Using the lower limit of normal for the FEV₁/FVC ratio reduces the misclassification of airway obstruction

Authors (The Pulmonaria Group)
1. Maureen P. Swanney, Respiratory Physiology Laboratory, Christchurch Hospital, Canterbury District Health Board, Christchurch, New Zealand.
2. Gregg Ruppel, Pulmonary Function Laboratory, Saint Louis University Hospital, St. Louis MO, USA.
3. Paul L. Enright, College of Public Health, The University of Arizona, Tucson, AZ, USA.
4. Ole F. Pedersen, Institute of Public Health, Aarhus University, Aarhus, Denmark.
5. Robert O. Crapo, Pulmonary Division, LDS Hospital and University of Utah, Salt Lake City, Utah, USA.
6. Martin R. Miller, Dept of Medicine, University Hospital Birmingham NHS Trust, Birmingham B29 6JD, UK.
7. Robert L. Jensen, Pulmonary Division, LDS Hospital and University of Utah, Salt Lake City, Utah, USA.
8. Emanuela Falaschetti, Department of Epidemiology and Public Health, University College London, London, UK.
9. Jan P. Schouten, Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.
10. John L. Hankinson, Valdosta, Georgia, USA.
12. Philip H. Quanjer, Department of Pulmonary Diseases, Erasmus Medical Centre, Erasmus University, Rotterdam, The Netherlands.

Corresponding author:
Maureen P. Swanney, Respiratory Physiology Laboratory, 4TH Floor Riverside Building, Christchurch Hospital, Private Bag 4710, Christchurch 8140, New Zealand.
E-mail: maureen.swanney@cdhb.govt.nz
Telephone: +64 3 3640924, Facsimile: +64 3 3640878

Key words: spirometry, COPD, airway obstruction, misclassification, interpretation
Word count: 2915

“The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd and its Licensees to permit this article to be published in Thorax editions and any other BMJPG products to exploit all subsidiary rights, as set out in our licence http://thorax.bmjournals.com/ifora/licence.pdf”
ABSTRACT

Rationale: The prevalence of airway obstruction varies widely with the definition used.

Objectives: To study differences in the prevalence of airway obstruction when applying four international guidelines to three population samples using four regression equations.

Methods: We collected predicted values for FEV₁/FVC and its lower limit of normal (LLN) from the literature. FEV₁/FVC from 40,646 adults (including 13,136 asymptomatic, never-smokers) aged 17-90+ years were available from American, English, and Dutch population-based surveys. The prevalence of airway obstruction was determined by the LLN for FEV₁/FVC, and by using the GOLD, ATS/ERS, or BTS guidelines, initially in the healthy subgroup and then the entire population.

Results: The LLN for FEV₁/FVC varied between prediction equations (57 available for men and 55 for women), and demonstrated marked negative age dependency. The median age at which the LLN fell below 0.70 in healthy subjects was 42 and 48 years in men and women respectively.

When applying the reference equations (Health Survey for England 1995-1996, NHANESIII, ECCS/ERS and a Dutch population study) to the selected population samples, the prevalence of airway obstruction in healthy never-smokers aged over 60 years varied for each guideline: 17-45% of men and 7-26% of women for GOLD; 0-18% of men and 0-16% of women for ATS/ERS; 0-9% of men and 0-11% of women for BTS. GOLD guidelines caused false positive rates of up to 60% when applied to entire populations.

Conclusions: Airway obstruction should be defined by FEV₁/FVC and FEV₁ being below the LLN using appropriate reference equations.

[Abstract word count = 248]
INTRODUCTION

COPD is a major public health concern as a cause of chronic morbidity and mortality [1]. COPD starts insidiously in adulthood causing a rapid decline in FEV₁. [2] Initially airway obstruction was defined as an FEV₁/FVC ratio below the lower fifth percentile of a large healthy reference group (the LLN, statistically defined lower limit of normal) [3,4,5], a widely accepted standard for interpreting physiological and biochemical measurements [6] (see Table 1).

Table 1. Criteria for assessing airway obstruction (COPD) according to various organisations, listed by year.

<table>
<thead>
<tr>
<th>Organisation</th>
<th>Year</th>
<th>Reference</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECCS</td>
<td>1983</td>
<td>3</td>
<td>FEV₁/VC or FEV₁/FVC &lt; LLN</td>
</tr>
<tr>
<td>ATS</td>
<td>1987</td>
<td>8</td>
<td>FEV₁/FVC &lt; 0.75</td>
</tr>
<tr>
<td>ATS</td>
<td>1991</td>
<td>4</td>
<td>FEV₁/FVC &lt; LLN</td>
</tr>
<tr>
<td>ECCS/ERS</td>
<td>1993</td>
<td>5</td>
<td>FEV₁/VC or FEV₁/FVC &lt; LLN</td>
</tr>
<tr>
<td>ERS</td>
<td>1995</td>
<td>7</td>
<td>FEV₁/VC &lt; 88% predicted (males) or 89% (females)</td>
</tr>
<tr>
<td>BTS</td>
<td>1997</td>
<td>9</td>
<td>FEV₁/FVC &lt; 0.70 and FEV₁ &lt; 80% predicted</td>
</tr>
<tr>
<td>NLHEP</td>
<td>2000</td>
<td>12</td>
<td>FEV₁/FVC or FEV₁/FEV₆&lt;LLN and FEV₁&lt;LLN</td>
</tr>
<tr>
<td>GOLD</td>
<td>2006</td>
<td>1</td>
<td>FEV₁/FVC &lt; 0.70 post-bronchodilator</td>
</tr>
<tr>
<td>NICE</td>
<td>2004</td>
<td>10</td>
<td>FEV₁/FVC &lt; 0.70 and FEV₁ &lt; 80% predicted</td>
</tr>
<tr>
<td>ATS/ERS</td>
<td>2004</td>
<td>11</td>
<td>FEV₁/FVC &lt; 0.70 post-bronchodilator</td>
</tr>
<tr>
<td>ATS/ERS</td>
<td>2005</td>
<td>13</td>
<td>FEV₁/VC &lt; LLN</td>
</tr>
</tbody>
</table>

The ERS modified this definition by expressing FEV₁/FVC as a percentage of predicted [7]. In an attempt to promote the early detection of chronic airway obstruction by applying a simple rule of thumb that did not require prediction equations and computer use, expert panels introduced a fixed cut-off for FEV₁/FVC. In a 1987 ATS document airway obstruction was considered to be present if FEV₁/FVC was less than 0.75. [8] In 1997 the British Thoracic Society (BTS) defined airway obstruction as a FEV₁/FVC ratio less than 0.70, followed by the GOLD, NICE, and ATS/ERS COPD guidelines. [9,1,10,11] The BTS and NICE guidelines define airway obstruction using the fixed ratio of 0.70 AND an FEV₁ below 80% predicted, while GOLD and ATS/ERS COPD guidelines only use the 0.70 fixed ratio (the FEV₁ can be in the normal range). The National Lung Health Education Program in the United States interprets an FEV₁/FVC (or FEV₁/FEV₆) AND FEV₁ below their respective LLN values as evidence of airway obstruction. [12] In 2005, an ATS/ERS pulmonary function interpretation guideline strongly recommended the use of the LLN for FEV₁/VC to define airway obstruction. [13]

The fixed 0.70 cut-off has the potential to misdiagnose cases of airway obstruction, because FEV₁/FVC has been shown to vary with age, height, and gender. In addition, no evidence has been published validating the concept that the 0.70 cut-off with a normal FEV₁ (GOLD stage 1) identifies subjects with airway disease or COPD; in fact, doubt has been expressed whether GOLD stage I represents clinically relevant disease. [14] Several reports have shown that using the fixed cut-off leads to under-diagnosis in young adults and over-diagnosis of airway obstruction in adults aged over 40 years [15-24]; however, these reports were limited because they applied single spirometry reference equations to single
population samples. By contrast, the current study has investigated the prevalence of airway obstruction using three different geographical populations using a more thorough analysis than in previously published papers. [15-24]

**The objectives of this study were:**

1. To review published equations for different ethnic groups for the LLN for FEV₁/FVC for healthy lifelong non-smokers, and compare the LLNs to the 0.70 fixed cut-off.

2. To investigate whether the various criteria of airway obstruction lead to acceptable results when applied to the healthy subset of never-smoking, asymptomatic adults from 3 large population studies.

3. To determine differences in the prevalence of airway obstruction when applying 3 popular prediction equations for the LLN for FEV₁/FVC to all adults from the 3 population samples (including smokers and asthmatics).

**METHODS**

**Objective 1: LLN for FEV₁/FVC according to worldwide literature**

We retrieved from the literature 57 spirometry reference equations for FEV₁/FVC for men and 55 for women, including previously unpublished equations derived from a large Dutch population-based study (see the “Online Data Supplement”). In some of these studies no equations were presented, but they could be derived from tables or graphs in the original publications. Details from all of the studies are available at www.spirxpert.com/GOLD.html. [25] If the LLN was not stated for the studies, it was calculated by subtracting 1.645· residual standard deviation (RSD) from the predicted mean (assuming a Gaussian distribution of the residuals). For each study we computed the age at which the LLN for FEV₁/FVC fell below 0.70, at a standing height of 175 cm in men, and 165 cm in women. In some this required extrapolation beyond the studied age range. If the LLN did not fall below 0.70 at a younger age than 80, we arbitrarily allocated it to age 80. We then computed the median ages at which the LLN for FEV₁/FVC fell below 0.70.

**Objective 2: Prevalence of obstruction in HEALTHY adults**

We used the databases of all spirometry results from three large studies: the NHANES III from the United States [26], the Health Survey for England 1995-1996 (HSE9596) [20], and a longitudinal study of two Dutch populations [27], to investigate the proportions of healthy, asymptomatic never smokers with airway obstruction according to several definitions (see Table 1). The public NHANES III dataset used the ATS-1987 standards; J Hankinson recalculated the values for the 1994 standards. Details of subject selection and methodology of the three studies have been extensively published. The healthy subset from NHANES III included 1706 men and 2924 women. The healthy subset from the HSE9596 study included 3107 men and 4195 women. For the purpose of this analysis, a cross-section of the Dutch cohorts was created by selecting one record per person to obtain a relatively even age distribution of adults. Reference values for FEV₁/FVC for the Dutch population (see the Online Data Supplement for the equations) were derived from the healthy subset of 321 men and 883 women who were never-smokers and who had been free of respiratory symptoms, as judged by answers to the MRC/ECCS questionnaire administered by trained interviewers. [28] No post-bronchodilator results were available in any of these three studies.
Objective 3: Prevalence of obstruction in ALL adults
Using all subjects that were tested from these three databases (including smokers and those with respiratory symptoms), we calculated the prevalence of airway obstruction.

For each reference set, a true positive result, for the presence of airway obstruction, was defined when the FEV₁/FVC was below the LLN, and a true negative when the FEV₁/FVC was above or equal to the LLN. A false positive was defined when the FEV₁/FVC was less than 0.70 but above or equal to the LLN, and a false negative when the FEV₁/FVC was below the LLN and above or equal to 0.70.

This analysis was performed by computing the LLN using the regression equations appropriate for the three populations: Falaschetti et al. [20] for HSE9596, Hankinson et al. for NHANESIII [26], and the equations (see details in the Online Data Supplement) for the Dutch data. Since the ECCS/ERS prediction equations [5] are widely used, particularly in Europe, they were additionally applied to the English, American, and Dutch databases to gauge the effect of applying different prediction equations to the same data.

Statistical analyses
All analyses were performed using SPSS 14.0 for Windows (SPSS Inc., Chicago, Ill., USA). Percentiles in the Dutch study were estimated with quantile regression models (see Online Data Supplement).

RESULTS

Objective 1: LLN for FEV₁/FVC according to worldwide literature
The LLN for FEV₁/FVC varied widely between studies but invariably fell with age; it fell below 0.70 at various ages from all but 5 of the reference studies for adult men and 9 for women (see Figure 1). Standing height was also an independent predictor of FEV₁/FVC in most studies. See the detailed results online [25] or the On-line data Supplement. In 57 prediction equations for FEV₁/FVC for men, and 55 for women, the lower limit of normal for FEV₁/FVC was less than 0.70 at a median age of 42 and 48 years, respectively.

Objective 2: Prevalence of obstruction in HEALTHY adults
The proportion of subjects who had airway obstruction according to the 0.70 fixed cut-off was small in younger age groups, but increased up to values as high as 45% (male Dutch population) in those aged 60 or more (see Figure 2). The percentages of healthy subjects deemed to have airway obstruction according to ERS, BTS and GOLD guidelines are also shown in Table 2.

Differences are, at least in part, due to systematic differences in FEV₁/FVC between populations. After accounting for age and height, FEV₁/FVC in healthy Dutch women is on average 0.01 (confidence interval CI 0.004-0.016) below that of American women, and 0.029 (CI 0.024-0.034) below that of English women; corresponding figures for men are 0.011 (CI 0.002-0.020) and 0.024 (0.018-0.030), respectively.

In all three data sets, the 0.70 fixed cut-off resulted in an increase in the percentage of healthy subjects being labelled as having airway obstruction. Further, the percentage of healthy subjects classified with obstruction increased with age (see Figure 2 and Table 2): 17% to 45% of healthy men, and 7% to 26% of healthy women over 60 years. The prevalence of airway obstruction using the 0.70 fixed cut-point in the healthy Dutch population was systematically higher than in the other two populations.
Table 2. Relative frequency (%) by age group of observations of FEV₁/FVC below 0.70 (GOLD), below 88% predicted in women and 89% predicted in men (ERS), below 0.70 and FEV₁ < 80% predicted (BTS), or below the lower limit of normal (LLN). Data relate exclusively to asymptomatic, never-smokers. Predicted values were derived from equations that fitted each data set [20, 26, and the Dutch equations derived in this study].

<table>
<thead>
<tr>
<th>Age</th>
<th>Number of Subjects</th>
<th>GOLD FEV₁/FVC &lt;0.7</th>
<th>ERS FEV₁/FVC &lt; 88% pred.</th>
<th>BTS FEV₁/FVC &lt; 0.70 &amp; FEV₁ &lt; 80% pred.</th>
<th>Proposed FEV₁/FVC &lt; LLN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dutch</td>
<td>UK*</td>
<td>USA†</td>
<td>Dutch</td>
<td>UK*</td>
</tr>
<tr>
<td>16-29</td>
<td>110</td>
<td>709</td>
<td>804</td>
<td>4.6</td>
<td>2.1</td>
</tr>
<tr>
<td>30-39</td>
<td>108</td>
<td>675</td>
<td>372</td>
<td>9.3</td>
<td>2.8</td>
</tr>
<tr>
<td>40-49</td>
<td>58</td>
<td>448</td>
<td>211</td>
<td>6.9</td>
<td>3.4</td>
</tr>
<tr>
<td>50-59</td>
<td>28</td>
<td>312</td>
<td>96</td>
<td>21.4</td>
<td>10.3</td>
</tr>
<tr>
<td>60-69</td>
<td>11</td>
<td>213</td>
<td>127</td>
<td>45.4</td>
<td>17.8</td>
</tr>
<tr>
<td>≥70</td>
<td>6</td>
<td>122</td>
<td>96</td>
<td>16.7</td>
<td>18.1</td>
</tr>
<tr>
<td>Total</td>
<td>321</td>
<td>2479</td>
<td>1706</td>
<td>4.4</td>
<td>5.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>Number of Subjects</th>
<th>GOLD FEV₁/FVC &lt;0.7</th>
<th>ERS FEV₁/FVC &lt; 88% pred.</th>
<th>BTS FEV₁/FVC &lt; 0.70 &amp; FEV₁ &lt; 80% pred.</th>
<th>Proposed FEV₁/FVC &lt; LLN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dutch</td>
<td>UK*</td>
<td>USA†</td>
<td>Dutch</td>
<td>UK*</td>
</tr>
<tr>
<td>16-29</td>
<td>135</td>
<td>856</td>
<td>1055</td>
<td>0</td>
<td>0.4</td>
</tr>
<tr>
<td>30-39</td>
<td>195</td>
<td>863</td>
<td>1058</td>
<td>6.7</td>
<td>1.4</td>
</tr>
<tr>
<td>40-49</td>
<td>200</td>
<td>671</td>
<td>871</td>
<td>6.5</td>
<td>3.9</td>
</tr>
<tr>
<td>50-59</td>
<td>220</td>
<td>454</td>
<td>674</td>
<td>15.4</td>
<td>3.1</td>
</tr>
<tr>
<td>60-69</td>
<td>115</td>
<td>410</td>
<td>525</td>
<td>26.1</td>
<td>7.1</td>
</tr>
<tr>
<td>≥70</td>
<td>18</td>
<td>302</td>
<td>285</td>
<td>5.6</td>
<td>13.3</td>
</tr>
<tr>
<td>Total</td>
<td>883</td>
<td>3556</td>
<td>4468</td>
<td>5.3</td>
<td>6.2</td>
</tr>
</tbody>
</table>

Footnotes: * UK = Data from the Health Survey for England 1995-1996. † USA = data from NHANESIII
The overall airway obstruction rate using the LLN was slightly higher than the expected 5% in the NHANES III data (5.5%) and in the HSE study of women (6.2%) (see Table 2). In Dutch males aged 50-69 years, the predicted value for LLN did not appear to fit very well, which may be due to the limited number of men (39) in this age range. The prevalence of airway obstruction also increased with age using the ERS criteria. The ERS LLN for FEV₁/FVC (<88% of predicted for males and 89% of predicted for females) results in an increased prevalence of obstruction in all age groups, particularly in women. The BTS criteria differ from the GOLD guidelines in that not only must the FEV₁/FVC be below 0.70, but FEV₁ must also be less than 80% predicted. While an age-related trend in the prevalence of airway obstruction was seen with the BTS criteria, it remained below 5% except in the older subjects.

Objective 3: Prevalence of obstruction in ALL adults

We expressed the false positive results as a percentage of all positive test results (true positive plus false positive). This represented the percentage of subjects who were being diagnosed as having airway obstruction because their FEV₁/FVC was below the fixed 0.70 cut-point, but above the LLN.

There was a steep increase with advancing age in the percentage of subjects who were erroneously identified as having airway obstruction (see Figure 3), reflecting the widening gap between the LLN and the 0.70 fixed cut-point (see Figure 1). The ascending lines run largely in parallel, demonstrating that the age dependency of the LLN differs little between the prediction equations. The USA data were shifted to the left by about 7 years compared to the English data, the Dutch data were shifted 3 decades to the left.

Applying ECCS/ERS reference values [5] for the LLN for FEV₁/FVC brought the English, American and Dutch data together (see Figure 3), because for the same age and gender there was no longer a difference in the predicted LLN, any remaining differences being related to differences in the distribution of age and male/female ratio in the different populations, and of course the true prevalence of airway obstruction. There was a 10-15 year shift of the curves to the left because the ECCS/ERS prediction equation led to a lower LLN for FEV₁/FVC than equations due to Falaschetti [20] in English data and Hankinson [26] in American data, but higher than the Dutch equations applied to the Dutch population.

Applying internal reference values to the Caucasian American, English and Dutch adults led to small differences in the overall prevalence of airway obstruction (see Table 3), 12% in the Netherlands, 14% in England and 16% in the USA. Except for the English population, these percentages were systematically smaller than those obtained by using the 0.70 fixed cut-off. This was fortuitous, because the false positive test results were balanced by the false negative ones in these wide age ranges. The level of FEV₁/FVC was lower in the Dutch population than in the other populations, so applying the fixed cut-off nearly doubled the prevalence of airway obstruction in those above age 50 yr, as did applying regression equations due to Hankinson [26] and Falaschetti. [20] The English and Dutch populations did not have a sample of Black or Hispanic participants, so the comparisons in Table 3 and in Figure 3 were only done for Caucasians. The predicted values for the English and American populations did not differ by much; hence their effect on the prevalence of airway obstruction in these populations was rather small. Adding the requirement for a low FEV₁ (<80% predicted) (BTS and NICE) reduced the prevalence of airway obstruction by approximately a half.
Table 3. The prevalence (as a percentage) of airway obstruction in three large population-based samples using three reference equations and different definitions of airway obstruction applied to each population: the GOLD criterion; FEV$_1$/FVC below the LLN for three reference equations; and the last three columns apply the FEV$_1$%pred < 80% (BTS and NICE) criteria together with the FEV$_1$/FVC < LLN criteria for each reference equation.

<table>
<thead>
<tr>
<th>Age</th>
<th>N</th>
<th>GOLD</th>
<th>NHANES III</th>
<th>HSE</th>
<th>NL</th>
<th>NHANES III</th>
<th>HSE</th>
<th>NL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dutch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;50</td>
<td>2,965</td>
<td>12.2</td>
<td>17.2</td>
<td>20.7</td>
<td>11.1</td>
<td>4.6</td>
<td>5.2</td>
<td>4.4</td>
</tr>
<tr>
<td>&gt;50</td>
<td>1,592</td>
<td>36.1</td>
<td>26.6</td>
<td>29.3</td>
<td>14.6</td>
<td>15.5</td>
<td>14.9</td>
<td>9.9</td>
</tr>
<tr>
<td>All</td>
<td>4,557</td>
<td>20.5</td>
<td>20.4</td>
<td>23.7</td>
<td>12.3</td>
<td>8.4</td>
<td>8.6</td>
<td>6.3</td>
</tr>
<tr>
<td>English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;50</td>
<td>14,506</td>
<td>6</td>
<td>9.2</td>
<td>11.8</td>
<td>6.7</td>
<td>2.4</td>
<td>2.4</td>
<td>2.6</td>
</tr>
<tr>
<td>&gt;50</td>
<td>10,098</td>
<td>24</td>
<td>15.7</td>
<td>17.3</td>
<td>9.1</td>
<td>14.2</td>
<td>13.5</td>
<td>6.7</td>
</tr>
<tr>
<td>All</td>
<td>24,604</td>
<td>13.4</td>
<td>11.9</td>
<td>14.1</td>
<td>7.7</td>
<td>7.2</td>
<td>6.9</td>
<td>4.3</td>
</tr>
<tr>
<td>American</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;50</td>
<td>3,196</td>
<td>6.3</td>
<td>10.2</td>
<td>12.4</td>
<td>6.9</td>
<td>2.7</td>
<td>2.8</td>
<td>4.4</td>
</tr>
<tr>
<td>&gt;50</td>
<td>3,633</td>
<td>35.5</td>
<td>20.2</td>
<td>21.4</td>
<td>10.2</td>
<td>16.8</td>
<td>15.1</td>
<td>7.8</td>
</tr>
<tr>
<td>All</td>
<td>6,829</td>
<td>21.8</td>
<td>15.5</td>
<td>17.2</td>
<td>8.7</td>
<td>10.2</td>
<td>9.4</td>
<td>5.8</td>
</tr>
</tbody>
</table>

Footnotes: The most appropriate reference equations are internal to the study (indicated by the bold numbers). Analysis in Americans was limited to Caucasian subjects. Reference equation sources: NHANES III [26]; HSE [20]; NL = Dutch reference equations (see Appendix in the Online Data Supplement).
DISCUSSION

This study demonstrates the large differences in the prevalence of airway obstruction defined by spirometry that arise from using different definitions and/or prediction equations. These differences overlay variation in the true prevalence of airway obstruction between populations, and differences in the level of FEV₁/FVC. However, when prediction equations are used that fit healthy lifelong non-smokers in the relevant population, there is fair agreement in the prevalence of airway obstruction, as demonstrated in this study using three large population samples (prevalence 12.3-15.5%, see Table 3).

Large differences in the prevalence of airway obstruction due to seemingly minor variants in defining airway obstruction were recognized many years ago [29] and recently confirmed. [15, 16, 23, 30, 31] The choice of these definitions should be dictated by an examination of evidence that demonstrates minimal disease misclassification in population-based samples rather than a non-validated ‘rule of thumb’ [16]. While a fixed cut-off seems easy to apply it comes at the expense of erroneous classification of disease.

Guidelines based only on the 0.70 fixed cut-off broaden the definition of COPD and cause a situation comparable to that documented recently for several other diseases, where changes in abnormality threshold values led to stunning increases in prevalence within the population. [32, 33] As shown in Table 2 and Figure 2, using a fixed cut-off for FEV₁/FVC turns a considerable proportion of healthy non-smokers into patients. The proportion of subjects over age 50 in a general population that is incorrectly identified as having airway obstruction is alarming, going up to as high as 68 percent (Figure 3). Thus a declining FEV₁/FVC, a normal phenomenon with advancing age, is equated to disease, without clinical evidence.

A limitation of our study is that post-bronchodilator values for the FEV₁/FVC were not obtained, as required by the GOLD guidelines. When applying GOLD guidelines in population-based samples of adults, one-third of those with airway obstruction pre-BD did not have evidence of obstruction after a bronchodilator was administered. [33] Hence the shortcomings of the use of a fixed ratio can only partly be remedied by bronchodilator effects. Indeed, the GOLD committee recently acknowledged that “Using the lower limit of normal (LLN) values for FEV₁/FVC, that are based on the normal distribution and classify the bottom 5% of the healthy population as abnormal, is one way to minimize the potential misclassification.” [1]

The ATS/ERS, BTS, NICE, and GOLD guidelines define airway obstruction for the purpose of diagnosing COPD (in patients with respiratory symptoms seeking help from physicians), but spirometry is also frequently used for detecting COPD by public screening outside of medical care settings, to include/exclude subjects for research, or for case-finding in primary care offices. The high false positive rates reported in this study and others when following the current GOLD guidelines can be reduced by 1) only testing people with a high pre-test probability of COPD (as determined by using a short questionnaire) [34]; 2) using the statistical LLN for FEV₁/FVC instead of the 0.70 fixed cut-off; 3) requiring that the FEV₁ also be low (below the LLN) as recommended by NLHEP [12], and akin to recommendations by BTS [9] and NICE [10]; and 4) by confirming that airway obstruction persists after inhaling a bronchodilator.

Reference equations that provide predictive normal values for spirometric indices are necessary for meaningful clinical interpretation. The normal range for most physiological and
biochemical variables is traditionally defined so that 5% of healthy subjects fall below the LLN, and this is entirely appropriate for the FEV₁/FVC. Until there is evidence to the contrary, the LLNs should be derived from prediction equations that have been shown to fit the population being tested. [13]

Our current list of international spirometry reference studies for adults includes 57 publications; the details of these studies as well as the reference equations can be found at www.spirxpert.com/GOLD.html in the help file associated with the Pulmonaria software [25], which may be freely downloaded. As can be seen by using the software, the LLN for FEV₁/FVC varies not only by age, height and gender, but also by racial or ethnic group. Therefore, each country should carefully select the set of spirometry reference equations which best matches the methodology, ethnicity and age range of their population.

We need to be careful when identifying someone below an arbitrary cut-off as having COPD because the medical profession can offer limited help. [36] While spirometry is an established test for diagnosing airway obstruction it is not necessarily the only or the best tool available. An obvious limitation of spirometry is that it cannot identify disease at an early stage. Other (bio-)markers of disease are urgently needed. [37]

In conclusion we recommend that spirometry for COPD case-finding should only be done for patients with symptoms or prior exposure to noxious substances who therefore have a high pre-test probability of COPD. Airway obstruction should be defined by a post-bronchodilator FEV₁/FVC and FEV₁ below the LLN derived from appropriate reference equations (having taken into account age, height, gender, and ethnicity). Defenders of a fixed ratio for FEV₁/FVC claim that the simplicity of this measure over-shadows its disadvantage. However, modern technology can easily provide the lower limit of normal, so why continue to accept false positive and negative lung function tests in the diagnosis of COPD?

FIGURE LEGENDS

Figure 1. The LLN for FEV₁/FVC falls with age in healthy adults. Reference values for the lower limit of the normal range from 57 studies of healthy men and women from around the world. Data includes the age range between 20 – 80 years. The solid horizontal line indicates the 0.70 fixed cut-off.

Figure 2. Percentage of healthy adults (asymptomatic never-smokers) whose FEV₁/FVC was less than 0.70. [Objective 2 results]

Figure 3. False positive test results expressed as a percentage of all positive results for airway obstruction in 3 large population-based samples of adults, including smokers and subjects with respiratory symptoms. [Objective 3 results]

Footnote to figure 2: NL = Dutch data and internal reference values (see Appendix in the Online Data Supplement); HSE = Health Survey for England 1995-1996 using internal reference values [20]; American data for Caucasians, using internal reference values (NHANES). [26]

Footnote to figure 3: NL = Dutch data and internal reference values (see Appendix in the Online Data Supplement); HSE = Health Survey for England 1995-1996 using internal reference values [20]; American data for Caucasians, using internal reference values (NHANES). [26] ECCS/ERS prediction equations [5] for the LLN of FEV₁/FVC were also applied to the three population samples (dashed lines).
Funding: none
REFERENCES

8 American Thoracic Society. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. Am Rev Respir Dis 1987; 136: 225-244.
14 Enright PL. GOLD stage I is not a COPD risk factor. Letter to the Editor. Thorax 2007;62:1107-1109
19 Medbø A, Melbye H. Lung function testing in the elderly - Can we still use the FEV1/FVC...
<70% as a criterion for COPD? Respir Med 2007; 101: 1097-1053


21 Culver HB. Interpretation of spirometry: we can do better than the GOLD standard. Chest 2006; 51: 719-721.


33 Moynihan R, Cassels A. Selling sickness: How the world’s biggest pharmaceutical companies are turning us all into patients. 2005. Nation Books, New York.


Appendix (for Online Data Supplement)

Reference values from the subset of healthy adults from two Dutch population-based cohorts

Population
The data were obtained in field surveys conducted at 3-yr intervals between 1965 and 1990 in a rural area (Vlagtwedde) and in an industrial town (Vlaardingen) in the Netherlands. Each survey was completed in 1 week in October. The Vlagtwedde sample was a dynamic cohort. In every field survey all men and women born between 1921 and 1952 and living in a specified area were invited to take part. The Vlaardingen sample was a fixed cohort selected from the municipal register. In 1965 a random sample of men and women born between 1901 and 1925 was selected, and in 1969 an additional sample was selected from those born between 1930 and 1954.

Measurements
Anthropometric and lung function measurements were obtained in a standardized way by trained operators. BTPS measurements of FEV₁ and FVC were derived from maximal expiratory flow-volume curves, using pneumotachography. Measurements and procedures complied with recommendations of the European Community for Coal and Steel ¹. Calibration was performed approximately every hour.


Data selection
For each person up to 5 measurements were available. From amongst persons who had never smoked and had been free of cardio-respiratory symptoms one record per person was selected using an algorithm designed to obtain the most even age distribution attainable. Thus 321 records were obtained for men, and 882 for women.

Data analysis
Statistical analysis of data was performed using R version 2.6.1 for Windows [2], including models for nonlinear regression. The 5th percentile was computed by quantile regression with bootstrapping to estimate standard errors and confidence intervals (package quantreg in R [3]).


Various models are used in the literature to describe the relationship of ventilatory indices with standing height (H) and age (A). We tested the following models:

\[
\text{Index} = a + b\cdot H + c\cdot A \\
\text{Index} = a + b\cdot H + c\cdot A + d\cdot A^2 \\
\text{Index} = (a + b\cdot A)\cdot H^k \\
\text{Index} = a + b\cdot H^2 + c\cdot A + d\cdot A^2 \\
\ln(\text{Index}) = a + b\cdot\ln(H) + c\cdot A \\
\ln(\text{Index}) = (a + b\cdot A)\cdot H + c \\
\ln(\text{Index}) = a + b\cdot\ln(H) + c\cdot A + d\cdot A^2
\]

where \(\ln\) is the natural logarithm, and height and age are in centimetres and years, respectively.
Results
Standing height in women varied between 144-190 cm (mean 163.5, SD 5.90), age between 17-74 yr (mean 44.74, SD 13.26). Corresponding data for men were height 150-205 cm (mean 178.53, SD 7.51), age 17-79 yr (mean 36.10, SD 12.46).

Differences in explained variance for FEV₁%FVC were small; they varied between 0.187 and 0.196 in men, and between 0.188 and 0.197 in women. There was a slightly greater range in explained variance for FEV₁; between 0.604 and 0.634 in men, and between 0.554 and 0.588 in women. Given the very small differences in the performance of the various models we selected the model which combined a high explained variance with a distribution of residuals that came closest to a Gaussian distribution as indicated by Shapiro-Wilk index and inspection of Q-Q plots. Predicted values were as follows:

<table>
<thead>
<tr>
<th>Gender</th>
<th>Index</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>FEV₁%FVC</td>
<td>111.23 - 0.1246·H - 0.3040·A + 0.00064·A²</td>
</tr>
<tr>
<td></td>
<td>LLN FEV₁%FVC</td>
<td>144.33 - 0.3754·H - 0.2259·A - 0.00094·A²</td>
</tr>
<tr>
<td>Female</td>
<td>FEV₁%FVC</td>
<td>130.69 - 0.2112·H - 0.5068·A + 0.00283·A²</td>
</tr>
<tr>
<td></td>
<td>LLN FEV₁%FVC</td>
<td>123.72 - 0.2501·H - 0.2621·A - 0.00138·A²</td>
</tr>
<tr>
<td>Male</td>
<td>FEV₁</td>
<td>0.0504·H - 0.0121·A - 0.000254·A² - 3.797</td>
</tr>
<tr>
<td></td>
<td>LLN FEV₁</td>
<td>0.0438·H + 0.0297·A - 0.00080·A² - 4.142</td>
</tr>
<tr>
<td>Female</td>
<td>FEV₁</td>
<td>0.0372·H - 0.0174·A - 0.00012·A² - 2.135</td>
</tr>
<tr>
<td></td>
<td>LLN FEV₁</td>
<td>0.0298·H - 0.0448·A + 0.00019·A² - 1.002</td>
</tr>
</tbody>
</table>

LLN = 5th percentile
List of publications on predicted values for adult men and women. Some publications provided prediction equations for more than one ethnic group.


As cited in 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: 737-743.


45. Quanjer: This manuscript.
Appendix (for Online Data Supplement)

Reference values from the subset of healthy adults from two Dutch population-based cohorts

Population
The data were obtained in field surveys conducted at 3-yr intervals between 1965 and 1990 in a rural area (Vlagtwedde) and in an industrial town (Vlaardingen) in the Netherlands. Each survey was completed in 1 week in October. The Vlagtwedde sample was a dynamic cohort. In every field survey all men and women born between 1921 and 1952 and living in a specified area were invited to take part. The Vlaardingen sample was a fixed cohort selected from the municipal register. In 1965 a random sample of men and women born between 1901 and 1925 was selected, and in 1969 an additional sample was selected from those born between 1930 and 1954.

Measurements
Anthropometric and lung function measurements were obtained in a standardized way by trained operators. BTPS measurements of FEV$_1$ and FVC were derived from maximal expiratory flow-volume curves, using pneumotachography. Measurements and procedures complied with recommendations of the European Community for Coal and Steel\(^1\). Calibration was performed approximately every hour.


Data selection
For each person up to 5 measurements were available. From amongst persons who had never smoked and had been free of cardio-respiratory symptoms one record per person was selected using an algorithm designed to obtain the most even age distribution attainable. Thus 321 records were obtained for men, and 882 for women.

Data analysis
Statistical analysis of data was performed using R version 2.6.1 for Windows [2], including models for nonlinear regression. The 5\(^{th}\) percentile was computed by quantile regression with bootstrapping to estimate standard errors and confidence intervals (package quantreg in R [3]).


Various models are used in the literature to describe the relationship of ventilatory indices with standing height (H) and age (A). We tested the following models:

Index = \[ a + b \cdot H + c \cdot A \] \hspace{1cm} (1)
\[ a + b \cdot H + c \cdot A + d \cdot A^2 \] \hspace{1cm} (2)
\[ (a + b \cdot A) \cdot H^k \] \hspace{1cm} (3)
\[ a + b \cdot H^2 + c \cdot A + d \cdot A^2 \] \hspace{1cm} (4)

\[ \ln(\text{Index}) = a + b \cdot \ln(H) + c \cdot A \] \hspace{1cm} (5)
\[ (a + b \cdot A) \cdot H + c \] \hspace{1cm} (6)
\[ a + b \cdot \ln(H) + c \cdot A + d \cdot A^2 \] \hspace{1cm} (7)

where ln is the natural logarithm, and height and age are in centimetres and years, respectively.
Results
Standing height in women varied between 144-190 cm (mean 163.5, SD 5.90), age between 17-74 yr (mean 44.74, SD 13.26). Corresponding data for men were height 150-205 cm (mean 178.53, SD 7.51), age 17-79 yr (mean 36.10, SD 12.46).

Differences in explained variance for FEV1%FVC were small; they varied between 0.187 and 0.196 in men, and between 0.188 and 0.197 in women. There was a slightly greater range in explained variance for FEV1: between 0.604 and 0.634 in men, and between 0.554 and 0.588 in women. Given the very small differences in the performance of the various models we selected the model which combined a high explained variance with a distribution of residuals that came closest to a Gaussian distribution as indicated by Shapiro-Wilk index and inspection of Q-Q plots. Predicted values were as follows:

<table>
<thead>
<tr>
<th>Gender</th>
<th>Index</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>FEV1%FVC</td>
<td>111.23 - 0.1246·H - 0.3040·A + 0.00064·A²</td>
</tr>
<tr>
<td></td>
<td>LLN FEV1%FVC</td>
<td>144.33 - 0.3754·H - 0.2259·A - 0.00094·A²</td>
</tr>
<tr>
<td>Female</td>
<td>FEV1%FVC</td>
<td>130.69 - 0.2112·H - 0.5068·A + 0.00283·A²</td>
</tr>
<tr>
<td></td>
<td>LLN FEV1%FVC</td>
<td>123.72 - 0.2501·H - 0.2621·A - 0.00138·A²</td>
</tr>
<tr>
<td>Male</td>
<td>FEV1</td>
<td>0.0504·H - 0.0121·A - 0.000254·A² - 3.797</td>
</tr>
<tr>
<td></td>
<td>LLN FEV1</td>
<td>0.0438·H + 0.0297·A - 0.00080·A² - 4.142</td>
</tr>
<tr>
<td>Female</td>
<td>FEV1</td>
<td>0.0372·H - 0.0174·A - 0.00012·A² - 2.135</td>
</tr>
<tr>
<td></td>
<td>LLN FEV1</td>
<td>0.0298·H - 0.0448·A + 0.00019·A² - 1.002</td>
</tr>
</tbody>
</table>

LLN = 5th percentile
List of publications on predicted values for adult men and women. Some publications provided prediction equations for more than one ethnic group.


45. Quanjer: This manuscript.