Pulmonary function interpretative strategies: from statistics to clinical practice

Vito Brusasco[1] and Riccardo Pellegrino[2]

[1]Dipartimento di Medicina Sperimentale, Università di Genova, Genova, Italy. [2]Centro Medico Pneumologico Torino, Torino, Italy.

Corresponding author: Vito Brusasco (vito.brusasco@gmail.com)

Shareable abstract (@ERSpublications)
The application of suitable reference equations based on a robust statistical approach is central to pulmonary function testing, but the complexity of lung physiology makes the interpretation of results in individual patients challenging https://bit.ly/3gUpXz5

Cite this article as: Brusasco V, Pellegrino R. Pulmonary function interpretative strategies: from statistics to clinical practice. *Eur Respir J* 2022; 60: 2200317 [DOI: 10.1183/13993003.00317-2022].

Received: 10 Feb 2022
Accepted: 13 Feb 2022

In 2005, the *European Respiratory Journal* published a series of American Thoracic Society (ATS)/European Respiratory Society (ERS) consensus documents on the standardisation of pulmonary function testing (PFT), which ended with an article on interpretative strategies [1]. Despite being well received by the scientific community (with over 3500 citations so far), that article generated controversy in a couple of areas and left some questions open due to insufficient evidence available at that time. After more than 10 years, the two societies felt the need to update the whole series, with the main purpose of addressing recent technical advancements. The final document published in the current issue of the *European Respiratory Journal* [2] is again on interpretative strategies. The difference between the two interpretative documents is that the current one is focused on technical sources of uncertainty, rather than on algorithms for lung function planning and interpretation in the clinical context.

The choice of reference equations is probably the major source of uncertainty in PFT, a problem that was left open by the 2005 ATS/ERS committee that realised that no single set of reference values could be recommended at that time and more work was necessary in this area. One of the greatest and most important advancements of this new document is definitely the recommendation in favour of the reference equations provided by the Global Lung Function Initiative (GLI). These were generated using large datasets by a robust statistical approach taking into account mean, variability and skewness of distributions over extended ranges of age [3–5]. This avoids uncertainties due to extrapolation beyond age for oldest subjects, the discontinuity in the transition from adolescence to adult age and, for spirometry, ethnic differences. Moreover, the GLI equations provide the 5th and 95th percentiles of distributions, which can be used, as for many biological variables, to define lower and upper limits of normality. These criteria had been already recommended by the committee of the 2005 document [1] but, unfortunately, most clinical documents and guidelines in the respiratory field have maintained definitions of lung function abnormalities based on age- and sex-biased thresholds. An important step forward made by the committee of the present document is the introduction of the z-score, which is a measure of how far a measured value is from the predicted one, expressed as standard deviations. Since z-scores can be converted to percentiles of frequency distribution they provide an estimate of uncertainty in ruling disease in or out, particularly in the grey area of transition between health and disease. Moreover, z-score has also been proposed for grading the severity of abnormalities to replace the size-biased percentage of predicted [6]. Ultimately, the advent of the z-score in respiratory medicine should help remove inconsistencies due to differences in the definitions of the same type of lung function abnormality by different clinical committees [7, 8].

Since most PFT is undertaken to address a clinical problem, a question is whether such a statistical approach, however robust, will help dismiss all uncertainties we may have when interpreting the results for an individual subject. Recent studies have demonstrated that GLI-defined spirometric abnormalities and z-score grading identify normal ageing phenotypes [9] and meaningful phenotypes of either obstructive or

restrictive disorders [10]. A general principle of decision making is that the post-test probability of disease should be estimated considering the pre-test probability; z-score would allow this, avoiding pre-determined thresholds. Unfortunately, information for estimating pre-test probability is often not provided by the referring clinician and this should be encouraged by disease-related documents and guidelines.

The current ERS/ATS interpretative document [2] has certainly filled some gaps in the 2005 document [1] by providing separate and more detailed interpretative flow charts for spirometry, lung volumes and diffusing capacity. However, a step-by-step approach exploiting all available tests is often necessary to interpret the PFT of individual patients, owing to the complexity of physiological factors, with possible contrasting effects on different tests. Generally, forced expiratory volume in 1 s (FEV_1) and the ratio of FEV_1 to forced vital capacity (FVC) are taken as the basic markers of lung impairment, not only for their associations with respiratory diseases and the ease with which they can be measured, but also because they reflect major determinants of respiratory mechanics. However, it is well known that different pathological conditions may importantly interfere with their measurement and interpretation. Here are some examples. As early as 1965, Woolcock and Read [11] reported clinical improvement from severe asthma attack not reflected by FEV_1, which they attributed to return of total lung capacity and lung elastic recoil to normality. In obesity, static lung volumes tend to decrease whereas FEV_1/FVC tends to increase, which may obscure the presence of airflow obstruction in patients with combined obesity and COPD [12]. In combined pulmonary fibrosis and emphysema, the overlapping restrictive and obstructive abnormalities may have opposing effects, by which spirometry and lung volumes may even remain within the range of normality [13], while they have additive effects on diffusing capacity. In some purely restrictive disorders, FVC and lung volumes may be occasionally found within normal ranges despite the presence of significant interstitial fibrosis [14]. This may be explained by pre-morbid values within the upper limits of normality and diffusing capacity would be required to reveal abnormal lung function.

There are also technical aspects related to manoeuvres that may influence the interpretation of spirometry in different conditions. First, negative effort-dependence of maximal expiratory flow was documented several years ago [15]. This phenomenon, which is due to intrathoracic gas compression, will cause FEV_1 and FEV_1/FVC to be higher in submaximal than maximal efforts, thus possibly masking spirometric abnormalities. In severe emphysema this effect may account for up to 50% of the FEV_1 reduction [16], thus overrating the severity classification in comparison with COPD patients with prevailing airway obstructive disease. Moreover, in subjects with large lung volume, gas compression may exaggerate both bronchoconstrictor and bronchodilator responses to pharmacological agents [17], thus potentially causing misclassification of asthma or COPD in tall subjects. Second, several studies over the past 25 years have documented that the deep inspiration preceding the forced expiration may affect spirometric measurements differently depending on type of disease, *e.g.* COPD *versus* asthma, and phase of disease [18]. Finally, in very few asthma patients, serial spirometric manoeuvres show progressive bronchoconstriction [19]. This makes the repeatability criteria inapplicable under this condition but gives a clinically useful information. Altogether, the above examples show how the complexity of lung physiology adds uncertainty to the interpretation of single or few indices based uniquely on statistical principles, rigorous as they may be.

Another aspect of this document that will certainly stir up discussion within our scientific community is the role of bronchodilator testing in clinical practice. Historically, the test was born in the 1970s with the hope of differentiating bronchial asthma from COPD. Yet since the mid-1980s, important trials have let the purpose of the test remain unfulfilled because of a large overlap of positive responses between groups [20, 21]. In a recent study including 35 628 subjects, the authors concluded that "bronchodilator reversibility was at least as common in participants with COPD as those with asthma. This indicates that measures of reversibility are of limited value for distinguishing asthma from COPD in population studies" [22]. Yet to date, the test is being used in clinical practice to support the diagnosis of COPD [7] or bronchial asthma [8]. In the present interpretative document, it is recognised that the results of the test "are limited for clinically meaningful thresholds across a range of diseases and age groups", but "…there is evidence related to survival to support a threshold-based approach". Moreover, it is known that acute bronchodilator responses are not predictive of responses to long-term treatment. Many colleagues would, therefore, wonder why there is still that much emphasis on a test unable to address any clinical question and whether prediction of survival could be a good reason for the patients to undergo this test. The conundrums with the clinical use of bronchodilator tests are mainly the heterogeneous mechanisms of airflow obstruction, *e.g.* airway smooth muscle contraction *versus* reduced mechanical interdependence between airways and lung parenchyma in different types and stages of the underlying disease, the poor reproducibility, and the choice of thresholds defining positive response.

In the current document [2], an increase >10% of predicted for FEV_1 and/or FVC after inhaling a bronchodilator agent is recommended, based on population studies. As co-authors of the previous

interpretative document [1], we acknowledge that the criteria proposed at that time (>200 mL and >12% from baseline) were imperfect, but this seems to be the case also for the new proposal. For example, in a tall patient with predicted FEV_1 of 3.00 L and baseline FEV_1 of 1.20 L, a 0.25-L post-bronchodilator increase would be said to be insignificant, even if its magnitude is like those found in the majority of clinical trials with bronchodilators in COPD, being associated with beneficial effects on patient-related outcomes.

In conclusion, we praise the committee members for having filled some technical gaps in the interpretation of PFT proposing robust predictive values and a rigorous statistical approach. Yet, what remains to be done is to provide recommendations for choosing the appropriate tests to answer clinical questions in individual patients considering the complexity of mechanisms underlying pulmonary function abnormalities and make PFT interpretation more uniform among different disease-specific guidelines.

Conflict of interest: The authors declare no conflict of interest.

References

1 Pellegrino R, Viegi G, Brusasco V, *et al.* Interpretative strategies for lung function tests. *Eur Respir J* 2005; 26: 948–968.
2 Stanojevic S, Kaminsky DA, Miller MR, *et al.* ERS/ATS technical standard on interpretive strategies for routine lung function tests. *Eur Respir J* 2022; 60: 2101499.
3 Quanjer PH, Stanojevic S, Cole TJ, *et al.* Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40: 1324–1343.
4 Stanojevic S, Graham BL, Cooper BG, *et al.* Official ERS technical standards: Global Lung Function Initiative reference values for the carbon monoxide transfer factor for Caucasians. *Eur Respir J* 2017; 50: 1700010.
5 Hall GL, Filipow N, Ruppel G, *et al.* Official ERS technical standard: Global Lung Function Initiative reference values for static lung volumes in individuals of European ancestry. *Eur Respir J* 2020; 57: 2000289.
6 Quanjer PH, Pretto JJ, Brazzale DJ, *et al.* Grading the severity of airways obstruction: new wine in new bottles. *Eur Respir J* 2014; 43: 505–512.
7 Global Initiative for Obstructive Lung Disease. The Global Strategy for Diagnosis, Management and Prevention of COPD (updated 2022). www.goldcopd.org
8 Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention, 2021. www.ginasthma.org
9 Vaz Fragoso CA, McAvay G, Van Ness PH, *et al.* Phenotype of normal spirometry in an aging population. *Am J Respir Crit Care Med* 2015; 192: 817–825.
10 Vaz Fragoso CA, McAvay G, Van Ness PH, *et al.* Phenotype of spirometric impairment in an aging population. *Am J Respir Crit Care Med* 2016; 193: 727–735.
11 Woolcock A, Read J. Improvement in bronchial asthma not reflected in forced expiratory volume. *Lancet* 1965; 2: 1323–1325.
12 O'Donnell DE, Deesomchok A, Lam YM, *et al.* Effects of BMI on static lung volumes in patients with airway obstruction. *Chest* 2011; 140: 461–468.
13 Barisione G, Brusasco C, Garlaschi A, *et al.* Lung function testing in COPD: when everything is not so simple. *Respirol Case Rep* 2014; 2: 141–143.
14 Barisione G, Garlaschi A, Occhipinti M, *et al.* Value of lung diffusing capacity for nitric oxide in systemic sclerosis. *Physiol Rep* 2019; 7: e14149.
15 Krowka MJ, Enright PL, Rodarte JR, *et al.* Effect of effort on measurement of forced expiratory volume in one second. *Am Rev Respir Dis* 1987; 136: 829–833.
16 Pellegrino R, Crimi E, Torchio R, *et al.* Severity grading of chronic obstructive pulmonary disease: the confounding effect of phenotype and thoracic gas compression. *J Appl Physiol (1985)* 2015; 118: 796–802.
17 Pellegrino R, Antonelli A, Crimi E, *et al.* Dependence of bronchoconstrictor and bronchodilator responses on thoracic gas compression volume. *Respirology* 2014; 19: 1040–1045.
18 Pellegrino R, Sterk P, Sont JK, *et al.* Assessing the effect of deep inhalation on airway calibre: a novel approach to lung function in bronchial asthma and COPD. *Eur Respir J* 1998; 12: 1219–1227.
19 Marthan R, Woolcock AJ. Is a myogenic response involved in deep inspiration-induced bronchoconstriction in asthmatics? *Am Rev Respir Dis* 1989; 140: 1354–1358.
20 Eliasson O, Degraff AC Jr. The use of criteria for reversibility and obstruction to define patient groups for bronchodilator trials. Influence of clinical diagnosis, spirometric, and anthropometric variables. *Am Rev Respir Dis* 1985; 132: 858–864.
21 Anthonisen NR, Wright EC, the IPPB TRIAL GROUP. Bronchodilator response in chronic obstructive pulmonary disease. *Am Rev Respir Dis* 1986; 133: 814–819.
22 Janson C, Malinovschi A, Amaral AFS, *et al.* Bronchodilator reversibility in asthma and COPD: findings from three large population studies. *Eur Respir J* 2019; 54: 1900561.