finding that the lower confidence limit for FEV_1/FVC in this study greatly exceeded 0.70 for young women and remained well above it, even in the elderly, cannot be generalized to other populations. As shown in Figure 1, the NHANES III data for white women are very similar to that for men, with LLN values near 0.75 at age 20, equaling 0.70 at age 52, and falling to 0.65 in the mid-70s age range. Separately developed equations for African-American and Mexican-American men and women from the same study give FEV_1/FVC ratios about 1–2% higher for these groups than for the white data shown.

The COPD guidelines are not meant to be applied to young individuals, and it is particularly important that an accurate LLN be used, to avoid failing to recognize airflow obstruction and undertreating asthma in this group. When applied to a population more likely to be screened for COPD, 40–60 years of age, the 0.70 cutoff is a reasonable first approximation to the true LLN for FEV_1/VC ratio, and in those with a smoking history or respiratory symptoms the pre-test probability of disease is sufficiently high that even borderline low values may more likely be associated with disease than normality. However, in the elderly population the risk of over-diagnosis by the fixed 0.70 value is much higher. In a study of elderly men and women, who were asymptomatic never-smokers, drawn from a population sample in Bergen, Norway, the likelihood of an FEV_1/FVC ratio < 0.70 was 22% for individuals age 70–79 and 50% for those above age 80,¹⁵ which emphasizes the importance of using an age-specific LLN.

A second issue addressed in the paper by Aggarwal et al¹⁴ is the use of 80% of the predicted value for the LLN for FEV_1 and FVC, which is a practice probably even more widespread than the use of a fixed FEV_1/FVC ratio, and is often extended to lung volumes and diffusing capacity as well. It has been known for many years, and reconfirmed with every new set of reference equations, that the variance of normal FEV_1 and FVC values above and below the predicted regression line is similar for large and small values, and thus the LLN remains a constant increment below the predicted value.^{16,17} This statistically accurate LLN will therefore occur at a lower percentage of a small predicted value than a high one. Like the 0.70 ratio, the 80%-of-predicted “rule of thumb” approximates the true LLN of FVC and FEV_1 (but not of all pulmonary function variables) for individuals of average age and height, but will be expected to under-diagnose abnormality in younger, taller individuals and over-diagnose abnormality in older, shorter individuals with smaller predicted values. The data from Aggarwal et al, collected in a real-world experience, are entirely consistent with these

expectations, and show that the normal range determined by confidence interval is much more likely to extend below 80%, and in some cases markedly below, than to stop above it.

So why do we—and I frequently find myself among the guilty—continue to use 80%? First, of course, it's easy, especially given the way pulmonary-function data are typically presented to us by the output of our equipment. Virtually every apparatus will readily print out a predicted value along with the measured value, and calculate a percent of predicted. Creating a column to display the properly calculated LLN may require special programming, if it is possible at all, or transferring the data to a report format separate from the testing equipment. Without this, one must keep track of confidence intervals for each test, which often differ by sex. The time is overdue to ask equipment manufacturers to provide a way to display appropriate lower limits for each test in their reporting formats.

Second, I think we often see our role as finding and documenting abnormalities, so we tend to resist lowering the LLN for FVC or FEV₁. There may be some justification for this in symptomatic patients, in whom early disease is often associated with low normal values. But if we wish to increase sensitivity in a high-risk population, it would be more rational to choose a narrower confidence interval rather than the age- and height-dependent error of the percent of predicted. When we are testing individuals with a low or unknown probability of disease (eg, at health fairs, routine physicals, or occupational screening), we should be scrupulous in using a statistically appropriate LLN, to avoid mislabeling healthy people whose values happen to be among the lower portion of a normal population distribution.

Bruce H Culver MD

Division of Pulmonary and Critical Care Medicine
Department of Medicine
University of Washington
Seattle, Washington

Correspondence: Bruce H Culver MD, Division of Pulmonary and Critical Care Medicine, University of Washington Medical Center, Box 356522, Seattle WA 98195. E-mail: bculver@u.washington.edu.

REFERENCES

1. Crapo RO, Jensen RL. Standards and interpretive issues in lung function testing. *Respir Care* 2003;48(8):764–772.
2. Enright PL, Kaminsky DA. Strategies for screening for chronic obstructive pulmonary disease. *Respir Care* 2003;48(12):1194–1201; discussion 1201–1203.
3. Morris JF, Koski A, Johnson LC. Spirometric standards for healthy nonsmoking adults. *Am Rev Respir Dis* 1971;103(1):57–67.
4. Crapo RO, Morris AH, Gardner RM. Reference spirometric values using techniques and equipment that meet ATS recommendations. *Am Rev Respir Dis* 1981;123(6):659–664.
5. Knudson RJ, Lebowitz MD, Holberg CJ, Burrows B. Changes in the normal maximal expiratory flow-volume curve with growth and aging. *Am Rev Respir Dis* 1983;127(6):725–734.
6. American Thoracic Society. Lung function testing: selection of reference values and interpretative strategies. *Am Rev Respir Dis* 1991;144(5):1202–1218.
7. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, et al. Interpretative strategies for lung function tests. *Eur Respir J* 2005;26(5):948–968.
8. Pauwels RA, Buist AS, Calverly PM, Jenkins CR, Hurd SS; GOLD Scientific Committee. Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. *Am J Respir Crit Care Med* 2001;163(5):1256–1276.
9. Fabbri LM, Hurd SS; GOLD Scientific Committee. Global strategy for the diagnosis, management and prevention of COPD: 2003 update (editorial). *Eur Respir J* 2003;22(1):1–2.
10. Celli BR, MacNee W; ATS/ERS Task Force. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J* 2004;23(6):932–946. *Erratum in: Eur Respir J* 2006;27(1):242.
11. National Collaborating Center for Chronic Conditions. Chronic obstructive pulmonary disease. National clinical guideline on management of chronic obstructive pulmonary disease in adults in primary and secondary care. *Thorax* 2004;59 Suppl 1:1–232.
12. Pierson DJ. Clinical practice guidelines for chronic obstructive pulmonary disease: a review and comparison of current resources. *Respir Care* 2006;51(3):277–288.
13. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999;159(1):179–187.
14. Aggarwal AN, Gupta D, Behera D, Jindal SK. Comparison of fixed percentage method and lower confidence limits for defining limits of normality for interpretation of spirometry. *Respir Care* 2006;51(7):737–743.
15. Hardie JA, Buist AS, Vollmer WM, Ellingsen I, Bakke PS, Morkve O. Risk of over-diagnosis of COPD in asymptomatic elderly never-smokers. *Eur Respir J* 2002;20(5):1117–1122.
16. Sobol BJ. Some cautions in the use of routine spirometry. *Arch Intern Med* 1966;118(4):335–339.
17. Miller MR, Pincock AC. Predicted values: how should we use them? (editorial). *Thorax* 1988;43(4):265–267.