

Standardisation of lung function testing: the authors' replies to readers' comments

To the Editors:

A few questions have been raised following the publication in 2005 of the joint American Thoracic Society (ATS)/European Respiratory Society (ERS) series of documents on standardising lung function testing and these are answered below.

SPIROMETRY

The following questions and answers pertain to the standardisation document for spirometry [1].

Start of test criteria

Should blows be rejected solely on the basis of a poor back extrapolated volume (EV)?

Reply

Usually. The forced vital capacity (FVC) may be usable, but the forced expiratory volume in 1 s (FEV₁) is likely to be falsely high or low.

Rationale

The acceptability criteria for spirometry were designed to help technologists improve the subject's technique in order to get the best and most reliable result. EV is important for determining that a fast start to the blow was achieved and this is crucial for getting the best values for FEV₁ and peak expiratory flow (PEF).

End of test criteria

In the original document, there was an error in table 5.

Reply

In table 5, the within-manoeuvre criteria for a satisfactory completion of a blow should have read "Duration of ≥ 6 s (3 s for children) and a plateau in the volume-time curve, or if the subject cannot or should not continue to exhale." The original table had "or" in twice, whereas the accompanying text was correct, as above.

Rationale

The end of test (EOT) criteria are applied in order to ensure that efforts are made to achieve the best estimate of FVC. When a subject cannot meet the plateau criterion (<25 mL exhaled in the previous second of the blow) this may be for reasons other than premature volitional cessation of the blow. For example, in some young subjects or patients with a rigid chest wall, it is chest wall limitation that suddenly causes exhalation to stop [2] and it is difficult for them to achieve a volume-time plateau of >1 s. Under these circumstances, the FVC values returned will be repeatable, whereas if the cessation of flow is volitional, then the repeatability is usually poor. The length of time necessary to achieve a plateau is dependent on many factors, including, but not limited to, the subject's age, body size and presence of lung disease. In general, children and adolescents

reach their plateau more quickly than adults, as do adults with small body frames (e.g. Asian females). Patients with restrictive disease can also reach their EOT more quickly than patients with airflow obstruction. The 3 s expiratory time attributed to children was given as a guide to the fact that longer expiratory times, ≥ 6 s, can be difficult to achieve in younger subjects.

If a subject fails to meet the EOT criteria, their results must not be discarded, since the result obtained may still give valuable clinical information. The results obtained should be interpreted with the caveat that the EOT criteria were not met, so the FVC result might be an underestimate and, thereby, falsely increase the FEV₁/FVC ratio. Variation in FVC can also arise from a failure to inhale fully to total lung capacity (TLC) at the start of the manoeuvre, or instrument errors. Failure to meet EOT criteria is a prompt to the technician to ensure the subject tries harder on subsequent blows to continue exhaling at the end to achieve the best FVC result. Under no circumstances should a lower FVC value be taken from a smaller blow that did meet EOT criteria and used in place of a larger FVC value from an otherwise acceptable blow that did not meet EOT criteria. Similarly, the FEV₁ can be taken from curves that did not meet EOT criteria but were otherwise acceptable blows.

Calculating the FEV₁/FVC ratio

It was unclear whether the FEV₁/FVC ratio is calculated from the values from the "best" blow (greatest sum of FEV₁ and FVC) or from the individual best values.

Reply

The largest obtained values of FEV₁ and FVC are used in the ratio even if they are obtained from different blows.

Rationale

The FEV₁/FVC ratio is critical for determining if a subject has airflow obstruction. The values of PEF and FEV₁ are the largest values recorded from manoeuvres without a high EV. The FVC is the largest FVC recorded from manoeuvres without extra breaths, coughs, pauses or a zero-flow error. In the original document [1], figure 3 indicated that the single blow with the largest sum of FVC and FEV₁ was used for deriving "other indices". Other indices were not explicitly stated, but might include time domain or moment analysis, forced expiratory flow at 25–75% of FVC and instantaneous flows other than PEF; these are lung function indices that are currently not recommended or supported as suitable for making clinical decisions on patients.

The interpretation document [3] referred to using the FEV₁/vital capacity (VC) for assessing airflow obstruction. VC was intended to refer generically to the highest acceptable VC recorded, whether it be from FVC manoeuvres or slow VC manoeuvres (inspiratory or expiratory). If more than one of these indices are available, the largest should be used to derive

the FEV₁/(F)VC. We understand that the FVC is most commonly available. In healthy subjects, there is little difference between the FVC, expiratory VC and inspiratory VC, so the reference equations for FEV₁/FVC can be used as an approximation [3, 4]. However, caution should be used when using VC measurements that come from tests with potentially different calibrations or measurement transducers.

Test signals for PEF meter testing

The requirements concerning repeatability errors appeared inconsistent.

Reply

There was an error in the section on repeatability testing, which has been revised as below. "Flow waveforms 1, 4, 8 and 25 are discharged three times to each of 10 production meters. The span of readings for each meter with each waveform is ascertained (40 span results). The repeatability validation limits are $\pm 6\%$ or $\pm 15 \text{ L}\cdot\text{min}^{-1}$, whichever is the greater, and these limits include 1% for waveform-generator variability. A repeatability error occurs if the span exceeds these limits. Acceptable performance is defined as two or fewer errors in the 40 test results (*i.e.* maximum error rate of 5%).".

Rationale

The original text left ambiguity as to how the span of results was derived and judged. This has now been clarified.

Units for resistance

The units for resistance were in error and all should read as either: $\text{cmH}_2\text{O}\cdot\text{L}^{-1}\cdot\text{s}$ or $\text{kPa}\cdot\text{L}^{-1}\cdot\text{s}$.

SINGLE-BREATH DETERMINATION OF CARBON MONOXIDE UPTAKE IN THE LUNG

Typographical error

There was a small typographical error in table 5 [4].

Reply

A small correction is necessary in table 5 in relation to the formula for when H₂O and CO₂ are removed from sampled gas. The correct formulae are:

$$V_{A,BTPS} = (V_{I,ATPD} - V_{D,INST} - V_{D,ANAT}) \times (F_{I,Tr} / ((1 - F_{A,CO_2}) \times F_{S,Tr})) \times (P_B / (P_B - 47)) \times (310 / (273 + T))$$

$$V_{A,STPD} = (V_{I,ATPD} - V_{D,INST} - V_{D,ANAT}) \times (F_{I,Tr} / ((1 - F_{A,CO_2}) \times F_{S,Tr})) \times (P_B / 760) \times (273 / (273 + T))$$

INTERPRETATIVE STRATEGIES

Lower limits of normality

In the original document (p. 949) there was some confusion between percentiles and confidence intervals [3].

Reply

The text should be: "If the reference data have a normal distribution (such as the NHANES III spirometry data), the 5th percentile can be estimated as the lower 90% confidence limit using two-tailed Gaussian statistics ($\text{mean} - 1.645 \times \text{SD}$). Fixed values, such as 80% predicted for FEV₁ or FVC, or 0.70 for FEV₁/FVC, should never be used to determine the lower limit of the normal range (LLN). If the distribution is skewed, the

LLN should be estimated with a nonparametric technique to derive the 5th percentile value."

Rationale

The 90% confidence limits of a Gaussian distribution are defined by the mean value $\pm 1.645 \times \text{SD}$ and there will be 5% of the normal distribution with values below the mean $- 1.645 \times \text{SD}$ and 5% above the mean $+ 1.645 \times \text{SD}$. Regression equations quote the standard error of the estimate (SEE) or residual standard deviation (RSD) for the prediction, which are exactly the same thing, and are the relevant SD to use in estimating the prediction confidence limits or limits of normality. For the spirometric indices FEV₁, forced expiratory volume in 6 s, FVC and PEF, any supranormal values found in the population are not usually deemed to be "abnormal", in the sense that they are not thought to be excessively high due to disease. Thus, the lower 90% confidence limit is an estimate of the 5th percentile for the population and is taken as the LLN. This can also be considered the 95% confidence interval using one-tailed Gaussian statistics.

All of the indices derived from measuring static lung volumes and the index of single-breath carbon monoxide diffusing capacity can have values that are abnormally high or abnormally low, and so the upper limit of normal needs to be defined, as well as the LLN. In clinical laboratory lung function testing, there is a low *a priori* probability that the client population is going to be "normal", and so many laboratories use predicted value $\pm 1.645 \times \text{SD}$, which would encompass 90% of a normal population, for all indices. Thus, if a large number of subjects without any disease were to be tested against these "normal" limits, the 5% of normal subjects with values $> 1.645 \text{ SD}$ above the predicted value and the 5% of normal subjects with values $> 1.645 \text{ SD}$ below the predicted value will be classified as "abnormal". This means the testing is more sensitive in detecting patients with possible disease, but there will be a reduction in specificity because of the above "false positives". If an investigator decides to limit false positives to just 5% of any normal individuals tested (not 10% as above), for example, if one was screening a random population for evidence of a disease, then indices with possible abnormalities in either the upper or lower range would require the normal range to be defined as the predicted value $\pm 1.96 \times \text{SD}$ (which defines the two sided 95% confidence limits).

Severity of airflow obstruction

Is the severity of airflow obstruction assessed from the level of the FEV₁/(F)VC ratio or the level of FEV₁?

Reply

The FEV₁/(F)VC is used to determine whether airflow obstruction is present, by judging it against its LLN. The FEV₁ is then used to assess the severity.

Rationale

The 2005 interpretation statement [3] continued the earlier 1991 recommendation [5]. Using the FEV₁ to estimate the severity of obstruction is appropriate when only obstruction is present, but can be misleading if there is coexisting restriction, since the FEV₁ may be reduced by restriction as well as by obstruction. Static lung volume measurements are necessary to distinguish

restrictive from mixed-restrictive and obstructive abnormalities. When only spirometry is available, it is not possible to determine whether restriction is present, because the (F)VC may be reduced by restriction due to a reduced TLC or by elevation of the residual volume due to air trapping.

M.R. Miller^{*}, J. Hankinson[#], V. Brusasco[¶], F. Burgos[†], R. Casaburi[§], A. Coates[‡], P. Enright^{**}, C. van der Grinten^{##}, P. Gustafsson^{¶¶}, R. Jensen⁺⁺, N. MacIntyre^{§§}, R.T. McKay^{‡‡}, O.F. Pedersen^{***}, R. Pellegrino^{###}, G. Viegi^{¶¶¶} and J. Wanger⁺⁺⁺

^{*}Dept of Medicine, University Hospitals Birmingham, Birmingham, UK. [#]Hankinson Consulting Inc., Athens, GA, [§]Los Angeles Biomedical Research Institute, Harbor-UCLA Medical Center, Torrance, CA, ^{**}College of Public Health, The University of Arizona, Tucson, AZ, ⁺⁺LDS Hospital, Pulmonary Division, Salt Lake City, UT, ^{§§}Dept of Medicine, Duke University Medical Center, Durham, NC, ^{‡‡}University of Cincinnati, Cincinnati, OH, ⁺⁺⁺PRA International, Lenexa, KS, USA. [¶]Dept of Internal Medicine, University of Genoa, Genoa, ^{###}Allergologia e Fisiopatologia Respiratoria, ASO S. Croce e Carle, Cuneo, ^{¶¶¶}CNR Institute of Clinical Physiology, Pisa, Italy. [†]Respiratory Dept, Hospital Clínic (IDIBAPS), CIBERes, Barcelona, Spain. [‡]Dept of Paediatrics, Hospital for Sick Children, University of Toronto, Toronto, Canada. ^{##}Maastricht University Medical Center, Maastricht, The Netherlands. ^{¶¶}Dept of Pediatrics, Central

Hospital, Skövde, Sweden. ^{***}Institute of Public Health, Dept of Environmental and Occupational Medicine, University of Aarhus, Aarhus, Denmark.

Correspondence: M.R. Miller, Dept of Medicine, 5th Floor Nuffield House, Queen Elizabeth Hospital, Birmingham, B15 2TH, United Kingdom. E-mail: martin.miller@uhb.nhs.uk

Statement of Interest: A statement of interest for J. Hankinson can be found at www.erj.ersjournals.com/site/misc/statements.xhtml

REFERENCES

- 1 Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.
- 2 Davis C, Campbell EJ, Openshaw P, *et al.* Importance of airway closure in limiting maximal expiration in normal man. *J Appl Physiol* 1980; 48: 695–701.
- 3 Pellegrino R, Viegi G, Brusasco V, *et al.* Interpretative strategies for lung function tests. *Eur Respir J* 2005; 26: 948–968.
- 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: 720–735.
- 5 American Thoracic Society. Lung function testing: selection of reference values and interpretative Strategies. *Am Rev Respir Dis* 1991; 144: 1202–1218.

DOI: 10.1183/09031936.00130010