

- pulmonary disease. *N Engl J Med* 2005;**352**:1967–76.
- 9 Ito K, Lim S, Caramori G, et al. A molecular mechanism of action of theophylline: induction of histone deacetylase activity to decrease inflammatory gene expression. *Proc Natl Acad Sci USA* 2002;**99**:8921–6.
- 10 Cosio BG, Tsaprouni L, Ito K, et al. Theophylline restores histone deacetylase activity and steroid responses in COPD macrophages. *J Exp Med* 2004;**200**:689–95.
- 11 Ito K, Yamamura S, Essilfie-Quaye S, et al. Histone deacetylase 2-mediated deacetylation of the glucocorticoid receptor enables NF- $\kappa$ B suppression. *J Exp Med* 2006;**203**:7–13.
- 12 Barnes PJ, Ito K, Adcock IM. A mechanism of corticosteroid resistance in COPD: inactivation of histone deacetylase. *Lancet* 2004;**363**:731–3.
- 13 Ito K, Tomita T, Barnes PJ, et al. Oxidative stress reduces histone deacetylase (HDAC)2 activity and enhances IL-8 gene expression: role of tyrosine nitration. *Biochem Biophys Res Commun* 2004;**315**:240–5.
- 14 Ichinose M, Sugiura H, Yamagata S, et al. Increase in reactive nitrogen species production in chronic obstructive pulmonary disease airways. *Am J Respir Crit Care Med* 2000;**160**:701–6.

## Standardisation of lung function testing

# Standardisation of lung function testing: helpful guidance from the ATS/ERS Task Force

G Laszlo

## A critical overview of the new ATS/ERS guidelines

The American Thoracic Society and the European Respiratory Society have jointly issued a new revision of their guidelines for the performance of spirometry, lung volumes, and carbon monoxide transfer factor. These have been published as a series of documents in the *European Respiratory Journal*.<sup>1–5</sup> They contain much wisdom, some compromises, and a few new recommendations. Blood gases, sleep, exercise, and challenge testing have not yet been readdressed. This brief review highlights a few of the more important recommendations dealing with the performance and interpretation of the several tests.

### GENERAL CONSIDERATIONS

This first chapter is essential reading for laboratory staff and sets standards for hygiene, calibration, quality control, and housekeeping. Observance of these standards will reassure research workers as well as clinicians.

### SPIROMETRY

Peak flow is the topic of current research and the task force plans to introduce more stringent standards for home recording. It may be derived from the flow-volume plot or from a separate blow, ideally using a flow measuring device. The guideline emphasises the importance of rehearsal and the need to blow immediately after a full inspiration.

Relaxed expired and inspired vital capacity (EVC and IVC) have been rehabilitated, in spite of the fact that the various COPD guidelines—such as GOLD and others—chose to dispense with them for simplicity. When performing spirometry, the suggested

method for EVC is to take the best of three measurements made before the forced expiratory tests, instructing the patient to speed the expiration only at the beginning and end of the blow. There is still no validated standard pattern for this test; one suggestion might be “take a full breath in; breathe out gently but firmly”, going on to further encouragement after 2–3 seconds until flow is less than 0.25 l/s.

For forced vital capacity (FVC), the Working Party has retained the old ATS recommendation to record 14 seconds of forced expiration, using the same criterion to identify the end of the test, and emphasising the need to inspect the curves to identify glottal closure and other sources of error. The 6 second blow (FEV<sub>6</sub>) is fully documented as a surrogate for the more demanding FVC manoeuvre,<sup>6</sup> but the Task Force stopped short of recommending its use, perhaps because of the lack of European standards. This topic has been aired again since the guidelines were published.<sup>7–9</sup> FEV<sub>1</sub>/FEV<sub>6</sub> identifies 94% of those diagnosed as having airflow obstruction by FEV<sub>1</sub>/FVC using normal values for the latter. The only obstacle to replacing FVC is the lack of well documented reference data from Europe; crudely, FEV<sub>1</sub>/FVC <0.7 identifies the same population as FEV<sub>1</sub>/FEV<sub>6</sub> <0.73.<sup>7</sup>

Vital capacity (VC) is defined in the document as the maximal volume that can be displaced from the lung—that is, the greatest among EVC, FVC and IVC. EVC can be greater or less than FVC in healthy subjects and those with restrictive ventilatory disorders, according to the method used; EVC usually exceeds

FVC in patients with COPD. For identifying airflow obstruction it would therefore make sense to abandon the 14 second FVC, measure FEV<sub>1</sub> as a fraction of FEV<sub>6</sub>, EVC and IVC, and report the lowest of these ratios as FEV<sub>1</sub>/VC.

Numerous examples of characteristic flow-volume loops are provided; it is implied that their appearance is much more informative than numerical indices derived from them. The latter are dismissed without any description or references because it is said that all the clinical information they contain can be derived from FEV<sub>1</sub>, VC, PEF, and FEF<sub>25–75</sub>. FEF<sub>25</sub> is the approved term to describe the instantaneous flow when 25% of the FVC has been exhaled. Standards are set for equipment and there are some suggestions about the sequence of blows.

### LUNG VOLUMES

This section reflects the wide range of practice. Body plethysmography is described first, then nitrogen washout (which persists because of the useful indices of gas mixing such as lung clearance index that can be calculated as well), and then helium dilution. No great change is recommended. Regarding plethysmography, it is recommended that the term thoracic gas volume (TGV) should be abandoned and replaced by functional residual capacity (FRC), which implies that the operator will have to train the subject not to breathe in while waiting for the shutter to close.

There was no consensus about whether to breathe in or out first from FRC when measuring the subdivisions of lung volume; one of each would be a sensible compromise. Breathing to residual volume (RV) first followed by IVC gives a good estimate of both RV and total lung capacity (TLC), which can be compared with single breath RV and alveolar volume measured with the same manoeuvre during the measurement of carbon monoxide transfer factor (TLCO).

Rebreathing of helium is regarded as complete when the helium concentration falls by less than 0.02% in 30 seconds, the time to achieve this being “rarely longer than 10 minutes”. Being sufficiently reproducible, one technically satisfactory

test is regarded as sufficient, in contrast to all other tests. This is just as well because it can take half an hour to wash the residual helium from emphysematous bullae.

### CARBON MONOXIDE UPTAKE

The discussion on the diffusing capacity (transfer factor) for carbon monoxide ( $DL_{CO}$ ,  $Tl_{CO}$ ) is now much clearer than previously and merits detailed examination. Complying with the standards may call for some effort, given the level of automation that already exists.

The Task Force recognised that the term “transfer factor” was popular in Europe because the uptake of carbon monoxide in the whole lung, as opposed to the individual alveoli, is not limited by diffusion. However, “diffusing capacity” has been chosen for the sake of “consistency” and for historical reasons. Presumably they chose the term which would meet with the least resistance across the world. Arguments about units are justifiably dismissed; either can be used as there is a simple conversion factor.

The single breath breath-holding test has been selected as the standard, using helium or methane as the insoluble gas to measure accessible lung volume (VA). When new methods of measurement make it desirable to use gases with lower diffusivity in the alveolar gas, they will have to be validated against helium and a correction factor applied if necessary. The intra-breath method, which is available commercially and has some advantages, is not mentioned but, by inference, it can be used if adequate reference values are made available.

The calculation of  $Tl_{CO}$  is standardised in minute detail and laboratories will have to ascertain whether their software produces correct results. Instrumental dead space and expired gas absorption have to be allowed for in the calculation of VA; also there are recommended formulae for calculating anatomical dead space from weight and from height. The effect of the partial pressure of oxygen ( $PO_2$ ) on carbon monoxide uptake is described. The oxygen dependent (red cell) and independent (membrane) components of  $Tl_{CO}$  are explained; while these measurements are not often made, the theory behind them explains why  $Tl_{CO}$  is increased by more than 10% if the alveolar oxygen concentration after the breath-hold is 16% rather than 21%. Rather than recommending measuring expired oxygen and applying a correction, it is suggested that laboratories employ the same inspired gas as was used to obtain their reference values. This is not an available option for those using composite equations, so  $PO_2$  remains a source of variation.

Strictly, in the single breath test the subject should hold a breath at TLC for about 10 seconds without any Valsalva or Müller manoeuvre. This is not always achievable, and measurement at lower lung volumes has an alinear effect on the result; the surface area for absorption is reduced but the pulmonary capillary volume per unit lung volume is increased. In normal subjects the calculated  $Tl_{CO}$  is 93% of the true value when VA is 85% TLC, but there are no validated disease specific corrections. It is suggested that inspired volume should be reported and that tests with  $V_i/VC$  of less than 85% should usually be rejected.

Dealing with the problem of anaemia, the guidelines state that  $Tl_{CO}$  should be reported as it is calculated, and a theoretical correction applied to the predicted value rather than the measured result. This is a satisfactory procedure when the results are considered as a percentage of the mean reference value, but it fits uneasily into the approach which employs residual standard deviations to identify lower limits of normal. Women are supposed to have lower haemoglobin than the male value of 14.6 g/dl, but this may not apply to young non-smokers who take regular physical exercise. Carboxyhaemoglobin (COHb) can be ignored if it is 2% or less; otherwise, the interpretation can be improved using an empirical calculation of a 1% reduction of  $Tl_{CO}$  for each 1% increase in COHb.

In general the recommendations are helpful and not too prescriptive or excessively rigorous; the interpreter is encouraged to estimate potential errors and thus minimise the chance of clinical misjudgements. Any assumptions, corrections, and allowances for imperfect technique should be listed in the report; the authors might have added that this should also apply in published papers.

$Tl_{CO}$  is now recognised as measuring the CO uptake in the ventilated part of the lung. It is calculated by multiplying the transfer coefficient ( $K_{CO}$ ) by the alveolar volume measured at the same time; the option of using a separately measured TLC has been withdrawn.

### STRATEGIES FOR INTERPRETATION

The section on interpretation<sup>5</sup> pulls the whole together and provides advice for designing reporting algorithms. Reports should deal consecutively with (1) quality; (2) reference values; (3) patterns of abnormality; (4) comparisons with self (change). A fifth important component is to answer the implied question on the request form, which currently requires some human input.

Lower limits of normal are again taken from residual standard deviations

so that less than 5% of healthy subjects are misclassified, mainly those with borderline results. In 1983 the ECSC standards<sup>10</sup> included composite reference values calculated from a number of sources which are widely used in Europe but which introduced a number of errors. Up to date reference equations are urgently needed for European populations of ethnic backgrounds (and, indeed, for all ages). This may take some time; meanwhile, the results of Medline searches for reference equations are tabled and laboratories are invited to choose the most appropriate for their population. Allowance for the patient's self-defined ethnic group is better done from reference equations than from factors (for example,  $\times 0.94$  applied to American standards for Asian-American lung volumes). Extrapolation is strongly discouraged for both age and stature.

The old ATS recommendation to confirm the reference equations used by studying 40 healthy subjects locally, which has served well, is discarded on the grounds that at least 100 spirometric tests would have to be carried out to show a significant difference. Perhaps this misunderstands what I took to be the point of the recommendation, which is that any serious deficiency in the reference equations would be detected by 40 readings from a local population.

There is a simple reporting algorithm employing lower limits of normal. The starting point is  $FEV_1/VC$ . A low ratio is interpreted as airflow obstruction. There is said to be no need to measure TLC if VC is normal; TLC is said to be required only to confirm the presence of restriction when VC is low. In fact, when  $FEV_1$  and VC are normal, omitting lung volume measurements overlooks “only” 4% of restrictive defects—that is, those with low  $RV$ ,<sup>9</sup> but these are usually younger patients with sarcoidosis or non-specific interstitial pneumonia. I suspect that investigators will continue to measure lung volumes when testing patients with puzzling breathlessness or with chest radiographs suggesting alveolar disease.

There is useful advice for clinical laboratories about the use of indices such as  $FEV_1$  (ml)/PEF (l/min) (when this is  $>8$  the loop should be examined for central airway obstruction). There are sensible recommendations about assessment of severity following various previous documents. This is still based on  $FEV_1$ , but laboratories are encouraged to look at other indicators—for example, hyperinflation and flow limitation in patients with airflow obstruction. Interpreters are advised to recommend other tests when necessary, such as those estimating respiratory muscle power. In other words, clinical

tests should be reported in the light of clinical information and not just by using limits of normality. Further, the guidelines recommend that the reports should be screened and the whole system checked if there are repeated instances where the report does not tally with the eventual diagnosis. Epidemiologists, who require simple repeatable procedures, would do well to study these pages as they illustrate the limitations of single tests.

### INCREASE IN FEV<sub>1</sub> AND FVC AFTER BRONCHODILATOR ADMINISTRATION

The main indication for measuring bronchodilator responsiveness is to diagnose untreated asthma in primary care. The guidelines suggest a normal bronchodilator responsiveness of FEV<sub>1</sub> and VC of <12% in healthy subjects on the basis of a number of studies; by implication, a greater variability at least points towards a diagnosis of asthma. Disappointingly, the Task Force did not achieve a consensus on the interpretation of bronchodilator responsiveness in COPD and were unable to define a level of improvement in FEV<sub>1</sub> which, when present, called for asthma treatment (+10% of predicted normal has been employed to separate asthmatic from non-asthmatic COPD<sup>11</sup>). An increase of >12% was thought to be greater than random, but it is known that this figure is not very useful. In patients with non-asthmatic COPD the increase in FEV<sub>1</sub> and FVC after bronchodilator is more or less normally distributed.<sup>11 12</sup> Moreover,

these increases vary from day to day and are not a predictor for clinical response to treatment with any agent or a marker for any type of disease, so the guidelines emphasise the need for therapeutic trials to improve airway patency for several days regardless of the results of a single laboratory trial. FEV<sub>1</sub> after bronchodilator is a more stable measurement than the pre-bronchodilator reading, so it is worth measuring in longitudinal studies of COPD.<sup>11</sup> Guidelines for a number of dose schedules are given;<sup>1</sup> the recommended standard is 400 µg salbutamol given as four separate inhalations.<sup>3</sup> This is satisfactory for COPD; 200 µg is enough for subjects with no previous exposure to beta-adrenergic agents.

Brecht wrote "That which is sure, is not sure. As things are, they shall not remain".<sup>13</sup> The ATS/ERS guidelines have been programmed to die in 10 years. While thanking the Task Force for these thoughtful and, in the main, helpful recommendations, I would encourage them to enter another date in their diaries for somewhat sooner than 2015.

### ACKNOWLEDGEMENTS

I thank Dr Adrian Kendrick, Dr Mike Morgan, and Dr Evelyn Smith for their helpful comments on the first draft of this paper.

*Thorax* 2006;**61**:744–746.  
doi: 10.1136/thx.2006.061648

Correspondence to: Dr Gabriel Laszlo, Department of Respiratory Medicine, Bristol Royal Infirmary, Bristol BS2 8HW, UK; glaszlo11@aol.com

Competing interests: none.

### REFERENCES

- 1 Miller MR, Crapo R, Hankinson J, *et al*. General considerations for lung function testing. *Eur Respir J* 2005;**26**:153–61.
- 2 Miller MR, Hankinson J, Brusasco V, *et al*. Standardisation of spirometry. *Eur Respir J* 2005;**26**:319–38.
- 3 Wanger J, Clausen JL, Coates A, *et al*. Standardisation of the measurement of lung volumes. *Eur Respir J* 2005;**26**:511–22.
- 4 MacIntyre N, Crapo RO, Viegi G, *et al*. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. *Eur Respir J* 2005;**26**:720–35.
- 5 Pellegrino R, Viegi G, Brusasco V, *et al*. Interpretive strategies for lung function tests. *Eur Respir J* 2005;**26**:948–68.
- 6 Swanney MP, Jensen RL, Crichton DA, *et al*. FEV<sub>6</sub> is an acceptable surrogate for FVC in the spirometric diagnosis of airway obstruction and restriction. *Am J Respir Crit Care Med* 2000;**162**:917–9.
- 7 Pedersen OF. FEV<sub>6</sub>: a shortcut in spirometry? *Eur Respir J* 2006;**27**:245–7.
- 8 Akpinar-Elci M, Fedan KB, Enright PL. FEV<sub>6</sub> as a surrogate for FVC in detecting airways obstruction and restriction in workplace. *Eur Respir J* 2006;**27**:374–7.
- 9 Vandevoorde J, Verbanck S, Schuermans D, *et al*. Obstructive and restrictive spirometric patterns: fixed cut-offs for FEV<sub>1</sub>/FEV<sub>6</sub> and FEV<sub>6</sub>. *Eur Respir J* 2006;**27**:378–83.
- 10 Quanjer PH, ed. Standardized lung function testing. Report Working Party Standardization of Lung Function Tests. European Community for Coal and Steel. *Bull Eur Physiopathol Respir* 1983;**19**(Suppl 5):1–95.
- 11 Calverley PMA, Burge PS, Spencer S, for the ISOLDE Study Investigators, *et al*. Bronchodilator reversibility testing in chronic obstructive pulmonary disease. *Thorax* 2003;**58**:659–64.
- 12 Anthonisen NR, Lindgren PG, Tashkin DP *et al* for the Lung Health Study Research Group. Bronchodilator response in the lung health study over 11 years. *Eur Respir J* 2005;**26**:46–51.
- 13 Brecht B. *In Praise of Dialectics*, 1933. Available at <http://www.chtodelat.org/index.php?option=content&task=view&id=226&Itemid=98> (accessed 20 February 2006). Translation by Beaton GH. Available at <http://content.nejm.org/cgi/content/full/354/3/312> (accessed 20 February 2006).