The sources of variation in lung function testing and technical aspects of spirometry, lung volume measurements and $D_LCO$ measurement have been considered in other documents published in this series of Task Force reports [1–4] and in the American Thoracic Society (ATS) interpretative strategies document [5].

An interpretation begins with a review and comment on test quality. Tests that are less than optimal may still contain useful information, but interpreters should identify the problems and the
direction and magnitude of the potential errors. Omitting the quality review and relying only on numerical results for clinical decision making is a common mistake, which is more easily made by those who are dependent upon computer interpretations.

Once quality has been assured, the next steps involve a series of comparisons [6] that include comparisons of test results with reference values based on healthy subjects [5], comparisons with known disease or abnormal physiological patterns (i.e. obstruction and restriction), and comparisons with self, a rather formal term for evaluating change in an individual patient. A final step in the lung function report is to answer the clinical question that prompted the test.

Poor choices made during these preparatory steps increase the risk of misclassification, i.e. a falsely negative or falsely positive interpretation for a lung function abnormality or a change in lung function. Patients whose results are near the thresholds of abnormality are at a greatest risk of misclassification.

**REFERENCE EQUATIONS**

**General issues**

Interpretation of PFTs is usually based on comparisons of data measured in an individual patient or subject with reference (predicted) values based on healthy subjects. Predicted values should be obtained from studies of “normal” or “healthy” subjects with the same anthropometric (e.g. sex, age and height) and, where relevant, ethnic characteristics of the patient being tested. Ideally, reference values are calculated with equations derived from measurements observed in a representative sample of healthy subjects in a general population. Reference equations can also be derived from large groups of volunteers, provided that criteria for normal selection and proper distribution of anthropometric characteristics are satisfied. Criteria to define subjects as “normal” or healthy have been discussed in previous ATS and European Respiratory Society (ERS) statements [5, 7, 8].

Height and weight should be measured for each patient at the time of testing; technicians should not rely on stated height or weight. Height should be measured with a stadiometer, with shoes off, using standard techniques (patient standing erect with the head in the Frankfort horizontal plane) [9]. When height cannot be measured, options include using stated height or estimating height from arm span, as indicated in a previous document from this series and other publications [1, 10, 11].

Specific recommendations for selecting reference values to be used in any lung function laboratory have also been discussed [3]. These include the following: matching age-range, anthropometric, race/ethnic, socio-economic and environmental characteristics between subjects investigated by the laboratory and the reference population from which the prediction equations have been drawn; using similar instruments and lung function protocols in the reference population as in the laboratory; and using reference values derived by valid and biologically meaningful statistical models, taking into account the dependence of lung function with age. If possible, all parameters should be taken from the same reference source. For example, forced vital capacity (FVC), forced expiratory volume in one second (FEV1), and FEV1/FVC should come from the same reference source.

The subjects being tested should be asked to identify their own race/ethnic group, and race/ethnic-specific reference equations should be used whenever possible. If such equations are not available or are unsuitable for a particular setting, a race/ethnic adjustment factor based on published data may be used for lung volumes. The use of adjustment factors is not as good as specific race/ethnic equations [12]. An example of adjustment factors is the finding that populations using standing height as the measure of size tend to overpredict values measured in Black subjects by ~12% for total lung capacity (TLC), FEV1 and FVC, and by ~7% for functional residual capacity (FRC) and residual volume (RV) [5]. A race/ethnic adjustment factor of 0.94 is also recommended for Asian Americans based on two recent publications [13, 14]. Such adjustment factors should not be applied to the FEV1/FVC or FEV1/vital capacity (VC) ratios. The use of sitting height does not completely account for race/ethnic differences in pulmonary function [15]. If a race adjustment factor is used, a statement should be included in the report, along with the race adjustment value used.

Differences in the evaluation of lung function using different sets of reference equations have been documented [16, 17]. Ideally, spirometric reference values should be derived from a population similar to the individual subject using the same kind of instrument and testing procedure.

There have been recommendations to compare selected reference equations with measurements performed on a representative sample of healthy subjects tested in each laboratory. The reference equation that provides the sum of residuals (observed – predicted computed for each adult subject, or log observed – log predicted for each subject in the paediatric age range) closest to zero will be the most appropriate for that laboratory [7]. However, for spirometry, a relatively large number of subjects (i.e. n=100) is necessary to be confident that a significant difference between the published reference equations and the values from the local community does not exist [18]. Therefore, the suggestion is impractical for most laboratories.

When using a set of reference equations, extrapolation beyond the size and age of investigated subjects should be avoided [7]. If a patient’s age or height is outside the limits of the reference population, a statement in the interpretation should indicate that an extrapolation has been made.

Publications on reference equations should include explicit definitions of the upper and lower limits of the normal range, or provide information to allow the reader to calculate a lower range [5]. For each lung function index, values below the 5th percentile of the frequency distribution of values measured in the reference population are considered to be below the expected “normal range” [5]. If the reference data have a normal distribution, the lower 5th percentile can be estimated as the 95% confidence interval using Gaussian statistics. If the distribution is skewed, the lower limit should be estimated with a nonparametric technique, such as the 95th percentile. The practice of using 80% predicted as a fixed value for the lower limit of normal may be acceptable in children, but can lead to important errors when interpreting lung function in adults [5]. The practice of using 0.70 as a lower limit of the
FEV1/FVC ratio results in a significant number of false-positive results in males aged >40 yrs and females >50 yrs [12], as well as in a risk of overdiagnosis of chronic obstructive pulmonary diseases (COPD) in asymptomatic elderly never-smokers [19]. This discussion has been focused on the lower limit of the reference range. Upper limits are appropriate where the variable can be either too high or too low. Such variables include TLC, RV/TLC and DL CO. As equipment and techniques for lung function testing improve, advanced mathematical models to describe lung function data are implemented. Furthermore, the characteristics of the populations of “normal” subjects, with respect to nutrition, health status, environmental conditions and other factors, evolve (a phenomenon also described as “cohort effect”). Consideration should be given to updating reference equations on a regular basis, e.g. every 10 yrs, taking into account the applicability of the newer reference equations and the effect on interpretation of longitudinal patient follow-up.

Manufacturers should also provide software that allows users to easily select among a panel of reference equations. They should also allow easy insertion of new equations. The reference values used should be documented on every pulmonary function report with the first author’s last name (or organisation) and the date of publication.

**Spirometry**

The European Community for Coal and Steel (ECCS) [8, 20] and the ATS [5, 21] have both published comprehensive listings of published reference equations for spirometry. A number of additional studies on lung function reference values, dealing with a variety of ethnic/race groups and age ranges, have been published in the last 10 yrs [12, 14, 17, 22, 23].

Spirometric reference equations are usually derived from cross-sectional studies and are subject to “cohort effect”. Few authors have published longitudinal equations covering ages from childhood to the elderly [24–26], and there are few published sets of equations that cover volume and flow indices over a wide range of ages [27, 28]. Table 1 includes reference equations published from 1995 to August 2004. The table was created from known equations and a MEDLINE search using the keywords “reference equations” and “spirometry”. Its purpose is to recognise and encourage the continuing interest of worldwide researchers in deriving and using reference equations.

In the USA, ethnically appropriate National Health and Nutrition Examination Survey (NHANES) III reference equations are recommended for those aged 8–80 yrs [12]. For children aged <8 yrs, the equations of WANG et al. [29] are recommended. Other prediction equations may be used if there are valid reasons for the choice. In Europe, the combined reference equations published in the 1993 ERS statement [8] are often used for people aged 18–70 yrs, with a height range of 155–195 cm in males, and 145–180 cm in females, and those from QUANJER et al. [30] in paediatric ages. Currently, this committee does not recommend any specific set of equations for use in Europe, but suggests the need for a new Europe-wide study to derive updated reference equations for lung function.

**Lung volumes**

Lung volumes are related to body size, and standing height is the most important correlating variable. In children and adolescents, lung growth appears to lag behind the increase in standing height during the growth spurt, and there is a shift in the relationship between lung volume and height during adolescence [31, 32]. Height growth in young males between 12.5 and 18 yrs of age peaks ~1 yr before the growth rate of weight and FVC, and ~1.5 yrs before the growth rate of maximum flow at 50% FVC. In young females, growth rates of all spirometric indices decrease over the same age range. Using simple allometric relationships between stature and lung volumes, volume predictions are too high in the youngest age group and too low in the oldest adolescents.

Furthermore, for the same standing height, young males have greater lung function values than young females, and Whites have greater values than Blacks. Lung function increases linearly with age until the adolescent growth spurt at age ~10 yrs in females and 12 yrs in males. The pulmonary function versus height relationship shifts with age during adolescence. Thus, a single equation or the pulmonary function–height growth chart alone does not completely describe growth during the complex adolescent period. Nevertheless, race- and sex-specific growth curves of pulmonary function versus height make it easy to display and evaluate repeated measures of pulmonary function for an individual child [29].

Details of reference populations and regression equations for children and adolescents are summarised by QUANJER et al. [30]. Lung volume reference equations have been frequently derived from relatively small populations (<200 children) over a 6–12-yr age range when growth and developmental changes are extremely rapid. Relatively few studies have taken puberty or age into account.

A comprehensive listing of published reference equations for lung volumes was published in 1983 by the ECCS [20] and updated in 1993 [8]. A set of equations was created by combining the equations in this list with the intent to use the combined equations for adults aged 18–70 yrs with a height range of 155–195 cm in males, and 145–180 cm in females.


Differences due to ethnicity are not well defined [33–36]. These differences may be explained, in part, by differences in trunk length relative to standing height, but there are also differences in fat-free mass, chest dimensions and strength of respiratory muscles. Until better information is available, correction factors for Black and Asian children could be the same as those recommended for adults [7]. Reference values for RV, VC and TLC are, on average, 12% lower in Blacks than in Whites [35]; this difference may be smaller in elderly persons than in young adults [36]. Reference values for absolute lung volumes for adults of Asian ethnicity are generally considered to be lower than for Whites, but the magnitude of the differences is not well defined, and the difference may be less in Asians raised on “Western” diets during childhood [37]. According to the

**LUNG FUNCTION TEST INTERPRETATION**

R. PELLEGRINO ET AL.

950

VOLUME 26 NUMBER 5

EUROPEAN RESPIRATORY JOURNAL
### Table 1: Outcomes of a MEDLINE search using the keywords “reference equations” and “spirometry”

<table>
<thead>
<tr>
<th>First author</th>
<th>Year</th>
<th>Country/area/nature of study</th>
<th>Journal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kotaniemi</td>
<td>2004</td>
<td>Finland; adults</td>
<td>Int J Circumpolar Health 2004; 63: 129–139</td>
</tr>
<tr>
<td>Falaschetti</td>
<td>2004</td>
<td>England; prediction equations from the Health Survey</td>
<td>Eur Respir J 2004; 23: 456–463</td>
</tr>
<tr>
<td>Ben Saad</td>
<td>2003</td>
<td>Tunisia; vital capacity and peak expiratory flow rates in the elderly</td>
<td>Rev Mal Respir 2003; 20: 521–530 (French)</td>
</tr>
<tr>
<td>Perez-Padilla</td>
<td>2003</td>
<td>Mexico; comparison with Mexican American children</td>
<td>Pediatr Pulmonol 2003; 35: 177–183</td>
</tr>
<tr>
<td>Torres</td>
<td>2003</td>
<td>Brazil; height and arm span in children</td>
<td>Pediatr Pulmonol 2003; 36: 202–208</td>
</tr>
<tr>
<td>Golshani</td>
<td>2003</td>
<td>Iran</td>
<td>Eur Respir 2003; 22: 529–534</td>
</tr>
<tr>
<td>Mohamed</td>
<td>2002</td>
<td>Italy</td>
<td>Lung 2002; 180: 149–159</td>
</tr>
<tr>
<td>Boskabady</td>
<td>2002</td>
<td>Iran</td>
<td>Respiration 2002; 69: 320–326</td>
</tr>
<tr>
<td>Marion</td>
<td>2001</td>
<td>USA; American Indian</td>
<td>Chest 2001; 120: 489–495</td>
</tr>
<tr>
<td>Perez-Padilla</td>
<td>2001</td>
<td>Mexico; Mexican workers</td>
<td>Salud Publica Mex 2001; 43: 113–121 (Spanish)</td>
</tr>
<tr>
<td>Pistelli</td>
<td>2000</td>
<td>Italy</td>
<td>Am J Respir Crit Care Med 2000; 161: 899–905</td>
</tr>
<tr>
<td>Baltpoulos</td>
<td>2000</td>
<td>Greece; Greek elderly</td>
<td>Lung 2000; 178: 201–212</td>
</tr>
<tr>
<td>Quadrielli</td>
<td>1999</td>
<td>Italy</td>
<td>Respir Med 1999; 93: 523–535</td>
</tr>
<tr>
<td>Morato Rodriguez</td>
<td>1999</td>
<td>Spain; children of Basque Autonomic Community</td>
<td>An Esp Pediatr 1999; 51: 17–21 (Spanish)</td>
</tr>
<tr>
<td>Hankinson</td>
<td>1999</td>
<td>USA population sample, aged 8–80 yrs (NHANES III)</td>
<td>Am J Respir Crit Care Med 1999; 159: 179–187</td>
</tr>
<tr>
<td>Pan</td>
<td>1997</td>
<td>China; Taiwan</td>
<td>Chin J Physiol 1997; 40: 165–174</td>
</tr>
<tr>
<td>Luttmann</td>
<td>1997</td>
<td>Germany; 7–18-yr-old probands</td>
<td>Pneumologie 1997; 51: 47–54 (German)</td>
</tr>
<tr>
<td>Chin</td>
<td>1997</td>
<td>Singapore; nonsmoking adults</td>
<td>Respir Med 1997; 2: 143–149</td>
</tr>
<tr>
<td>Oyarzun</td>
<td>1996</td>
<td>Chile</td>
<td>Rev Med Chi 1996; 124: 1365–1367 (Spanish)</td>
</tr>
<tr>
<td>Gutierrez</td>
<td>1996</td>
<td>Chilean population ≥5 yrs old</td>
<td>Rev Med Chi 1996; 124: 1295–1306 (Spanish)</td>
</tr>
<tr>
<td>Enright</td>
<td>1996</td>
<td>USA; elderly Blacks</td>
<td>Chest 1996; 110: 1416–1424</td>
</tr>
<tr>
<td>Sirotkovic</td>
<td>1995</td>
<td>Croatia; school children from Dalmatia</td>
<td>Monaldi Arch Chest Dis 1995; 50: 258–263</td>
</tr>
</tbody>
</table>
ATS 1991 document [5], no race correction is used for TLC or RV in Hispanic or Native American subjects in North America. For African American, Asian American and East Indian subjects, race corrections of 0.88 for TLC, FRC, and 0.93 for RV, are used. Race corrections should not be used for RV/TLC.

Table 2 reports studies on reference equations published from 1993 to August 2004, and equations derived from a MEDLINE search under the keywords “reference equations” and “lung volumes”. Its purpose is to recognise and encourage the continuing interest of worldwide researchers in deriving and using reference equations.

In practice, many USA and European laboratories use the reference equations for TLC, FRC and RV recommended by the 1995 ATS/ERS workshop [7] or by the ECCS [8, 20].

**Diffusing capacity for carbon monoxide**

Selecting reference values for DLCO is more problematic than selecting reference values for spirometry because inter-laboratory differences are much larger for DLCO [38, 39]. Some of these differences can be attributed to the method of calculating DLCO and adjusting for haemoglobin concentration, carboxyhaemoglobin concentration and altitude. Laboratory directors should thoughtfully select reference values that match the numbers produced in their laboratories. Optimally, it would require individual laboratories to measure DLCO in a sample of healthy subjects and compare the results with several reference equations. At the very least, laboratory directors should be alert to frequent interpretations that do not match the clinical situation. Such mismatches may signal inappropriate reference values or problems with the DLCO measurement.

Predicted values for alveolar volume (VA) inspired volume (VI), DLCO and transfer coefficient of the lung for carbon monoxide (KCO) should be derived from the same source. As DLCO and KCO may be variably affected by factors previously described in this series of Task Force reports [4], a statement should be included describing which parameters might have been used to adjust the predicted values (e.g. VA, haemoglobin and carboxyhaemoglobin concentrations, and altitude).

---

**TABLE 1** (Continued)

<table>
<thead>
<tr>
<th>First author</th>
<th>Year</th>
<th>Country/area/nature of study</th>
<th>Journal</th>
</tr>
</thead>
</table>

**NHANES**: National Health and Nutrition Examination Survey; **ECCHS**: European Community Respiratory Health Survey; **MVV**: maximum voluntary ventilation.

**TABLE 2** Outcomes of a MEDLINE search under the keywords “reference equations” and “lung volumes”

<table>
<thead>
<tr>
<th>First author</th>
<th>Year</th>
<th>Country/area/nature of study</th>
<th>Journal</th>
</tr>
</thead>
<tbody>
<tr>
<td>TORRES</td>
<td>2003</td>
<td>Brazil; height and arm span in children</td>
<td>Pediatr Pulmonol 2003; 36: 202–208</td>
</tr>
<tr>
<td>VERMA</td>
<td>2003</td>
<td>India; dimensional statistics for estimation of lung volumes in children and adolescents</td>
<td>Anthropol Anz 2003; 61: 79–84</td>
</tr>
<tr>
<td>NEVE</td>
<td>2002</td>
<td>France; puberty and thoracic volume index</td>
<td>Eur Respir J 2002; 20: 1292–1298</td>
</tr>
<tr>
<td>ZHENG</td>
<td>2002</td>
<td>China; survey on clinical application</td>
<td>Zhonghua Jie He Hu Xi Za Zhi 2002; 25: 69–73 (Chinese)</td>
</tr>
<tr>
<td>COTES</td>
<td>2001</td>
<td>UK workers; body mass index</td>
<td>Thorax 2001; 56: 839–844</td>
</tr>
<tr>
<td>BAUR</td>
<td>1999</td>
<td>Germany; comparison of lung function reference values</td>
<td>Int Arch Occup Environ Health 1999; 72: 69–83</td>
</tr>
<tr>
<td>CHIN</td>
<td>1997</td>
<td>Singapore; nonsmoking adults</td>
<td>Respir Med 1997; 2: 143–149</td>
</tr>
<tr>
<td>ROSENTHAL</td>
<td>1993</td>
<td>UK; body plethysmographic gas volumes in prepubertal and pubertal school children in London</td>
<td>Thorax 1993; 48: 803–808</td>
</tr>
</tbody>
</table>
Table 3 shows studies on reference equations, published from 1995 to August 2004, and some derived from a MEDLINE search under the keywords “reference equations” and “diffusing capacity” or “diffusion”. Its purpose is to recognise and encourage the continuing interest of worldwide researchers in deriving and using reference equations.

A single “summary” prediction equation was proposed by the ERS [38] and suggested by the ATS [39]. At present, however, a single equation set for DL,CO cannot be recommended because of the relatively high inter-laboratory variability. Commonly used equations appear to be those from the 1993 ERS document [38] and those of CRAPO and MORRIS [40]. In Europe, equations from COTES et al. [41], PAOLETTI et al. [42], and ROCA et al. [43] are also used.

Table 4 provides a summary of the reference values used for general issues, spirometry, lung volumes and diffusing capacity.

### TYPES OF VENTILATORY DEFECTS

#### General issues

PFT interpretations should be clear, concise and informative. A mere statement of which values are normal or low is not helpful. Ideally, the principles of clinical decision making should be applied to the interpretation of the results of PFTs [44], where the post-test probability of disease is estimated after taking into consideration the pre-test probability of disease, the quality of the test results, the downside of a false-positive and false-negative interpretation, and, finally, the test results themselves and how they compare with reference values. This is often not possible because many, if not most, tests are interpreted in the absence of any clinical information. To improve this situation, it may be useful, whenever possible, to ask the physicians who are responsible for ordering tests to state the clinical question to be answered and, before testing, ask patients why they were sent for testing. Similarly, recording respiratory symptoms, such as cough, phlegm, wheezing and dyspnoea, as well as smoking status, and recent bronchodilator use could be helpful in this regard.

The interpretation will be most meaningful if the interpreter can address relevant clinical diagnoses, the chest radiograph appearance, the most recent haemoglobin value, and any suspicion of neuromuscular disease or upper airway obstruction (UAO).

#### Obstructive abnormalities

An obstructive ventilatory defect is a disproportionate reduction of maximal airflow from the lung in relation to the maximal volume (i.e. VC) that can be displaced from the lung [45–47]. It implies airway narrowing during exhalation and is defined by a reduced FEV1/VC ratio below the 5th percentile of the predicted value. A typical example is shown in figure 1a.

The earliest change associated with airflow obstruction in small airways is thought to be a slowing in the terminal portion of the spirogram, even when the initial part of the spirogram is barely affected [45–47]. This slowing of expiratory flow is most obviously reflected in a concave shape on the flow–volume curve. Quantitatively, it is reflected in a proportionally greater reduction in the instantaneous flow measured after 75% of the FVC has been exhaled (FEF75%) or in mean expiratory flow between 25% and 75% of FVC than in FEV1. However, abnormalities in these mid-range flow measurements during a forced exhalation are not specific for small airway disease in individual patients [48].

As airway disease becomes more advanced and/or more central airways become involved, timed segments of the spirogram such as the FEV1 will, in general, be reduced out of proportion to the reduction in VC.

<table>
<thead>
<tr>
<th>First author</th>
<th>Year</th>
<th>Country/area/nature of study</th>
<th>Journal</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEVE</td>
<td>2002</td>
<td>France; puberty and thoracic volume index</td>
<td>Eur Respir J 2002; 20: 1292–1298</td>
</tr>
<tr>
<td>ZHENG</td>
<td>2002</td>
<td>China, survey on clinical application</td>
<td>Zhonghua Jie He He Hu Xi Za Zhi 2002; 25: 69–73 (Chinese)</td>
</tr>
<tr>
<td>COTES</td>
<td>2001</td>
<td>UK workers; body mass index</td>
<td>Thorax 2001; 56: 839–844</td>
</tr>
<tr>
<td>JOHNSON</td>
<td>2000</td>
<td>USA; correction for VA for both DL,CO and Kco</td>
<td>Respir Med 2000; 94: 28–44</td>
</tr>
<tr>
<td>BAUR</td>
<td>1999</td>
<td>Germany; comparison of lung function reference values</td>
<td>Int Arch Occup Environ Health 1999; 72: 69–83</td>
</tr>
<tr>
<td>CHIN</td>
<td>1997</td>
<td>Singapore; nonsmoking adults</td>
<td>Respirology 1997; 2: 143–149</td>
</tr>
<tr>
<td>COLLARD</td>
<td>1996</td>
<td>Belgium; obstructive sleep apnoea and obesity</td>
<td>Chest 1996; 110: 1189–1193</td>
</tr>
</tbody>
</table>

Kco: transfer coefficient of the lung for carbon monoxide; TL: transfer factor of the lung; VA: alveolar volume; DL,CO: carbon monoxide diffusing capacity.
### TABLE 4 Summary of the usage of reference values

<table>
<thead>
<tr>
<th>Item</th>
<th>Reference values</th>
</tr>
</thead>
</table>
| **General**   | Predicted values should be obtained from studies of “normal” or “healthy” subjects with the same anthropometric (e.g. sex, age and height) and ethnic characteristics of the patient being tested.  
Height and weight should be measured for each patient at the time of testing.  
If possible, all parameters should be taken from the same reference source.  
When comparing selected reference equations with measurements performed on a sample of healthy subjects in a laboratory, it is suggested to choose the reference equation that provides the sum of residuals (observed – predicted computed for each adult subject, or log observed – log predicted for each subject in the paediatric age range) closest to zero.  
When using a set of reference equations, extrapolation beyond the size and age of the investigated subjects should be avoided.  
For each lung function index, values below the 5th percentile of the frequency distribution of values measured in the reference population are considered to be below the expected “normal range”. |
| **Spirometry**| In the USA, ethnically appropriate NHANES III reference equations published in 1999 for those aged 8–80 yrs, and the equations of Wang et al. [29] for children aged <8 yrs are recommended.  
In Europe, the ECCS combined reference equations published in 1993 [8] are often used for 18–70-yr-old people, and those from Quanjer et al. [30] for paediatric ages.  
Currently, a specific set of equations is not recommendable for use in Europe. A new Europe-wide study to derive updated reference equations for lung function is needed.  
Table 1 includes reference equations published from 1995 to August 2004. |
| **Lung volumes**| No specific set of equations can be recommended.  
In practice, many USA and European laboratories use the reference equations for TLC, FRC and RV recommended by the 1995 ATS/ERS workshop [7] or by the ECCS in 1993 [8].  
Table 2 reports studies on reference equations published from 1993 to August 2004. |
| **Diffusing capacity**| No specific set of equations is generally recommended.  
Commonly used equations appear to be those by the ECCS in 1993 [38] and those of Capo and Morris [40]. In Europe, equations from Gotes et al. [41], Paoletti et al. [42] and Rocca et al. [43] are also used.  
Table 3 shows studies on reference equations published from 1995 to August 2004. |


---

**FIGURE 1.** a, b) Examples of obstructive pulmonary defects with a low (a; forced expiratory volume in one second (FEV1) 38%; FEV1/vital capacity (VC) 46%; peak expiratory flow (PEF) 48%; total lung capacity (TLC) 101%) or normal (b; FEV1 57%; FEV1/VC 73%; PEF 43%; TLC 96%) ratio of FEV1/VC. In both cases, TLC is normal, and flows are less than expected over the entire volume range.  
c) Example of a typical restrictive defect (FEV1 66%; FEV1/VC 80%; PEF 79%; TLC 62%). The TLC is low and flow is higher than expected at a given lung volume.  
d) Example of a typical mixed defect characterised by a low TLC and a low FEV1/VC ratio (FEV1 64%; FEV1/VC 64%; PEF 82%; TLC 72%). – – – –: predicted flow–volume curves; ———: observed inspiratory and expiratory flow–volume curves (as indicated in a).
Special attention must be paid when FEV1 and FVC are concomitantly decreased and the FEV1/FVC ratio is normal or almost normal. This pattern most frequently reflects failure of the patient to inhale or exhale completely. It may also occur when the flow is so slow that the subject cannot exhale long enough to empty the lungs to RV. In this circumstance, the flow–volume curve should appear concave toward the end of the manoeuvre. TLC will be normal and FEF75 will be low. Measurement of slow VC (inspiratory or expiratory) may then give a more correct estimate of the FEV1/VC ratio. Another possible cause of this pattern is patchy collapse of small airways early in exhalation [8, 49–52]. Under these conditions, TLC may be normal, but RV is ordinarily increased. A typical example is shown in figure 1b. When this pattern is observed in a patient performing a maximal, sustained effort, it may be useful to repeat spirometry after treatment with an inhaled bronchodilator. Significant improvement in the FEV1, FVC or both would suggest the presence of reversible airflow obstruction.

Apart from this unusual circumstance, measurement of lung volumes is not mandatory to identify an obstructive defect. It may, however, help to disclose underlying disease and its functional consequences. For example, an increase in TLC, RV or the RV/TLC ratio above the upper limits of normal variability may suggest the presence of emphysema, bronchial asthma or other obstructive diseases [47], as well as the degree of lung hyperinflation.

Airflow resistance is rarely used to identify airflow obstruction in clinical practice. It is more sensitive for detecting narrowing of extrathoracic or large central intrathoracic airways than of more peripheral intrathoracic airways [47]. It may be useful in patients who are unable to perform a maximal forced expiratory manoeuvre.

**Restrictive abnormalities**

A restrictive ventilatory defect is characterised by a reduction in TLC below the 5th percentile of the predicted value, and a normal FEV1/VC. A typical example is shown in figure 1c. The presence of a restrictive ventilatory defect may be suspected when VC is reduced, the FEV1/VC is increased (>85–90%) and the flow–volume curve shows a convex pattern. Once again, the pattern of a reduced VC and a normal or even slightly increased FEV1/VC is often caused by submaximal inspiratory or expiratory efforts and/or patchy peripheral airflow obstruction, and a reduced VC by itself does not prove a restrictive ventilatory defect. It is associated with a low TLC no more than half the time [53, 54].

Pneumothorax and noncommunicating bullae are special cases characterised by a normal FEV1/VC and TLC measured in a body plethysmograph, but low FEV1 and VC values. In these conditions, TLC assessed by gas dilution techniques will be low.

A low TLC from a single-breath test (such as VA from the DLCO test) should not be interpreted as demonstrating restriction, since such measurements systematically underestimate TLC [55]. The degree of underestimation increases as airflow obstruction worsens. In the presence of severe airflow obstruction, TLC can be underestimated by as much as 3 L, greatly increasing the risk of misclassification of the type of PFT abnormality [55, 56]. A method of adjusting the single-breath VA for the effect of airway obstruction has been published, but needs further validation [57].

**Mixed abnormalities**

A mixed ventilatory defect is characterised by the coexistence of obstruction and restriction, and is defined physiologically when both FEV1/VC and TLC are below the 5th percentiles of their relevant predicted values. Since VC may be equally reduced in both obstruction and restriction, the presence of a restrictive component in an obstructed patient cannot be inferred from simple measurements of FEV1 and VC. A typical example is presented in figure 1d. If FEV1/VC is low and the largest measured VC (pre- or post-bronchodilator VC or VI in the DLCO test) is below its lower limits of normal (LLN), and there is no measurement of TLC by body plethysmography, one can state that the VC was also reduced, probably due to hyperinflation, but that a superimposed restriction of lung volumes cannot be ruled out [58]. Conversely, when FEV1/VC is low and VC is normal, a superimposed restriction of lung volumes can be ruled out [53, 54].

Table 5 shows a summary of the types of ventilatory defects and their diagnoses.

**COMMENTS ON INTERPRETATION AND PATTERNS OF DYSFUNCTION**

The definition of an obstructive pulmonary defect given in the present document is consistent with the 1991 ATS statement on interpretation [5], but contrasts with the definitions suggested by both Global Initiative for Chronic Obstructive Lung Disease (GOLD) [59] and ATS/ERS guidelines on COPD [60], in that FEV1 is referred to VC rather than just FVC and the cut-off value of this ratio is set at the 5th percentile of the normal distribution rather than at a fixed value of 0.7. This committee feels that the advantage of using VC in place of FVC is that the ratio of FEV1 to VC is capable of accurately identifying more obstructive patterns than its ratio to FVC, because FVC is more dependent on flow and volume histories [61]. In contrast with a fixed value of 0.7, the use of the 5th percentile does not lead to an overestimation of the ventilatory defect in older people with no history of exposure to noxious particles or gases [62].

The assumption that a decrease in major spirometric parameters, such as FEV1, VC, FEV1/VC and TLC, below their relevant 5th percentiles is consistent with a pulmonary defect is a useful simple approach in clinical practice. Problems arise, however, when some or all of these variables lie near their upper limits of normal or LLN. In these cases, a literal interpretation of the functional pattern is too simplistic and could fail to properly describe the functional status.

The current authors suggest that additional studies should be done in these circumstances if they are indicated by the clinical problem being addressed. Such tests could include bronchodilator response, DLCO, gas-exchange evaluation, measurement of respiratory muscle strength or exercise testing.

Caution is also recommended when TLC is at the LLN and coexists with a disease expected to lead to lung restriction. A typical example is lung resection. The expected restrictive defect would be difficult to prove on the simple basis of TLC as per cent of predicted if the latter remains above the 5th
percentile of predicted as a result of subsequent lung growth or of a large TLC before surgery. Similar care must be taken in cases where diseases with opposing effects on TLC coexist, such as interstitial lung disease (ILD) and emphysema.

While patterns of physiological abnormalities can be recognised, they are seldom pathognomonic for a specific disease entity. The types of clinical illness most likely to produce an observed set of physiological disturbances can be pointed out. Regardless of the extent of testing, it is important to be conservative in suggesting a specific diagnosis for an underlying disease process based only on pulmonary function abnormalities.

The VC, FEV\textsubscript{1}, FEV\textsubscript{1}/VC ratio and TLC are the basic parameters used to properly interpret lung function (fig. 2). Although FVC is often used in place of VC, it is preferable to use the largest available VC, whether obtained on inspiration (IVC), slow expiration (SVC) or forced expiration (i.e. FVC). The FVC is usually reduced more than IVC or SVC in airflow obstruction [61]. The FEV\textsubscript{6} may be substituted for VC if the TLC <5th percentile of predicted measurement with the predicted value adjusted for haemoglobin. In the mixed defect group, the measurement of absolute lung volumes may assist in diagnosing extrathoracic airway obstruction. Confirmation of airway obstruction requires measurement of lung volumes. Measurement of absolute lung volumes may assist in the diagnosis of emphysema, bronchial asthma and chronic bronchitis. It may also be useful in assessing lung hyperinflation. Measurements of airflow resistance may be helpful in patients who are unable to perform spirometric manoeuvres.

A concomitant decrease in FEV\textsubscript{1} and VC is most commonly caused by poor effort, but may rarely reflect airflow obstruction. A low TLC from a single-breath test should not be seen as evidence of restriction.

Although FVC is often used in place of VC, it is preferable to use the largest available VC, whether obtained on inspiration (IVC), slow expiration (SVC) or forced expiration (i.e. FVC). The FVC is usually reduced more than IVC or SVC in airflow obstruction [61]. The FEV\textsubscript{6} may be substituted for VC if the TLC <5th percentile of predicted measurement with the predicted value adjusted for haemoglobin. In the mixed defect group, the measurement of absolute lung volumes may assist in diagnosing extrathoracic airway obstruction. Confirmation of airway obstruction requires measurement of lung volumes. Measurement of absolute lung volumes may assist in the diagnosis of emphysema, bronchial asthma and chronic bronchitis. It may also be useful in assessing lung hyperinflation. Measurements of airflow resistance may be helpful in patients who are unable to perform spirometric manoeuvres.

A low TLC from a single-breath test should not be seen as evidence of restriction.

A reduced VC does not prove a restrictive pulmonary defect. It may be suggestive of lung restriction when FEV\textsubscript{1}/VC is normal or increased.

TABLE 5 Types of ventilatory defects and their diagnoses

<table>
<thead>
<tr>
<th>Abnormality</th>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obstruction</td>
<td>FEV\textsubscript{1}/VC &lt;5th percentile of predicted</td>
</tr>
<tr>
<td></td>
<td>A decrease in flow at low lung volume is not specific for small airway disease in individual patients</td>
</tr>
<tr>
<td></td>
<td>A concomitant decrease in FEV\textsubscript{1} and VC is most commonly caused by poor effort, but may rarely reflect airflow obstruction.</td>
</tr>
<tr>
<td></td>
<td>Confirmation of airway obstruction requires measurement of lung volumes</td>
</tr>
<tr>
<td></td>
<td>Measurement of absolute lung volumes may assist in the diagnosis of emphysema, bronchial asthma and chronic bronchitis. It may also be useful in assessing lung hyperinflation. Measurements of airflow resistance may be helpful in patients who are unable to perform spirometric manoeuvres.</td>
</tr>
<tr>
<td>Restriction</td>
<td>TLC &lt;5th percentile of predicted</td>
</tr>
<tr>
<td></td>
<td>A reduced VC does not prove a restrictive pulmonary defect. May be suggestive of lung restriction when FEV\textsubscript{1}/VC is normal or increased</td>
</tr>
<tr>
<td>Mixed defect</td>
<td>FEV\textsubscript{1}/VC and TLC &lt;5th percentile of predicted</td>
</tr>
</tbody>
</table>

FEV\textsubscript{1}: forced expiratory volume in one second; VC: vital capacity; TLC: total lung capacity.

![FIGURE 2](image-url) A simplified algorithm that may be used to assess lung function in clinical practice. It presents classic patterns for various pulmonary disorders. As in any such diagram, patients may or may not present with the classic patterns, depending on their illnesses, severity and lung function prior to the disease onset (e.g. did they start with a vital capacity (VC) close to the upper or lower limits of normal (LLN)). The decisions about how far to follow this diagram are clinical, and will vary depending on the questions being asked and the clinical information available at the time of testing. The forced expiratory volume in one second (FEV\textsubscript{1})/VC ratio and VC should be considered first. Total lung capacity (TLC) is necessary to confirm or exclude the presence of a restrictive defect when VC is below the LLN. The algorithm also includes diffusing capacity for carbon monoxide (DL\textsubscript{CO}) measurement with the predicted value adjusted for haemoglobin. In the mixed defect group, the DL\textsubscript{CO} patterns are the same as those for restriction and obstruction. This flow chart is not suitable for assessing the severity of upper airway obstruction. PV: pulmonary vascular; CW: chest wall; NM: neuromuscular; ILD: interstitial lung diseases; CB: chronic bronchitis.

accompanied by an FEV\textsubscript{1} within the normal range is unclear. This pattern is probably due to “dysanaptic” or unequal growth of the airways and lung parenchyma [67] (referred to in a previous ATS document as a possible physiological variant when FEV\textsubscript{1} was \( \geq 100\% \) pred [5]). Whether this pattern represents airflow obstruction will depend on the prior
probability of obstructive disease and possibly on the results of additional tests, such as bronchodilator response, DLCO, gas-exchange evaluation, and measurement of muscle strength or exercise testing. Expiratory flow measurements other than the FEV1 and FEV1/VC should be considered only after determining the presence and clinical severity of obstructive impairment using the basic values mentioned previously. When the FEV1 and FEV1/VC are within the expected range, the clinical significance of abnormalities in flow occurring late in the maximal expiratory flow–volume curve is limited. In the presence of a borderline value of FEV1/VC, however, these tests may suggest the presence of airway obstruction. The same is true for average flows, such as mid-expiratory flow (MEF25–75%), especially in children with cystic fibrosis [68, 69]. Even with this limited use, the wide variability of these tests in healthy subjects must be taken into account in their interpretation.

The maximal voluntary ventilation (MVV) is not generally included in the set of lung function parameters necessary for diagnosis or follow-up of the pulmonary abnormalities because of its good correlation with FEV1 [70]. However, it may be of some help in clinical practice. For example, a disproportionate decrease in MVV relative to FEV1 has been reported in neuromuscular disorders [71, 72] and UAO [73]. In addition, it is also used in estimating breathing reserve during maximal exercise [74], although its application may be of limited value in mild-to-moderate COPD [75, 76]. For these purposes, the current authors suggest that MVV should be measured rather than estimated by multiplying FEV1 by a constant value, as is often done in practice.

**SEVERITY CLASSIFICATION**

A method of categorising the severity of lung function impairment based on the FEV1 % pred is given in table 6. It is similar to several previous documents, including GOLD [59], ATS 1986 [77], ATS 1991 [5], and the American Medical Association (AMA) [78]. The number of categories and the exact cut-off points are arbitrary.

Severity scores are most appropriately derived from studies that relate pulmonary function test values to independent indices of performance, such as ability to work and function in daily life, morbidity and prognosis [79–82]. In general, the ability to work and function in daily life is related to pulmonary function, and pulmonary function is used to rate impairment in several published systems [77–79, 83]. Pulmonary function level is also associated with morbidity, and the patients with lower function have more respiratory complaints [82].

Lung function level is also associated with prognosis, including a fatal outcome from heart as well as lung disease [84, 85], even in patients who have never smoked [86]. In the Framingham study, VC was a major independent predictor of cardiovascular morbidity and mortality [84, 85]. In several occupational cohorts, FEV1 and FEV1/FVC were independent predictors of all-cause or respiratory disease mortality [87–89]. In addition, a meta-analysis of mortality in six surveys in various UK working populations showed that the risk of dying from COPD was related to the FEV1 level. In comparison to those whose FEV1 at an initial examination was within 1 SD of average, those whose FEV1 was >2 SD below average were 12 times more likely to die of COPD, more than 10 times as likely to die of non-neoplastic respiratory disease, and more than twice as likely to die of vascular disease over a 20-yr follow-up period [90]. Although there is good evidence that FEV1 correlates with the severity of symptoms and prognosis in many circumstances [79, 82, 90], the correlations do not allow one to accurately predict symptoms or prognosis for individual patients.

The DLCO is also an important predictor of mortality both in the general population [91] and in patients after pulmonary resection [92].

Though the FEV1 % pred is generally used to grade severity in patients with obstructive, restrictive and mixed pulmonary defects, it has little applicability to patients with UAO, such as tracheal stenosis, where obstruction could be life-threatening and yet be classified as mildly reduced by this scheme. In addition, there is little data documenting the performance of other functional indexes, such as FRC in airflow obstruction or TLC in lung restriction as indices to categorise severity of impairment.

VC is reduced in relation to the extent of loss of functioning lung parenchyma in many nonobstructive lung disorders. It is also of some use in assessing respiratory muscle involvement in certain neuromuscular diseases. VC may be only slightly impaired in diffuse interstitial diseases of sufficient severity to lead to marked loss of diffusing capacity and severe blood gas abnormalities [63]. The onset of a severe respiratory problem in patients with a rapidly progressive neuromuscular disease may be associated with only a small decrement in VC [47, 93].

FEV1 and FVC may sometimes fail to properly identify the severity of ventilatory defects, especially at the very severe stage for multiple reasons. Among them are the volume history effects of the deep breath preceding the forced expiratory manoeuvre on the bronchial tone and, thus, calibre [94–98], and the inability of these parameters to detect whether tidal breathing is flow limited or not [99–102]. The FEV1/VC ratio should not be used to determine the severity of an obstructive disorder, until new research data are available. Both the FEV1 and VC may decline with the progression of disease, and an FEV1/VC of 0.5/1.0 indicates more impairment than one of 2.0/4.0, although the ratio of both is 50%. While the FEV1/VC ratio should not routinely be used to determine the severity of an obstructive disorder, it may be of value when persons having genetically large lungs develop obstructive disease. In

---

**TABLE 6** Severity of any spirometric abnormality based on the forced expiratory volume in one second (FEV1)

<table>
<thead>
<tr>
<th>Degree of severity</th>
<th>FEV1 % pred</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild</td>
<td>&gt;70</td>
</tr>
<tr>
<td>Moderate</td>
<td>60–69</td>
</tr>
<tr>
<td>Moderately severe</td>
<td>50–59</td>
</tr>
<tr>
<td>Severe</td>
<td>35–49</td>
</tr>
<tr>
<td>Very severe</td>
<td>&lt;35</td>
</tr>
</tbody>
</table>

% pred: % predicted.
these cases, the FEV1/VC ratio may be very low (60%), when the FEV1 alone is within the mild category of obstruction (i.e., >70% pred).

Recent studies have stressed the importance of additional measurements in assessing the severity of the disease. For example, when airflow obstruction becomes severe, FRC, RV, TLC and RV/TLC tend to increase as a result of decreased lung elastic recoil and/or dynamic mechanisms [47, 103, 104]. The degree of hyperinflation parallels the severity of airflow obstruction [58]. On one hand, lung hyperinflation is of benefit because it modulates airflow obstruction, but, on the other hand, it causes dyspnoea because of the increased elastic load on inspiratory muscles [47]. In a recent investigation, resting lung hyperinflation, measured as inspiratory capacity (IC)/TLc, was an independent predictor of respiratory and all-cause mortality in COPD patients [105]. In addition, in either severe obstructive or restrictive diseases, tidal expiratory flow often impinges on maximum flow [98, 99, 102]. This condition, denoted as expiratory flow limitation during tidal breathing (EFL), is relatively easy to measure in practice by comparing tidal and forced expiratory flow–volume loops. Its clinical importance is that it contributes to increased dyspnoea [100], puts the inspiratory muscles at a mechanical disadvantage [43] and causes cardiovascular side-effects [106]. Although there currently isn’t sufficient evidence to recommend the routine use of measurements of hyperinflation or EFL to score the severity of lung function impairment, they may be helpful in patients with disproportionate differences between spirometric impairment and dyspnoea.

Finally, the reported increase in RV in obstruction is deemed to be a marker of airway closure [47, 103]. Although its clinical relevance remains uncertain, especially with regard to assessment of severity, RV may be useful in special conditions, including predicting the likelihood of lung function improvement after lung volume-reduction surgery [104].

Table 7 shows the summary of the considerations for severity classification.

**BRONCHODILATOR RESPONSE**

Bronchial responsiveness to bronchodilator medications is an integrated physiological response involving airway epithelium, nerves, mediators and bronchial smooth muscle. Since the within-individual difference in response to a bronchodilator is variable, the assumption that a single test of bronchodilator response is adequate to assess both the underlying airway responsiveness and the potential for therapeutic benefits of bronchodilator therapy is overly simplistic [107]. Therefore, the current authors feel that the response to a bronchodilator agent can be tested either after a single dose of a bronchodilator agent in the PFT laboratory or after a clinical trial conducted over 2–8 weeks.

The correlation between bronchoconstriction and bronchodilator response is imperfect, and it is not possible to infer with certainty the presence of one from the other. There is no consensus about the drug, dose or mode of administering a bronchodilator in the laboratory. However, when a metered dose inhaler is used, the following procedures are suggested in order to minimise differences within and between laboratories. Short-acting β2-agonists, such as salbutamol, are recommended. Four separate doses of 100 μg should be used when given by metered dose inhaler using a spacer. Tests should be repeated after a 15-min delay. If a bronchodilator test is performed to assess the potential therapeutic benefits of a specific drug, it should be administered in the same dose and by the same route as used in clinical practice, and the delay between administration and repeated spirometric measurements should reflect the reported time of onset for that drug.

The first step in interpreting any bronchodilator test is to determine if any change greater than random variation has occurred. The per cent change in FVC and FEV1 after bronchodilator administration in general population studies [108–110] and patient populations [101, 111–113] are summarised in table 8. Studies show a tendency for the calculated bronchodilator response to increase with decreasing baseline VC or FEV1, regardless of whether the response was considered as an absolute change or as a per cent of the initial value. Bronchodilator responses in patient-based studies are, therefore somewhat higher than those in general population studies.

There is no clear consensus about what constitutes reversibility in subjects with airflow obstruction [111, 114]. In part, this is because there is no consensus on how a bronchodilator response should be expressed, the variables to be used, and, finally, the kind, dose and inhalation mode of bronchodilator agent. The three most common methods of expressing bronchodilator response are per cent of the initial spirometric value, per cent of the predicted value, and absolute change.

Expressing the change in FEV1 and/or FVC as a per cent of predicted values has been reported to have advantages over per cent change from baseline [115]. When using per cent change from baseline as the criterion, most authorities require a 12–15% increase in FEV1 and/or FVC as necessary to define a meaningful response. Increments of <8% (or <150 mL) are likely to be within measurement variability [107, 115]. The current authors recommend using the per cent change from baseline and absolute changes in FEV1 and/or FVC in an individual subject to identify a positive bronchodilator response. Values >12% and 200 mL compared with baseline during a single testing session suggest a “significant” bronchodilatation. If the change in FEV1 is not significant, a
have assessed the utility of MEF 25–75%. The results were proposed to solve this problem [116, 117]. At least two studies patients [101, 109, 111, 114]. In addition, it must be also shown to be capable of clearly separating the two classes of patients, the response to a bronchodilator has never been flow and volume after inhaling a dilator agent than COPD instance, even if asthmatics tend to show a larger increase in and depends on the reasons for undertaking the test. For important. This aspect of the interpretation is harder to define the next step is to determine if this change is clinically if the change is above the threshold of natural variability, then upper 95% CI change after placebo. Per cent change from baseline CL for absolute change. 95% CL per cent change of baseline.

<table>
<thead>
<tr>
<th>Population Studies</th>
<th>Agent/mode of delivery</th>
<th>FVC</th>
<th>FEV1</th>
<th>MEF25-75% or MEF50%</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1063 subjects 8–75 yrs of age; general population [108]</td>
<td>IP 2 puffs via MDI</td>
<td>10.7% (0.40 L)</td>
<td>7.7% (0.31 L)</td>
<td>20%</td>
<td>95th percentile for per cent change from baseline</td>
</tr>
<tr>
<td>2609 subjects; random sample of 3 areas in Alberta, Canada [109]</td>
<td>TB 500 μg via spacer</td>
<td>Males 9% (0.34 L); females 9% (0.22 L)</td>
<td></td>
<td></td>
<td>95th percentile for per cent change from baseline in asymptomatic never-smokers with FEV1 &gt;80% pred</td>
</tr>
<tr>
<td>75 selected normal subjects [110]</td>
<td>Two puffs via MDI</td>
<td>5.1% (0.23 L)</td>
<td>10.1% (0.36 L)</td>
<td>48.3%</td>
<td>Upper 95% CL (two-tailed) for per cent change from baseline</td>
</tr>
</tbody>
</table>

The MEF25–75% is a highly variable spirometric test, in part because it depends on FVC, which increases with expiratory time in obstructed subjects. If FVC changes, post-bronchodilator MEF25–75% is not comparable with that measured before the bronchodilator. Volume adjustment of MEF25–75% has been proposed to solve this problem [116, 117]. At least two studies have assessed the utility of MEF25–75%. The results were disappointing; only 8% of asthmatics [117] and 7% of patients with COPD were identified as outside the expected range by MEF25–75% criteria alone. Tests such as the FEV1/VC ratio and instantaneous flows measured at some fraction of the VC may also be misleading in assessing bronchodilator response if expiratory time changes are not considered and if flows are not measured at the same volume below TLC.

If the change is above the threshold of natural variability, then the next step is to determine if this change is clinically important. This aspect of the interpretation is harder to define and depends on the reasons for undertaking the test. For instance, even if asthmatics tend to show a larger increase in flow and volume after inhaling a dilator agent than COPD patients, the response to a bronchodilator has never been shown to be capable of clearly separating the two classes of patients [101, 109, 111, 114]. In addition, it must be also acknowledged that responses well below the significant thresholds may be associated with symptom improvement and patient performance [118]. The possible reasons are discussed as follows.

Quite often, responses to bronchodilator therapy are unpredictably underestimated by FEV1 and/or FVC in comparison to airway resistance or flow measured during forced expiratory manoeuvres initiated from a volume below TLC (partial expiratory flow–volume manoeuvres) in both healthy subjects and patients with chronic airflow obstruction [8, 101, 102, 119–122]. These findings are probably due to the fact that deep inhalations tend to reduce airway calibre, especially after a bronchodilator [101, 120]. In patients with airflow obstruction, the increase in expiratory flow after bronchodilation is often associated with a decrease in FRC or an increase in IC of similar extent at rest and during exercise [101, 123]. The improvement of the lung function parameters in the tidal breathing range and not following a deep breath may explain the decrease in shortness of breath after inhaling a bronchodilator, despite no or minimal changes in FEV1 and/or FVC. Short-term intra-individual variabilities for partial flows and IC have been reported [101]. Therefore, the lack of increase of FEV1 and/or FVC after a bronchodilator is not a good reason to avoid 1–8-week clinical trial with bronchoactive medication.

An isolated increase in FVC (>12% of control and >200 mL) not due to increased expiratory time after salbutamol is a sign of...
Therefore, an increased ratio of FEV\textsubscript{1} divided by PEF is necessary to evaluate for central or upper airway obstruction. It is critical that the patient’s inspiratory and expiratory efforts are near maximal and the technician should confirm this in the quality notes. When patient effort is good, the pattern of a repeatable plateau of forced inspiratory flow, with or without a forced expiratory plateau, suggests a variable extrathoracic central or upper airway obstruction (fig. 3). Conversely, the pattern of a repeatable plateau of forced expiratory flow, along with the lack of a forced inspiratory plateau suggests a variable, intrathoracic central or upper airway obstruction. The pattern of a repeatable plateau at a similar flow in both forced inspiratory and expiratory flows suggests a fixed central or upper airway obstruction (fig. 3).

In general, maximum inspiratory flow is largely decreased with an extrathoracic airway obstruction, because the pressure surrounding the airways (which is almost equal to atmospheric) cannot oppose the negative intraluminal pressure generated with the inspiratory effort. In contrast, it is little affected by an intrathoracic airway obstruction, for the pressure surrounding the intrathoracic airways (which is close to pleural pressure) strongly opposes the negative intraluminal pressure on inspiration, thus limiting the effects of the obstruction on flow. With unilateral main bronchus obstruction, a rare event, maximum inspiratory flow tends to be higher at the beginning than towards the end of the forced inspiration because of a delay in gas filling (fig. 4).

Maximum expiratory flow at high lung volume (especially peak flow) is generally decreased in both intrathoracic and extrathoracic lesions [126–129]. In contrast, maximum flows may be normal in the presence of a variable lesion, such as vocal cord paralysis. Flow oscillations (saw-tooth pattern) may be occasionally observed on the either inspiratory or expiratory phase, and probably represent a mechanical instability of the airway wall.

The effects of anatomical or functional lesions on maximum flows depend on the site of the obstruction, kind of lesion (variable or fixed) and the extent of anatomical obstruction [61, 127, 130]. Typical cases of extra- and intrathoracic central or upper airway obstruction are reported in figures 3 and 4. The absence of classic spirometric patterns for central airway

### Table 9

Summary of the procedures relating to bronchodilator response

<table>
<thead>
<tr>
<th>Procedures suggested to minimise differences within and between laboratories</th>
<th>Assess lung function at baseline and repeatable maximal forced inspiratory and expiratory flow–volume loops (Fig. 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administer salbutamol in four separate doses of 100 µg through a spacer</td>
<td>Assesses a repeatable plateau of forced expiratory flow, along with the lack of a forced inspiratory plateau.</td>
</tr>
<tr>
<td>Re-assess lung function after 15 min.</td>
<td>The pattern of a repeatable plateau at a similar flow in both forced inspiratory and expiratory flows suggests a fixed central or upper airway obstruction.</td>
</tr>
<tr>
<td>An increase in FEV\textsubscript{1} and/or FVC (&gt;12%) of control and (&gt;200) mL constitutes a positive bronchodilator response.</td>
<td></td>
</tr>
<tr>
<td>In the absence of a significant increase in FEV\textsubscript{1} and/or FVC, an improvement in lung function parameters within the tidal breathing range, such as increased partial flows and decrease of lung hyperinflation, may explain a decrease in dyspnoea.</td>
<td></td>
</tr>
<tr>
<td>The lack of a bronchodilator response in the laboratory does not preclude a clinical response to bronchodilator therapy.</td>
<td></td>
</tr>
</tbody>
</table>

FEV\textsubscript{1}: forced expiratory volume in one second; FVC: forced vital capacity.

bronchodilation [124]. This may, in part, be related to the fact that deep inhalations tend to reduce airway calibre and/or airway wall stiffness, especially after a bronchodilator [101, 120].

Table 9 shows a summary of the suggested procedures for laboratories relating to bronchodilator response.

### CENTRAL AND UPPER AIRWAY OBSTRUCTION

Central airway obstruction and UAO may occur in the extrathoracic (pharynx, larynx, and extrathoracic portion of the trachea) and intrathoracic airways (intrathoracic trachea and main bronchi). This condition does not usually lead to a decrease in FEV\textsubscript{1} and/or VC, but PEF can be severely affected. Therefore, an increased ratio of FEV\textsubscript{1} divided by PEF (mL-L\textsuperscript{-1}·min\textsuperscript{-1}) can alert the clinician to the need for an inspiratory and expiratory flow–volume loop [125]. A value \(>8\) suggests central or upper airway obstruction may be present [126]. Poor initial effort can also affect this ratio.

At least three maximal and repeatable forced inspiratory and forced expiratory flow–volume curves are necessary to
obstruction does not accurately predict the absence of pathology. As a result, clinicians need to maintain a high degree of suspicion for this problem, and refer suspected cases for visual inspection of the airways. The authors feel that, although maximum inspiratory and expiratory flow–volume loops are of great help to alert clinicians to the possibility of central or upper airway obstruction, endoscopic and radiological techniques are the next step to confirm the dysfunction.

The parameters presented in table 10 may help to distinguish intrathoracic from extrathoracic airway obstructions.

Table 11 gives a summary of the relevant issues concerning UAO.

### INTERPRETATION OF CHANGE IN LUNG FUNCTION

Evaluation of an individual’s change in lung function following an intervention or over time is often more clinically valuable than a single comparison with external reference (predicted) values. It is not easy to determine whether a measured change reflects a true change in pulmonary status or is only a result of test variability. All lung function measurements tend to be more variable when made weeks to months apart than when repeated at the same test session or even daily [25, 131]. The short-term repeatability of tracked parameters should be measured using biological controls. This is especially important for the Dl,CO [132, 133], since small errors in measurements of inspiratory flows or exhaled gas concentrations translate into large Dl,CO errors. The variability of lung volume measurements has recently been reviewed [134].

The optimal method of expressing the short-term variability (measurement noise) is to calculate the coefficient of repeatability (CR) instead of the more popular coefficient of variation [135]. Change measured for an individual patient that falls outside the CR for a given parameter may be considered significant. The CR may be expressed as an absolute value (such as 0.33 L for FEV1 or 5 units for Dl,CO) [136] or as a percentage of the mean value (such as 11% for FEV1) [137].

It is more likely that a real change has occurred when more than two measurements are performed over time. As shown in table 12, significant changes, whether statistical or biological,
vary by parameter, time period and the type of patient. When there are only two tests available to evaluate change, the large variability necessitates relatively large changes to be confident that a significant change has in fact occurred. Thus, in subjects with relatively “normal” lung function, year-to-year changes in FEV1 over 1 yr should exceed 15% before confidence can be given to the opinion that a clinically meaningful change has occurred [5].

For tracking change, FEV1 has the advantage of being the most repeatable lung function parameter and one that measures changes in both obstructive and restrictive types of lung disease. Two-point, short-term changes of >12% and >0.2 L in the FEV1 are usually statistically significant and may be clinically important. Changes slightly less than these may, perhaps, be equally significant, depending on the reproducibility of the pre- and post-bronchodilator results. Other parameters such as VC, IC, TLC and DLCO may also be tracked in patients with ILD or severe COPD [138, 140–142]. Tests like VC and FVC may be relevant to COPD because they may increase when FEV1 does not, and changes in DLCO, in the absence of change in spirometry variables, may be clinically important. Again, when too many indices of lung function are tracked simultaneously, the risk of false-positive indications of change increases.

The clinician seeing the patient can often interpret results of serial tests in a useful manner, which is not reproducible by any simple algorithm. Depending on the clinical situation, statistically insignificant trends in lung function may be meaningful to the clinician. For example, seemingly stable test results may provide reassurance in a patient receiving therapy for a disease that is otherwise rapidly progressive. The same test may be very disappointing if one is treating a disorder that is expected to improve dramatically with the therapy prescribed. Conversely, a statistically significant change may be of no clinical importance to the patient. The largest errors occur in attempting to interpret serial changes in subjects without disease, because test variability will usually far exceed the true annual decline, and reliable rates of change for an individual subject cannot be calculated without prolonged follow-up [143].

Test variability can be reduced when lung function standards and guidelines are followed strictly. Simple plots (i.e. trending) of lung function with time can provide additional information to help differentiate true change in lung function from noise. Measuring decline in lung function as a means of identifying individuals (such as smokers) who are losing function at excessive rates has been proposed. However, establishing an accelerated rate of loss in an individual is very difficult, and requires many measurements over several years with meticulous quality control of the measurements.

Table 13 shows a summary of the considerations involved in interpreting lung function changes.

**DLCO INTERPRETATION**

The lower 5th percentile of the reference population should be used as LLN for DLCO and KCO (if the latter is used). Table 14 presents a scheme to grade the severity of reductions in DLCO.

<table>
<thead>
<tr>
<th>Degree of severity</th>
<th>DLCO % pred</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild</td>
<td>&gt;60% and &lt;LLN</td>
</tr>
<tr>
<td>Moderate</td>
<td>40–60%</td>
</tr>
<tr>
<td>Severe</td>
<td>&lt;40</td>
</tr>
</tbody>
</table>

% pred: % predicted; LLN: lower limits of normal.

The pathophysiological importance of this test has been recently reviewed [144, 145].

Interpreting the DLCO, in conjunction with spirometry and lung volumes assessment, may assist in diagnosing the underlying disease (fig. 2). For instance, normal spirometry and lung volumes associated with decreased DLCO may suggest anaemia, pulmonary vascular disorders, early ILD or early emphysema. In the presence of restriction, a normal DLCO may be consistent with chest wall or neuromuscular disorders, whereas a decrease suggests ILDs. In the presence of airflow obstruction, a decreased DLCO suggests emphysema [146], but airway obstruction and a low DLCO are also seen in lymphangioleiomyomatosis [147]. Patients with ILD, sarcoidosis and pulmonary fibrosis usually have a low DLCO [135–137, 140]. A low DLCO is also seen in patients with chronic pulmonary embolism, primary pulmonary hypertension [148], and other pulmonary vascular diseases. These patients may or may not also have restriction of lung volumes [149].

A high DLCO is associated with asthma [150], obesity [151] and intrapulmonary haemorrhage [152].

Adjustments of DLCO for changes in haemoglobin and carboxyhaemoglobin are important, especially in situations where patients are being monitored for possible drug toxicity, and where haemoglobin is subject to large shifts (e.g. chemotherapy for cancer).

Adjusting DLCO for lung volume using DLCO/VA or DLCO/TLC is controversial [153, 154]. Conceptually, a loss of DLCO that is much less than a loss of volume (low DLCO but high DLCO/VA) might suggest an extraparenchymal abnormality, such as a pneumonectomy or chest wall restriction, whereas a loss of DLCO that is much greater than a loss of volume (low DLCO and low DLCO/VA) might suggest parenchymal abnormalities. The relationship between DLCO and lung
volume, however, is not linear and markedly less than 1:1, so these simple ratios as traditionally reported do not provide an appropriate way to normalise $D_{L, CO}$ for lung volume [154–159]. Nonlinear adjustments may be considered, but their clinical utility must be established before they can be recommended. Meanwhile, it is advisable to keep examining $D_{L, CO}$ in conjunction with spirometry and lung volumes may assist in diagnosing the underlying disease (fig. 2).

The relationship between $D_{L, CO}$ and lung volume is not linear, so $D_{L, CO}/VA$ or $D_{L, CO}/TLC$ do not provide an appropriate way to normalise $D_{L, CO}$ for lung volume. Nonlinear adjustments may be considered, but their clinical utility must be established before they can be recommended. Meanwhile, it is advisable to keep examining $D_{L, CO}/VA$ and $VA$ separately [153], in so far as it may provide information on disease pathophysiology that cannot be obtained from their product, the $D_{L, CO}$.

Table 15 shows a summary on the considerations for $D_{L, CO}$ interpretation.

### ABBREVIATIONS

Table 16 contains a list of abbreviations and their meanings, which have been used in this series of Task Force reports.
ACKNOWLEDGEMENTS

R. Pellegrino: Azienda Ospedaliera S. Croce e Carle, Cuneo, Italy; G. Viegi: CNR Institute of Clinical Physiology, Pisa, Italy; V. Brusasco: Università degli Studi di Genova, Genova, Italy; R.O. Crapo and R. Jensen: LDS Hospital, Salt Lake City, UT, USA; F. Burgos: Hospital Clinic Villarreal, Barcelona, Spain; R. Casaburi: Harbor UCLA Medical Center, Torrance, CA, USA; A. Coates: Hospital for Sick Children, Toronto, ON, Canada; C.P.M. van der Grinten: University Hospital of Maastricht, Maastricht, the Netherlands; P. Gustafsson: Queen Silvias Children’s Hospital, Gothenburg, Sweden; J. Hankinson: Hankinson Consulting, Inc., Valdosta, GA, USA; D.C. Johnson: Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA; N. MacIntyre: University Hospital Birmingham NHS Trust, Birmingham, UK; D. Navajas: Lab Biofisica I Bioenginyeria, Barcelona, Spain; O.F. Pedersen: University of Aarhus, Aarhus, Denmark; J. Wangen: Pharmaceutical Research Associates, Inc., Lenexa, KS, USA.

REFERENCES


Lebowitz MD, Quackenboss J, Camilli AE, Bronnimann D, Holberg CJ, Boyer B. The epidemiological importance of...


Hughes JM, Bates DV. Historical review: the carbon monoxide diffusing capacity (DLCO) and its membrane (DM) and red cell (Theta.Vc) components. *Respir Physiol Neurobiol* 2003; 138: 115–142.


