Should forced expiratory volume in six seconds replace forced vital capacity to detect airway obstruction?

J.E. Hansen, X-G. Sun and K. Wasserman

ABSTRACT: It has been suggested that forced expiratory volume in six seconds (FEV6) should be substituted for forced vital capacity (FVC) to measure fractions of timed expired volume for airflow obstruction detection. The present authors hypothesised that this recommendation might be questionable because flow after 6 s of forced expiration from more diseased lung units with the longest time constants was most meaningful and should not be ignored. Furthermore, previous studies comparing FEV6 and FVC included few subjects with mild or no disease.

The present study used spirometric data from the USA Third National Health and Nutrition Evaluation Survey with prior published ethnicity- and sex-specific equations for FEV1/FEV6, FEV1/FVC and FEV3/FVC, and new equations for FEV3/FEV6, all derived from ∼4,000 adult never-smokers aged 20–80 yrs.

At 95% confidence intervals, 21.3% of 3,515 smokers and 41.3% of smokers aged >51 yrs had airway obstruction; when comparing FEV1/FEV6 with FEV1/FVC, 13.5% were concurrently abnormal, 1.5% were false positives and 4.1% were false negatives; and when comparing FEV3/FEV6 with FEV3/FVC, 11.6% were concurrently abnormal, 3.3% were false positives and 5.7% were false negatives.

Substituting forced expiratory volume in six seconds for forced vital capacity to determine the fractional rates of exhaled volumes reduces the sensitivity of spirometry to detect airflow obstruction, especially in older individuals and those with lesser obstruction.

KEYWORDS: Airway obstruction, cigarette smoking, forced expiratory volume in six seconds, forced expiratory volume in three seconds, forced vital capacity, spirometry

In 1999, using the large National Health and Nutrition Examination Survey (NHANES) III database, a number of spirometric reference equations, including those for forced expiratory volume in six seconds (FEV6) and FEV1/FEV6, were published [1]. In 2000, a National Lung Health Education Program consensus statement [2] advocated replacement of forced vital capacity (FVC) and FEV1/FVC with FEV6 and FEV1/FEV6 to detect airways obstruction. Later, Swanney et al. [3] reported high sensitivity and specificity for FEV1/FEV6 compared to gold standard FEV1/FVC in 337 out of 502 patients tested in a tertiary hospital-based university laboratory. In a multicentred lung health study, Enright et al. [4] concluded that FEV1/FEV6 values could be useful for following the course of obstructive airways disease in smokers and for screening smokers for the presence of airway obstruction. Subsequently, Vandevenorde et al. [5] concluded from a large patient study that “the FEV1/FEV6 ratio can be used as a valid alternative for FEV1/ FVC in the diagnosis of airway obstruction, especially for screening purposes”. Other investigators recommended using FEV6 rather than FVC for the mean forced expiratory flow between 25 and 75% of FVC (FEF25–75%) and for measuring lung restriction [3, 6–9].

FEV1/FVC, FEV1/FEV6 and FEV3/FVC ratios, derived from the large never-smoking NHANES III database, all decrease in a linear fashion as age increases, indicating an increase in long time-constant lung units or 1-FEV3/FVC [1, 10]. Each of these formulae correctly identifies patients with severe airway obstruction. However, subjects with subtler obstruction also commonly exhale an important portion of their FVC after 6 s, i.e. from lung units discharging their gas late in exhalation. Consequently, FEV1/FEV6 and FEV3/FEV6 measurements, with denominators which exclude the FVC–FEV6 volumes, may be less discriminating than FEV1/FVC and FEV3/ FVC in detecting milder airway obstruction. In
screening for disease, it may be better to focus on detecting lung units with long time constants (after 3 or 6 s) rather than on shaving seconds off expiratory time and centilitres off forced expiratory volumes.

The present authors, therefore, hypothesised that FEV1/FEV6 and FEV3/FEV6 would be less reliable screening parameters than FEV1/FVC and FEV3/FVC in distinguishing changes in lung function due to normal ageing from those due to superimposed airway obstruction from smoking.

METHODS
Data meeting American Thoracic Society (ATS) standards [11] were extracted from the NHANES III database [12] for 13,113 adults, including ex-smokers, 5,943 never-smokers and 3,515 current smokers with no apparent skeletal or neuromuscular disease. Data were obtained nationwide with informed consent. Some data and equations derived from this population were previously published by others [9, 10] and by the present authors [10].

Regression equations [13] for mean and 95% confidence lower limits of normal (95% LLN) for FEV3/FEV6 (table 1) were derived for never-smokers identified ethnically as Black, Latin and White [10]. These never-smokers had similar FEV1/FVC derived for never-smokers identified ethnically as Black, Latin and White [10]. These never-smokers had similar FEV1/FVC values [8] to those of HANKINSON et al. [1], derived from the same database.

The 3,515 current smokers were then divided into four similar-sized groups according to age: 20–29.3 yrs; 29.4–38.1 yrs; 38.2–50.7 yrs; and 50.8–80 yrs. Using sex- and ethnicity-specific equations for FEV1/FVC and FEV1/FEV6 [1], FEV3/FVC [10] and the newly derived FEV3/FEV6, each value of the current smokers was categorised as “normal” or “abnormal”, depending on whether it was above or below 95% LLN values, as all ratios had normal distributions in never-smokers.

Given \( x = FEV1 \) or \( FEV3 \), deviations of discordant \( x/FEV6 \) values from \( x/FVC \) values, i.e. false positive or negative, were calculated as follows:

\[
\% \text{ deviation} = \frac{\% \text{ deviation} \text{ of FEV}3/\text{FEV}6}{\% \text{ deviation} \text{ of FEV}1/\text{FEV}6} \times 100
\]

When the paired ratios (x/FVC and x/FEVs) were both above their LLN, they were concordant normal; when both were below their LLN, they were concordant abnormal. However, when an individual FEV1/FVC was normal and the FEV1/FEVs was below LLN, for example, the FEV1/FEVs value was considered discordant and false positive. When an FEV3/FEVC was below LLN and FEV3/FEVs was normal, the FEV3/FEVs was considered discordant and false negative.

Deviations of false positive or negative x/FEVs values from x/ FVC values were calculated as follows:

\[
\% \text{ deviation} = \frac{\% \text{ deviation} \text{ of FEV}3/\text{FEV}6}{\% \text{ deviation} \text{ of FEV}1/\text{FEV}6} \times 100
\]

Ratios of false positive and false negative to concordant abnormal values were calculated for groups of differing age and severities of obstruction. In those with abnormal FEV1/ FVC and/or abnormal FEV3/FVC, severity of obstruction was based on FEV1 % predicted: severe <50%; moderate 50–65%; mild 65–80%; and minimal >80% and <120%. Two-by-two tables were created for the calculation of sensitivity, specificity, and positive and negative predictive values for each age group. To counter the potential criticism that a 95% LLN (mean–1.645 ×se) might be spurious or too strict, all analyses were repeated (but not necessarily reported) using a confidence limit of 99% (99% LLN; mean–2.33 ×se).

RESULTS
Figure 1 shows the FVC manoeuvre durations for 13,113 subjects in the NHANES III survey aged ≥20 yrs who had optimal tests. In individuals with longer forced expirations (late emptying of long time-constant units), the mean and variability of the volume differences from FEVs values increased markedly (almost similar in mL to the square of the duration of FVC in s). Table 2 shows that the volume differences between FVC and FEVs were higher for smokers and increased with age, especially in current smokers.

Using ethnicity- and sex-specific formulae, percentages of NHANES III current smokers found to have abnormal FEV1/ FVC or FEV3/FVC are displayed in figure 2 for each age group.

### TABLE 1

**Forced expiratory volume in three seconds (FEV3)/forced expiratory volume in six seconds (FEVs) per cent formula for never-smoking adults**

<table>
<thead>
<tr>
<th>Group</th>
<th>Subjects n</th>
<th>Age factor</th>
<th>Mean constanta</th>
<th>95% LLN constantb</th>
<th>se</th>
<th>( r^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Females</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>1149</td>
<td>-0.0867</td>
<td>99.40</td>
<td>96.50</td>
<td>2.09</td>
<td>0.605</td>
</tr>
<tr>
<td>Latin</td>
<td>1248</td>
<td>-0.0958</td>
<td>100.29</td>
<td>97.31</td>
<td>1.82</td>
<td>0.406</td>
</tr>
<tr>
<td>White</td>
<td>1440</td>
<td>-0.1026</td>
<td>100.69</td>
<td>97.30</td>
<td>2.06</td>
<td>0.442</td>
</tr>
<tr>
<td>All</td>
<td>3830</td>
<td>-0.0958</td>
<td>100.32</td>
<td>97.03</td>
<td>2.00</td>
<td>0.411</td>
</tr>
<tr>
<td><strong>Males</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>634</td>
<td>-0.0815</td>
<td>99.31</td>
<td>95.76</td>
<td>2.16</td>
<td>0.239</td>
</tr>
<tr>
<td>Latin</td>
<td>699</td>
<td>-0.0942</td>
<td>99.81</td>
<td>97.11</td>
<td>1.64</td>
<td>0.416</td>
</tr>
<tr>
<td>White</td>
<td>775</td>
<td>-0.0842</td>
<td>99.38</td>
<td>96.34</td>
<td>1.85</td>
<td>0.396</td>
</tr>
<tr>
<td>All</td>
<td>2113</td>
<td>-0.0865</td>
<td>99.50</td>
<td>96.40</td>
<td>1.88</td>
<td>0.367</td>
</tr>
</tbody>
</table>

LLN: confidence lower limits of normal. a: mean FEVs/FEVs = constant + age in yrs x age factor; b: LLN at 95% confidence limit = 95% LLN constant + age in yrs x age factor.
As expected, abnormalities increased in older age groups. Using 95% LLN values, >41% of 50.8–80-yr-old smokers had evidence of airway obstruction. Most commonly, in all age groups, both FEV1/FVC and FEV3/FVC were abnormal. If only one was abnormal, it was more likely to be FEV1/FVC in younger smokers and FEV3/FVC in older smokers. As noted in table 3, using FEV1/FVC and FEV3/FVC as standards, the overall incidence of airway obstruction exceeded 20% in smokers. Severe airway obstruction was rare except in the oldest age group.

Table 4 displays, across age groups, both 95% and 99% LLN values. Using 95% LLN for FEV1/FEV6 and FEV1/FVC, a total of 473 concordant abnormal pairs and 194 discordant pairs (51 false positives and 143 false negatives) were found with a discordant/concordant abnormal ratio of 315/408 = 77%. The number of false negatives increased strikingly with age. Using 95% LLN for both FEV1 and FEV3 comparisons, total specificities were relatively high, negative and positive predictive values were intermediate, while sensitivities were low.

Using 99% LLN for FEV1/FEV6 and FEV1/FVC pairs and FEV3/FEV6 and FEV3/FVC pairs, (table 4), proportions of discordant to concordant abnormal pairs actually increased while sensitivities declined. The high incidence of discordant values, false-negative values, low sensitivity and even FEV3/FEV6false positives confirm the low reliability of the FEV1/FEV6 to detect airway obstruction in this population.

Figure 3 shows that at 95% LLN, discord increased markedly as the severity of airway obstruction decreased. Table 5 shows that such false-negative discords also tended to be larger than false-positive discords; mean absolute mismatch of these ratios was 2.64%.

**DISCUSSION**

The present study found low sensitivities, and a high incidence of false negative FEV1/FEV6 and FEV3/FEV6 and a moderate
incidence of false positive FEVs/FEV6s in this NHANES III population, supporting the present authors’ hypothesis that the use of FEVs in place of FVC reduces the sensitivity of spirometry in detecting airway disease. Prior findings [3–7] that promote FEVs as an acceptable surrogate for FVC are therefore further detailed in table 6. The study from a university hospital-based laboratory [3], which used spirometry from 310 patients, found high sensitivities and specificities when comparing FEV1/FEV6 to FEV1/FVC. However, 53% of their patients had severe (35%) or moderate (18%) obstruction. Their conclusion, stated in table 6, might not be valid in populations with a lower severity of airway obstruction. A recent re-emphasis [14] suggests that patients with airway obstruction may have low sensitivities for older smokers. Interestingly, the higher incidence of false positives than false negatives is surprising, since eliminating flow after 6 s would favor a finding of false-negative FEV1/FEVs ratios, as in the present study.

First, the differences in the populations studied will be considered. In the NHANES III database, adult never-smokers outnumbered current smokers. Furthermore, <20% of current smokers had an FEV1/FVC below 95% LLN and only 3% had obstruction considered moderate or severe. This incidence is much lower than that of 53% and 25% of moderate-to-severe obstruction found in the two university hospital patient studies (table 6). The present authors believe that NHANES III smokers and nonsmokers in their study better represent the USA (or other) general populations likely to request or receive spirometric screening by primary care physicians or other providers. The NHANES III analyses disclose that discord increases as severity of obstructive airways declines (fig. 3) and average sensitivities (table 4) fall to 77 and 67% at 99% LLN for FEV1/FEV6 and FEV3/FEV6, respectively, for all NHANES III smokers, with even lower sensitivities for older smokers.

Secondly, as late flow occurs when longer time-constant lung units play a more prominent role in expiratory airflow, it is not surprising that exclusion of late flow by terminating flow, volume and ratio measurements at 6 s, causes low sensitivities and false negatives. As has been recognized and recently re-emphasized [14], patients with airway obstruction
frequently have slow or unforced vital capacity volumes exceeding those of forced manoeuvres. The fact that nearly one-quarter to one-third of smokers with airway obstruction as discerned using FEV1/FVC and FEV3/FVC at 95% LLN and one-third to one-half at 99% LLN are excluded by substituting FEV1/FEV6 and FEV3/FEV6 should curb enthusiasm for use of the latter measures to detect obstructive disease in a general population, including smokers.

Reasons for false positives in the present study using FEV1/FEV6 and FEV3/FEV6 are less obvious. On review of the present data, the correlations with age are inferior to those for FEV1/FVC and FEV3/FVC ratios for each ethnic and sex group. Perhaps more importantly, SE values for FEV6 ratios are invariably, but minimally, lower than SE values for FVC ratios for every ethnic and sex group. This results in defining narrower “windows” of abnormality, so that ratios using FEV6 may “find” airway obstruction outside those windows when it is not present.

It also appears (fig. 2) that FEV1/FVC identifies airway obstruction slightly less often than FEV3/FVC, especially in older smokers. Confirming that both ratios detect deterioration with smoking, the present authors previously found that, by middle age, both FEV1/FVC and FEV3/FVC values of current smokers are similar to those of never-smokers who are 20 yrs older [10]. Excluding FEV3/FVC from spirometric analyses misses some airway obstruction.

Possible limitations
As in all prior studies comparing values and ratios of FEVs to FVC, sharp cut-off lines were used in order to distinguish the actual differences between the equations. Although it could be argued that sharp cut-off lines are inappropriate, differences between equations cannot be detected and statistically analysed without using such limits. As populations have a greater variability of FVC or FEV1 than their ratios, subjects with abnormal ratios and FEV1 within normal limits (<0.80–120% pred) could be normal or be minimally obstructed, but have a higher morbidity [14].

Perspective
A recent publication [15] reviewed, discarded, selected and analysed a large number of past studies based on clinical evaluation, spirometry and questionnaires to evaluate the effect of multiple therapies on patients with or suspected of having obstructive lung disease. It concluded that: “80 percent of adults reporting a clinical diagnosis of chronic bronchitis or emphysema did not have current airflow obstruction”, spirometry increased 1-yr smoking cessation quit rates by only 1% and “COPD treatment trials including inhaled medications, pulmonary rehabilitation, disease management, or surgery, improved […] functional status […] less than considered clinically significant”. After stating that spirometry was a useful diagnostic tool in evaluating individuals with symptoms suggestive of COPD, Wilt et al. [15] concluded that: “spirometric testing is likely to label a large number of individuals (many who do not report respiratory symptoms) with disease and result in considerable testing and treatment costs and healthcare resource utilization”. In reaching their conclusions, which might not find agreement from other pulmonologists, it must be noted that many of their referenced studies inappropriately used fixed ratios of FEV1/FVC as criteria for obstruction, rather than ratios dependent on age.

Eaton et al. [16] placed quality spirometers in 30 primary care practices and assessed the results. They found: 1) 2 h of physician and nurse training, and further experience were important in improving quality of tracings; 2) spirometric manoeuvres were commonly terminated prematurely; 3) even with training it was rare to get two (33%) or three (19%) blows
<table>
<thead>
<tr>
<th>First author [ref.]</th>
<th>Sites</th>
<th>Subjects n</th>
<th>Subject characteristics</th>
<th>Severity of obstruction</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>Study observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swanney [3]</td>
<td>Hospital laboratory, New Zealand</td>
<td>337 out of 502</td>
<td>Aged 20–89 yrs</td>
<td>35% severe, 18% moderate, 12% mild</td>
<td>95.0</td>
<td>97.4</td>
<td>98.6</td>
<td>91.1</td>
<td>“FEV6 is an acceptable surrogate for FVC...”</td>
</tr>
<tr>
<td>Enright [4]</td>
<td>Clinical centres (10 sites) in USA and Canada</td>
<td>5887</td>
<td>From Lung Health Study; smokers aged &gt;35 yrs</td>
<td>Mean FEV1/FVC 63%</td>
<td>94.0</td>
<td>93.1</td>
<td>89.8</td>
<td>96.0</td>
<td>“…The study demonstrates that the use of the FEV1/FEV6 is a good substitute for the FEV1/FVC when screening smokers for the presence of airways obstruction...”</td>
</tr>
<tr>
<td>Vanfveoorde [5, 6]</td>
<td>University Hospital, Brussels</td>
<td>40% of 11676 with obstruction</td>
<td>Aged 20–82 yrs</td>
<td>13% of total tested and 32% of obstruction severe</td>
<td>94.0</td>
<td>93.1</td>
<td>89.8</td>
<td>96.0</td>
<td>“FEV1/FEV6 ratio can be used as a valid alternative to FEV1/FVC in the diagnosis of airway obstruction, especially for screening purposes...” Using FEV1/FVC of &lt;70% versus FEV6/FVC &lt;73% found PPV of 92.2 and NPV of 95.2.</td>
</tr>
<tr>
<td>Akpinar-Eli [7]</td>
<td>In the workplace, West Virginia (USA)</td>
<td>1139</td>
<td>Workers; 43% never-smokers</td>
<td>15% obstructed</td>
<td>92</td>
<td>98</td>
<td>92</td>
<td>98</td>
<td>Used excellent equipment technicians; recommended more studies of flow-sensing spirometers in occupational sites</td>
</tr>
<tr>
<td>The present study</td>
<td>Multiple NHANES III sites</td>
<td>3515 current smokers</td>
<td>Ambulatory non-patients aged 20–80 yrs</td>
<td>0.8% of total smokers and 4% of obstruction severe</td>
<td>76.7*</td>
<td>98.2*</td>
<td>90.3*</td>
<td>95.2*</td>
<td></td>
</tr>
<tr>
<td>The present study*</td>
<td>Multiple NHANES III sites</td>
<td>3515 current smokers</td>
<td>Ambulatory non-patients aged 20–80 yrs</td>
<td>0.8% of total smokers and 4% of obstruction severe</td>
<td>66.2**</td>
<td>99.4**</td>
<td>91.2**</td>
<td>96.5**</td>
<td></td>
</tr>
</tbody>
</table>

PPV: positive predictive value; NPV: negative predictive value; NHANES: National Health and Nutrition Examination Survey. *: compared FEV6/FEV6 with FEV1/FVC; all the rest compared FEV1/FEV6 with FEV1/FVC. **: values at p<0.05; ***: values at p<0.01.
meeting ATS criteria; 4) primary care physician interpretations were deemed to be correct only 53% of the time; and 5) only an average of 2.3 tests were performed weekly at each site. Nevertheless, the practitioners believed that 13% of the tests helped in counselling smokers. To place these findings in perspective, one might ask if 13% of radiographs, electrocardiograms, or mammograms performed and interpreted in a primary care practice are helpful in counselling patients. The present authors wonder whether physicians should rely on tests from equipment and personnel used so infrequently.

Despite the recommendation of the Global Initiative for Chronic Obstructive Lung Disease Committee [17], the present authors believe that it is unwise to ignore age and use fixed ratios of FEV1/FVC such as 70% to identify airway obstruction. Rather than identifying 5% of normal individuals at the 95% confidence limits as abnormal, such fixed limits are certain to underdiagnose airway obstruction in younger individuals and overdiagnose airway obstruction in older individuals.

Therefore, there are diverse approaches and recommendations regarding spirometry. On the one hand, there is the desire to reduce costs by having minimally trained personnel use simpler and cheaper equipment, with emphasis on measurement of FEV1, FEVs and FEF25-75% and subsequent interpretation at the primary care level. On the other hand, there are concerns that spirometry is costly and of limited value in detecting early lung disease, reducing the incidence of smoking, or following the effect of therapy in those with known lung disease. The present authors favour a third approach: referral of patients with pulmonary symptoms or a significant smoking history to sites where well-trained personnel with excellent equipment test a large number of patients per day, measuring FEV1, FEV3, FVC, and their ratios, adding, when indicated, measurements of the slow vital capacity, inspiratory capacity, expiratory reserve volume, inspiratory flow and total lung capacity, with interpretation of the values and tracings by experienced pulmonologists or other similarly well-trained physicians. In the present authors’ opinion, this latter approach would be cost-effective, result in more accurate diagnoses and be in everyone’s best interests.

It is concluded that quality spirograph measures, with proper reference standards, must be used to accurately identify airway obstruction. The perceived benefit of terminating forced expiratory manoeuvres at 6 s discards data from the most obstructed lung units and reduces the sensitivity of detection of obstructive lung disease.

REFERENCES
3 Swanney MP, Jensen RL, Crichton DA, Beckert LE, Cardno LA, Crapo RO. FEV(6) is an acceptable surrogate for FVC in the spirometric diagnosis of airway obstruction and restriction. _Am J Respir Crit Care Med_ 2000; 162: 917–919.