



ERS/ATS technical standard on interpretive strategies for routine lung function tests

Sanja Stanojevic ¹, David A. Kaminsky², Martin R. Miller ³, Bruce Thompson⁴, Andrea Aliverti⁵, Igor Barjaktarevic⁶, Brendan G. Cooper⁷, Bruce Culver⁸, Eric Derom⁹, Graham L. Hall¹⁰, Teal S. Hallstrand⁸, Joerg D. Leuppi^{11,12}, Neil MacIntyre¹³, Meredith McCormack¹⁴, Margaret Rosenfeld¹⁵ and Erik R. Swenson^{8,16}

¹Dept of Community Health and Epidemiology, Dalhousie University, Halifax, NS, Canada. ²Pulmonary Disease and Critical Care Medicine, University of Vermont Larner College of Medicine, Burlington, VT, USA. ³Institute of Applied Health Research, University of Birmingham, Birmingham, UK. ⁴Physiology Service, Dept of Respiratory Medicine, The Alfred Hospital and School of Health Sciences, Swinburne University of Technology, Melbourne, Australia. ⁵Dept of Electronics, Information and Bioengineering (DEIB), Politecnico di Milano, Milan, Italy. ⁶Division of Pulmonary and Critical Care Medicine, University of California, Los Angeles, CA, USA. ⁷Lung Function and Sleep, Queen Elizabeth Hospital, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK. ⁸Dept of Medicine, Division of Pulmonary, Critical Care and Sleep Medicine, University of Washington, Seattle, WA, USA. ⁹Dept of Respiratory Medicine, Ghent University, Ghent, Belgium. ¹⁰Children's Lung Health, Wal-yan Respiratory Research Centre, Telethon Kids Institute and School of Allied Health, Faculty of Health Science, Curtin University, Bentley, Australia. ¹¹University Clinic of Medicine, Cantonal Hospital Basel, Liestal, Switzerland. ¹²University Clinic of Medicine, University of Basel, Basel, Switzerland. ¹³Division of Pulmonary, Allergy, and Critical Care Medicine, Dept of Medicine, Duke University Medical Center, Durham, NC, USA. ¹⁴Pulmonary Function Laboratory, Pulmonary and Critical Care Medicine, Johns Hopkins University, Baltimore, MD, USA. ¹⁵Seattle Children's Hospital, Seattle, WA, USA. ¹⁶VA Puget Sound Health Care System, Seattle, WA, USA.

Corresponding author: Sanja Stanojevic (sanja.stanojevic@dal.ca)



Shareable abstract (@ERSpublications)

Data from pulmonary function tests must be complemented with clinical expertise and consideration of the inherent biological variability and uncertainty of the test result to ensure appropriate interpretation of an individual's lung function measurements <https://bit.ly/3ecluFc>

Cite this article as: 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 [DOI: 10.1183/13993003.01499-2021].

The content of this work is not subject to copyright. Design and branding are copyright ©ERS 2022. For reproduction rights and permissions contact permissions@ersnet.org

This article has an editorial commentary: <https://doi.org/10.1183/13993003.00317-2022>

Received: 26 May 2021
Accepted: 18 Nov 2021

Abstract

Background Appropriate interpretation of pulmonary function tests (PFTs) involves the classification of observed values as within/outside the normal range based on a reference population of healthy individuals, integrating knowledge of physiological determinants of test results into functional classifications and integrating patterns with other clinical data to estimate prognosis. In 2005, the American Thoracic Society (ATS) and European Respiratory Society (ERS) jointly adopted technical standards for the interpretation of PFTs. We aimed to update the 2005 recommendations and incorporate evidence from recent literature to establish new standards for PFT interpretation.

Methods This technical standards document was developed by an international joint Task Force, appointed by the ERS/ATS with multidisciplinary expertise in conducting and interpreting PFTs and developing international standards. A comprehensive literature review was conducted and published evidence was reviewed.

Results Recommendations for the choice of reference equations and limits of normal of the healthy population to identify individuals with unusually low or high results are discussed. Interpretation strategies for bronchodilator responsiveness testing, limits of natural changes over time and severity are also updated. Interpretation of measurements made by spirometry, lung volumes and gas transfer are described as they relate to underlying pathophysiology with updated classification protocols of common impairments.

Conclusions Interpretation of PFTs must be complemented with clinical expertise and consideration of the inherent biological variability of the test and the uncertainty of the test result to ensure appropriate interpretation of an individual's lung function measurements.

Introduction

Pulmonary function tests (PFTs)/respiratory function tests reflect the physiological properties of the lungs (*e.g.* airflow mechanics, volumes and gas transfer). These tests have been used for decades to help

diagnose lung disease, explain dyspnoea, and monitor disease progression and treatment response. In addition, PFTs have been employed in population studies of the association between exposures and lung health. The American Thoracic Society (ATS)/European Respiratory Society (ERS) Task Force on the standardisation of PFTs published a series of technical documents in 2005 [1–4]. The technical standards for spirometry [5] and single-breath carbon monoxide uptake in the lung (transfer factor (T_{LCO}) or diffusing capacity (D_{LCO})) [6] have recently been updated, and an update on lung volumes is forthcoming. This document is an update to the interpretation strategies of routine PFTs [3].

Interpretation of technically acceptable PFT results has three key aspects. 1) Classification of observed values as within/outside the normal range with respect to a population of healthy individuals. This involves consideration of the measurement error of the test, as well as the inherent biological variability of measurements both between individuals and between repeated measurements in the same individual. 2) Integration of knowledge of physiological determinants of test results into a functional classification of the identified impairments. 3) Integration of the identified patterns with other clinical data to inform differential diagnosis and guide therapy. These are three distinct, yet complementary aspects of interpretation. This document addresses only the first two aspects. The final integration of pulmonary function results into a diagnosis or management plan is beyond the scope of this technical guidance on physiological interpretation.

Appropriate interpretation of PFTs requires measurements that meet technical specifications for test performance and appropriate levels of quality [6–8]. Poorer quality tests must be interpreted with greater uncertainty as the measurements may not reflect functional impairments. Interpretation also relies on clear reporting of results; therefore, current ATS standards for reporting of PFTs are recommended [9]. Technical aspects of PFT measurement, equipment and biological controls are summarised in the ERS/ATS standards for each PFT [6–8].

This document considers the 2005 recommendations [1–4] and incorporates evidence from subsequent literature to establish new standards for PFT interpretation. The key distinction between the previous recommendations and the current ones is the emphasis on the uncertainty of measurement and interpretation.

A summary of the changes from the 2005 interpretation standard can be found in table 1.

Methods

Task Force members were selected by the ATS Proficiency Standards for Pulmonary Function Laboratories Committee, as well as ERS leadership. Conflicts of interest, including academic conflicts, were declared and vetted by the ATS throughout the duration of the Task Force. Six of the 16 Task Force members are current or past members of the Global Lung Function Initiative Network Executive. A comprehensive literature search was conducted by a professional librarian using the following databases: Ovid MEDLINE, Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily, Ovid MEDLINE 1946 to Present, Embase Classic, Embase 1947 to 29 March 2019 and Wiley Cochrane. The search terms are listed in figure 1. All identified publications were screened by two members of the Task Force at the title/abstract level. Publications identified as relevant for the Task Force were read in full by at least one member of the Task Force. The literature search was systematic, but not a formal systematic review of the evidence. Available literature was used to inform the discussions and recommendations. The reported standards were reached by consensus among the Task Force members and apply to all settings globally (clinical interpretation, research studies, and tertiary, community and primary care). Consensus was reached after all Task Force members agreed on the final version.

Comparison of measured values to a healthy population

Global Lung Function Initiative (GLI) reference equations for spirometry [10], diffusing capacity [11] and lung volumes [12] should be used to define the expected range of values in healthy individuals.

Summaries of data collected in otherwise healthy individuals provide meaningful benchmarks against which to compare an individual's PFT results. The range of values expected in a healthy population is expressed using population-based reference equations that, ideally, are derived from large and representative samples of healthy individuals (*i.e.* never-smokers, without a history of respiratory disease). There are hundreds of published reference equations for different populations and for each PFT. Comparison of published reference equations and individual results derived from different reference equations demonstrates large differences that may be attributed to real population differences in lung function or simply sampling variability with equations derived from small samples. The lack of standards

TABLE 1 Summary of differences between the American Thoracic Society (ATS) and European Respiratory Society (ERS) 2005 [3] and 2021 interpretation standards

	2005 ATS/ERS statement	2021 ATS/ERS technical standard
General comments	<ul style="list-style-type: none"> Using PFT interpretation to aid in clinical diagnosis and decision making 	<ul style="list-style-type: none"> More emphasis on using PFTs to classify physiology, not make a clinical diagnosis Emphasis on uncertainty of interpretation, especially near LLN
Reference equations	<ul style="list-style-type: none"> Use of race/ethnic-specific equations preferred over using adjustment factors Spirometry: <ul style="list-style-type: none"> In USA: NHANES III recommended In Europe: no specific equations recommended Lung volumes and D_{LCO}: <ul style="list-style-type: none"> In USA and Europe: no specific equations recommended 	<ul style="list-style-type: none"> Recommendation to use GLI reference equations for spirometry, lung volumes and D_{LCO} More emphasis on incomplete understanding of role of race/ethnicity on lung function Clarify that biological sex, not gender be used to interpret lung function
Defining normal range	<ul style="list-style-type: none"> General use of LLN=5th percentile Use of fixed ratio FEV₁/FVC <0.7 not recommended Use of 80% predicted to define normal not recommended 	<ul style="list-style-type: none"> General use of LLN=5th percentile and ULN=95th percentile Use of fixed ratio FEV₁/FVC <0.7 not recommended Use of 80% predicted to define normal not recommended
Bronchodilator response	<ul style="list-style-type: none"> ≥12% and 200 mL in FEV₁ or FVC from baseline 4 doses of 100 µg salbutamol; wait 15 min 	<ul style="list-style-type: none"> >10% of predicted value in FEV₁ or FVC Choice of protocol for administering bronchodilator not specified
Interpretation of change over time	<ul style="list-style-type: none"> Variable changes over time depending on normal <i>versus</i> COPD and time period (within a day, week to week, year to year) 	<ul style="list-style-type: none"> Conditional change score in children FEV₁Q in adults
Severity of lung function impairment	<ul style="list-style-type: none"> Using FEV₁ (includes obstruction or restriction): <ul style="list-style-type: none"> Mild: FEV₁ >70% predicted Moderate: 60–69% predicted Moderate-to-severe: 50–59% predicted Severe: 35–49% predicted Very severe: <35% predicted D_{LCO}: <ul style="list-style-type: none"> Mild: >60% predicted and <LLN Moderate: 40–60% predicted Severe: <40% predicted 	<ul style="list-style-type: none"> For all measures use z-score: <ul style="list-style-type: none"> Mild: –1.65 to –2.5 Moderate: –2.51 to –4.0 Severe: <–4.1
Classification of physiological impairments	<ul style="list-style-type: none"> Airflow obstruction: FEV₁/FVC <5th percentile, using largest VC; lung volumes to detect hyperinflation or air trapping; elevated airway resistance; central/upper airway obstruction Restriction: <ul style="list-style-type: none"> TLC <5th percentile and normal FEV₁/VC Mixed: FEV₁/VC and TLC <5th percentile Gas transfer impairment: <ul style="list-style-type: none"> D_{LCO}, K_{CO} <5th percentile Importance of adjustments for Hb, COHb 	<ul style="list-style-type: none"> Airflow obstruction: FEV₁/FVC <5th percentile, using FVC; lung volumes to detect hyperinflation or air trapping; dysanapsis; non-specific pattern and PRISM; central/upper airway obstruction Restriction: <ul style="list-style-type: none"> TLC <5th percentile Simple <i>versus</i> complex restriction Hyperinflation Mixed Gas transfer impairment: <ul style="list-style-type: none"> D_{LCO} <5th percentile Using V_A, K_{CO} to classify low D_{LCO}

PFT: pulmonary function test; LLN: lower limit of normal; NHANES: National Health and Nutrition Examination Survey; D_{LCO} : diffusing capacity of the lung for carbon monoxide; GLI: Global Lung Function Initiative; FEV₁: forced expiratory volume in 1 s; FVC: forced vital capacity; ULN: upper limit of normal; COPD: chronic obstructive pulmonary disease; FEV₁Q: FEV₁ divided by the sex-specific 1st percentile values of the absolute FEV₁ values found in adults with lung disease; VC: vital capacity; TLC: total lung capacity; PRISM: preserved ratio impaired spirometry; K_{CO} : transfer coefficient of the lung for carbon monoxide; V_A : alveolar volume; Hb: haemoglobin; COHb: carboxyhaemoglobin.

for how to derive and use PFT reference equations has led to considerable confusion in the interpretation of PFT results.

Typically, height, age and sex are used to estimate expected lung function in health, and account for the wide biological variability observed within and between populations. Height *per se* is not a direct determinant of lung size but is a reasonable proxy for chest size. Differences in height and body proportions (*e.g.* leg length and trunk length) have been observed between populations [13]. The determinants of the observed differences in height and chest size are multifaceted and must be considered

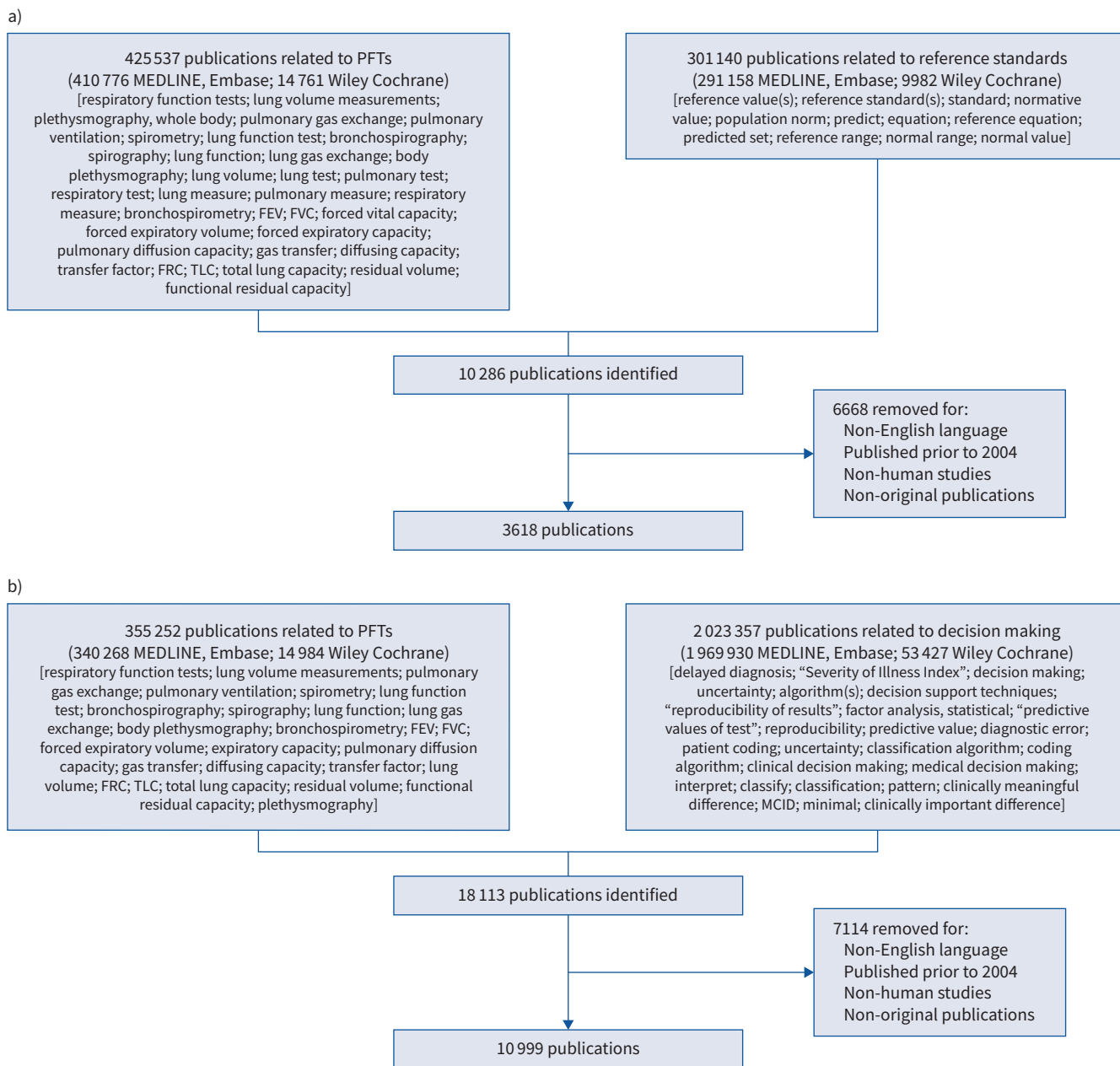


FIGURE 1 Summary of literature search terms and results: a) reference equations for lung function and b) interpretation of pulmonary function tests (PFTs). FEV: forced expiratory volume; FVC: forced vital capacity; FRC: functional residual capacity; TLC: total lung capacity; MCID: minimal clinically important difference.

during PFT interpretation. Age has two important contributions to the expected range of lung function in health. In childhood, somatic growth (*i.e.* height) is strongly linked to chronological age, except during periods of rapid growth and development, such as puberty, when there is asynchrony between height and thoracic volume, and thus disproportionate growth between lung parenchyma and airway calibre [14, 15]. In older adults the rigidity of the chest wall, chest wall muscles and the elasticity of the lung change with the normal ageing process [16, 17]. Sex is an important predictor of lung size, even after accounting for differences in height [18–21]. Thus, while gender identity should be respected, use of biological sex will yield a more accurate prediction of lung function. The effect of gender-affirming hormonal therapy on lung function is poorly understood, so the appropriate reference equation for transgender individuals is currently not known. Timing of gender reassignment, especially during adolescence, may impact lung growth and development, and thus needs to be considered when interpreting results during adulthood [22, 23].

Considerations for individuals in whom standing height cannot be measured are summarised in the technical standards for each PFT [6–8].

The reasons for observed differences in lung function between people around the world are multifactorial and not fully understood. The narrow definition of health may contribute to the observed differences, as “healthy” individuals may include people exposed to risk factors for poor lung health during their lifetime. There are ongoing efforts to better understand the geographical, environmental, genetic and social determinants of health that play a role in explaining these observed differences. The differences by population groupings that were observed in the GLI data may represent genetic differences or health disparities, and thus reflect social and environmental determinants of health. The specific contribution that genetic ancestry plays in the regional differences that were observed in GLI data remains uncertain. Furthermore, assigning ethnicity is challenging. It is important that individuals have their lung function assessed against the appropriate reference population for that individual. The historical approach of fixed adjustment factors for race is not appropriate and is unequivocally discouraged [24, 25]. As there are observed population differences in body proportions [13, 26, 27] and lung function [28, 29], in some contexts it may be relevant to interpret results for an individual relative to that of a similar ancestral grouping, whereas in others it may be more appropriate to compare to the whole population. Caution over which equation is applied is necessary to ensure the same reference equations are applied across serial encounters. An individual’s medical history, symptoms and social circumstances must be considered when applying PFT results to inform clinical decision making.

GLI equations

The GLI reference equations are available for spirometry [10], D_{LCO} [11] and lung volumes [12], and facilitate standardised reporting and interpretation of pulmonary function measurements. These three GLI equations (spirometry, D_{LCO} and lung volumes) are internally consistent, providing a single suite of PFT equations which will avoid discordant results between PFTs and potential misclassification of physiological phenotypes. The GLI equations include the largest samples of healthy individuals and represent a single standard to compare observed measurements applicable across all ages. The GLI equations also explicitly describe the between-subject variability across age, such that the limits of normal are age-specific. Despite the name, the GLI do not include individual data from all populations around the world and do not explicitly consider the factors that may contribute to the observed differences in lung function between populations. Spirometry equations are available for four specific population groupings as well as a composite “other” equation which represents a multi-ethnic population (table 2). The GLI “other” equation was mathematically derived from the four population-specific equations, including the White group, and represents an average across these populations. GLI reference ranges for the forced expiratory volume in 1 s (FEV_1)/forced vital capacity (FVC) ratio appear relatively independent of population differences and will result in more consistent interpretation between populations. GLI D_{LCO} equations and GLI static lung volumes are currently based on measurements predominantly from individuals of European ancestry due to insufficient reference data from other populations.

Further studies regarding the use of reference equations relating to specific population groupings are currently under development, so these recommendations are based on the current evidence designed to

TABLE 2 Summary of Global Lung Function Initiative (GLI) equations for spirometry and current evidence regarding application of these equations in different populations

GLI reference population	GLI data sources	Population/ancestral origin	Considerations
White	Europe, Israel, Australia, USA, Canada, Brazil, Chile, Mexico, Uruguay, Venezuela, Algeria, Tunisia	White (European); Hispanic (European)	Suitable for use in White European populations [36, 175, 176]
Black	African American	Black (North America)	
South East Asian	Thailand, Taiwan, China (including Hong Kong)	Asian	
North East Asian	Japan, Korea		North East Asian equations demonstrate poor fit when applied to contemporary populations [29]
Multi-ethnic	Average of the other four GLI groups	Multiracial; Black South Africa [177]; India [178]; unknown	Indian [178] and South African [177] data based on a single prospective study in children

increase the precision of determining whether the results are outside of the expected range for an individual. There is no single reference equation equally applicable to all populations. There is a trade-off between applying reference equations that are specific to population groupings *versus* a single standard for all. Different approaches may be warranted in different contexts. Therefore, at this time employing the appropriate GLI spirometry equations based on self-reported ancestral origins, if known, should be used as a way to standardise lung function measurements for sex, age and height. If ancestral origins are unknown or uncertain, the GLI “other” equation should be used. PFT reports and research publications must include the specific reference equation that is used.

Differences from the previous recommendations

The 2005 ATS/ERS interpretation strategy [3] recommended the use of the National Health and Nutrition Examination Survey (NHANES) III spirometry reference equations for individuals in North America. The NHANES III spirometry data are included within the GLI equations and, overall, the predicted values are similar, with a few notable differences. NHANES III derived separate equations for Mexican Americans and Caucasians, whereas the GLI equations do not make this distinction, as re-analysis of NHANES III data reveals minimal differences between expected lung function in these populations [30]. The GLI equations span a wider age range (3–95 years) than NHANES III (8–80 years). There have been notable differences observed in predicted values between the two equations for adults older than 65 years [31–33]. The 2005 ATS/ERS interpretation document [3] did not make specific recommendations for reference equations in Europe and elsewhere, although the European Community for Steel and Coal (ECSC) equations have been used widely. There are demonstrable differences between the predicted values from ECSC and GLI, where the predicted GLI values are consistently higher than ECSC values [34–36].

Special considerations for D_{LCO}

The overall recommendation to use the GLI reference equations also applies to D_{LCO} [11]. Interpretation of D_{LCO} values requires adjustment for equipment dead space and barometric pressure (altitude), which should be done by the equipment software before calculating predicted values [11]. Changes in haemoglobin, carboxyhaemoglobin and carbon monoxide back-pressure must also be considered when interpreting results. This is particularly important in situations where patients are being serially monitored for possible drug toxicity and where haemoglobin is subject to large shifts (*e.g.* chemotherapy for cancer) [6, 7]. The clinician must incorporate information about haemoglobin concentrations on an individual basis while interpreting results. It is recommended that the reference value be adjusted for measured haemoglobin concentration.

Special considerations for lung volumes

The overall recommendations for reference equations also apply to the interpretation of lung volumes. GLI [12] and other reference equations for lung volumes adjust for height but not weight. However, lung volumes can be affected by obesity, with significant reductions in functional residual capacity (FRC) and expiratory reserve volume (ERV) at body mass index (BMI) $>30 \text{ kg}\cdot\text{m}^{-2}$ [37, 38], with similar findings in children and adolescents when obesity is defined as >97 th percentile [39]. In extreme obesity, both obstructive and restrictive ventilatory impairment patterns are seen [40]. Nonetheless, measured lung volumes for the majority of obese individuals still fall within the normal range and total lung capacity (TLC) is usually not reduced until BMI $>40 \text{ kg}\cdot\text{m}^{-2}$ [37]. The typical patterns of obstruction and restriction may be altered in obesity; thus, in the context of obesity, results observed outside the normal range need to be interpreted with greater uncertainty [41]. Measurements of lung volumes are also impacted during pregnancy, and results need to be interpreted cautiously both during pregnancy and in the post-partum period [42].

Practical considerations

PFT reports must include the reference equations applied for each index [9]. Caution should be applied to interpretation of results where different reference equations or combinations of reference equations are used for each test (or indices) as there may be differences in the healthy populations used to derive the equations. A change in reference equations must be clearly documented and communicated, as an individual’s results may appear to change based solely on the change in reference equation [34, 43–45]. If reference equations are changed, interpretation of trends should include re-calculation of prior predicted values as well as comparison of raw values to avoid misinterpretation. If standing height, biological sex or ancestral background are not known, the report must clearly state what is assumed.

Validation of reference equations in individual PFT laboratories with a small sample of healthy individuals (*e.g.* 100) is not recommended. Differences due to sampling variability alone can be as large as 0.4 z-score

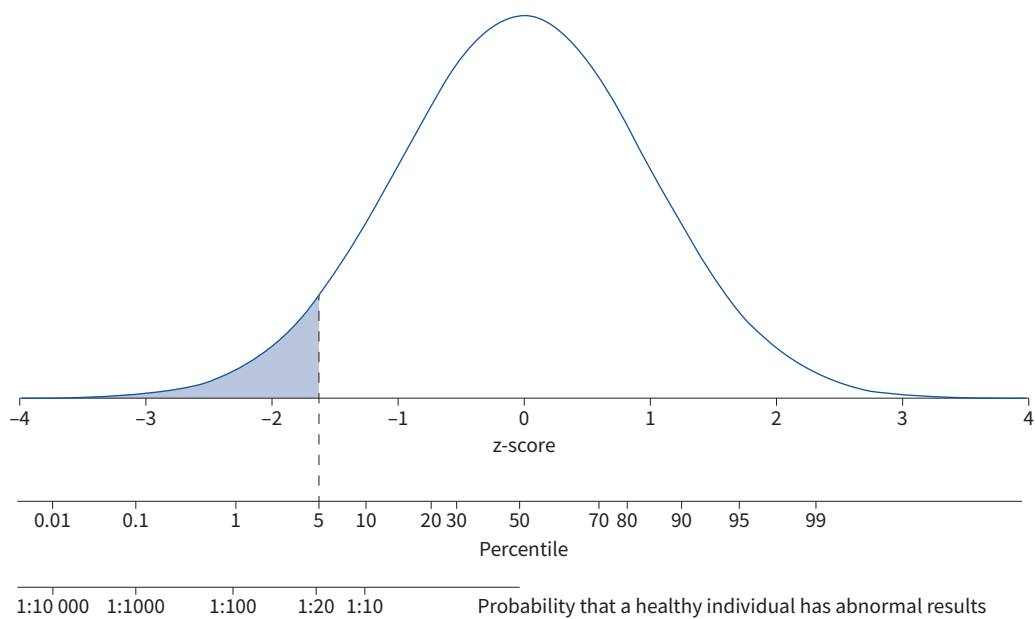


FIGURE 2 The normal distribution with z-scores and percentiles displayed. Percentile can be interpreted as the probability that a healthy individual has results inside the normal range (*i.e.* the false-positive rate).

(6–9% predicted) even when the same equipment and protocols are used and the sample size is at least 1000 [46].

Limits of normal

The 5th and 95th percentile limits (–1.645 and +1.645 z-score) of the healthy population can be used to identify individuals with unusually low or high results, respectively.

Ideally, limits of normal ought to be based on an individual's pre-disease measure or baseline. Further clinical decision making requires relevant thresholds based on prognosis or clinical risk of adverse outcomes. To date, no satisfactory outcome-based thresholds for lung function have been defined; therefore, careful consideration of the medical and exposure history of an individual is necessary when interpreting lung function results when using the limits of normal. Further research to establish a comprehensive disease-specific clinical approach to interpretation (not simply relying on whether results are within or outside the normal range) is necessary. It is the consensus of the Task Force that the percentile limits represent a standardised and unbiased approach to identify values outside the range of expected results from a normal population.

A reference range represents the distribution of values that are expected in a healthy population and the lower limit of normal (LLN) represents a cut-off to define results that are outside the range of values typically observed in health. This approach is used for many clinical outcomes in medicine [47–49]. Population-defined z-scores or percentile values describe the chance the observed result falls within the distribution of values in healthy individuals (figure 2). At the 5th percentile (corresponding to a z-score of –1.645), there is a 5% chance that the results in a healthy individual would be at or below this level, as shown in figure 2. At the 1st percentile, there would be a 1% chance. Since typically for spirometry, low values are considered abnormal, it has become standard to define the LLN as the 5th percentile, accepting that this will result in 5% of healthy individuals having a false-positive result (*i.e.* being incorrectly classified as having an abnormal result). The 5th percentile represents a trade-off between incorrectly classifying a low value in a healthy individual and missing a clinically significant reduction in lung function (*i.e.* increased sensitivity for less specificity compared with using a lower percentile). For tests that may be outside the normal range in either direction (*e.g.* lung volumes or D_{LCO}), the potential for false positives increases to 10% but the probability in a given individual for which these tests are requested based on concerns for lung disease is lower because there is a higher likelihood (pre-test probability) that lung function will be outside the normal range [50]. The LLN does not necessarily indicate a pathophysiological abnormality nor is it a clinically meaningful threshold to diagnose disease. It provides an indication of whether the observed result can be expected in otherwise healthy individuals of similar

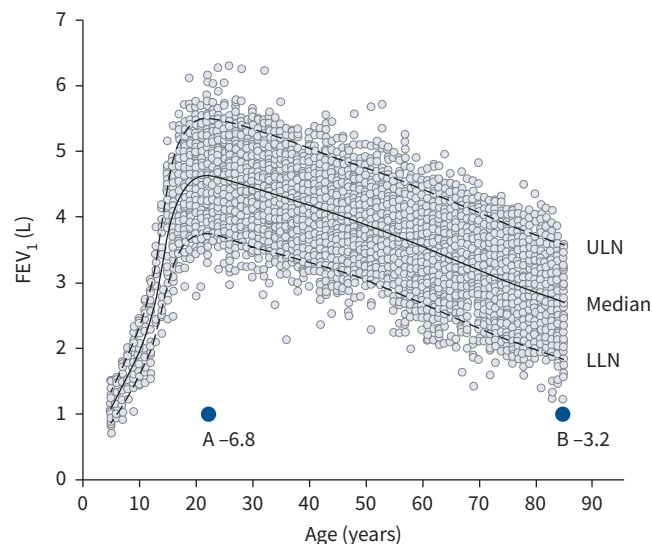


FIGURE 3 Plot of population forced expiratory volume in 1 s (FEV_1) data for males of median height for age between ages 5 and 85 years with the upper limit of normal (ULN; 95th percentile), lower limit of normal (LLN; 5th percentile) and median predicted shown as derived from Global Lung Function Initiative spirometry equations [10]. The LLN for a man aged 22 years is 81.1% predicted but is 67.9% predicted for a man of the same median height aged 85 years. Participants A and B both have an FEV_1 of 1.0 L, giving a z-score of -6.8 for individual A and -3.2 for individual B.

age, sex and height. A result within the expected range for a subject does not exclude the presence of a disease process impairing function. For example, a drop from the 95th percentile to the 10th percentile is a very significant change but still leaves lung function within the normal limits.

The LLN need not be the 5th percentile. With adequate supporting evidence, the LLN could be adjusted lower when PFTs are performed in the absence of elevated risk (*e.g.* screening the general population). For example, when screening a general population, a more conservative lower limit of 2.5% (-1.96 SD or z-score) or even 1% (-2.326 SD or z-score) will reduce the number of false positives. The specific LLN that is used must be clearly documented in PFT reports. Results that are close to the LLN should be interpreted with caution and considered in the context of the individual patient's medical history, physical findings and pre-test probability of disease. This further emphasises that the person interpreting PFTs should be informed of the patient's context and not solely rely on the numbers generated in reports.

The widely used cut-offs of 80% of predicted for FEV_1 (% predicted = observed \times 100 / predicted) and the 0.70 cut-off for the FEV_1/FVC ratio are strongly discouraged [51]. Percent predicted does not take into account the observed age-related changes in measurement variability (figure 3). These "rules of thumb" only approximate the LLN in the mid-range of age, where screening or case finding for obstructive disease is most likely to be conducted (figure 4). The simplicity of these cut-offs has resulted in their use across the age spectrum, leading to systematic misinterpretation of results, particularly for women, children and older adults [52, 53]. For example, the LLN for FEV_1 varies from 81% predicted at the age of 10 years to 68% predicted at the age of 85 years (figure 3 and table 3).

The limits of normal derived from data collected in healthy individuals represents a cross-sectional snapshot of an otherwise healthy population, and the range of values does not represent ideal lung growth and development expected under optimal social and environmental conditions. Therefore, neither simple cut-offs nor the 5th percentile should be used as absolute diagnostic criteria, as there is a gradual increase in risk the further away from the range of values observed in health (figure 2). There is considerable overlap in the range of values in health and disease, resulting in a "zone of uncertainty" (figure 5). Early-life exposures and cumulative environmental exposures have negative effects on growth and development of the lungs that predispose individuals to lung disease in later life [54, 55]. For some ventilatory impairments, development of airflow obstruction is characterised by a slowly progressive decline in FEV_1 relative to FVC [56] and it is likely that early stages of airflow obstruction will be present before the FEV_1/FVC value falls below the LLN.

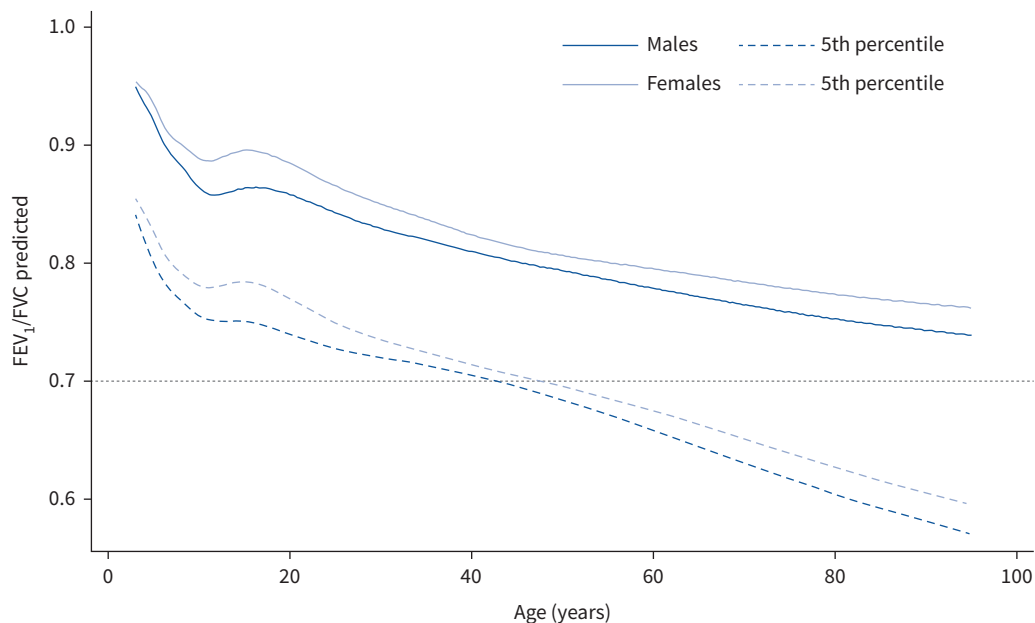


FIGURE 4 Forced expiratory volume in 1 s (FEV_1)/forced vital capacity (FVC) predicted and lower limits of normal (5th percentile) compared with the fixed cut-off of 0.7.

Future directions

There is an urgent need to develop more precise and individualised ways to define what normal lung function should be under ideal growth and environmental conditions. There is a need to understand the factors that contribute to population differences and environmental influences in lung function, and the impact of using ethnic-specific equations on clinical decisions in populations around the world. There is also a need for data to better define the relationship between risk factors, lung function and outcomes that would allow a shift from the interpretive dichotomy of normal/abnormal to a more realistic probability assessment as lung function declines through lower percentiles or z-scores.

Bronchodilator responsiveness testing

Changes in FEV_1 and FVC following bronchodilator responsiveness (BDR) testing should be expressed as the percent change relative to the individual’s predicted value. A change >10% of the predicted value indicates a positive response.

TABLE 3 The 5th percentile values (lower limit of normal (LLN)) for various lung function indices expressed as percent predicted for six individuals

	Male (age 10 years; height 137 cm)	Female (age 15 years; height 162 cm)	Male (age 25 years; height 175 cm)	Female (age 25 years; height 165 cm)	Male (age 80 years; height 175 cm)	Female (age 80 years; height 165 cm)
FEV_1	81.3	80.5	80.5	80.2	69.4	70.0
FVC	81.2	80.4	80.9	79.9	72.0	70.0
FEV_1/FVC	87.4	87.8	86.9	87.2	80.0	80.5
TLC	78.0	79.8	80.0	80.4	77.8	77.6
FRC	70.9	69.9	69.6	72.5	69.8	70.7
RV	40.6	40.9	49.1	52.5	55.7	57.7
D_{LCO}	75.4	77.5	79.0	77.8	72.4	74.5

FEV_1 : forced expiratory volume in 1 s; FVC: forced vital capacity; TLC: total lung capacity; FRC: functional residual capacity; RV: residual volume; D_{LCO} : diffusing capacity of the lung for carbon monoxide. Global Lung Function Initiative reference equations were used for all indices [10–12]. The table demonstrates that the equivalent percent predicted value at the LLN varies considerably for individuals of different ages and for each pulmonary function index, and highlights the potential bias introduced when using percent predicted thresholds for defining normal limits.

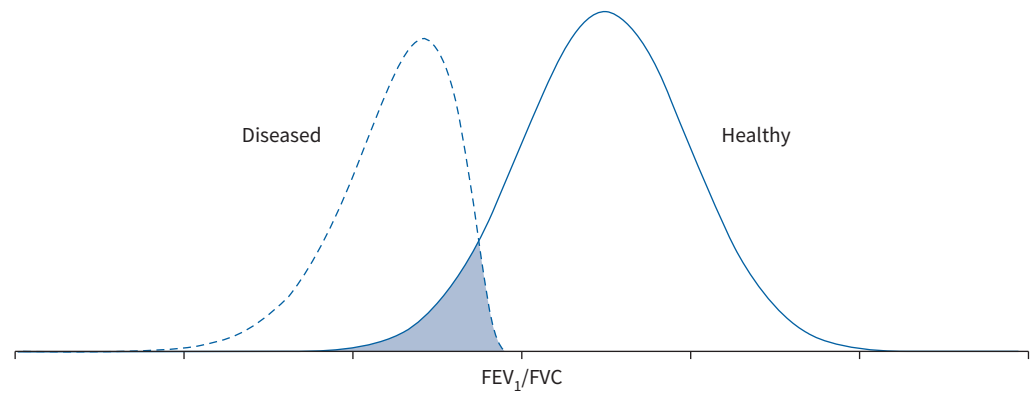


FIGURE 5 Theoretical distribution of health and disease. The shaded area is the zone of uncertainty. FEV₁: forced expiratory volume in 1 s; FVC: forced vital capacity.

When clinically indicated, the BDR test assesses the change in respiratory function in response to bronchodilator administration. The BDR result reflects the integrated physiological response of airway epithelium, nerves, mediators and airway smooth muscle, along with structural and geometric factors that affect airflow in the conducting airways [3, 57–59]. The choice of bronchodilator, dose and mode of delivery is a clinical decision. The relative merits of different protocols (*e.g.* delivered dose) are unclear. Recommended BDR protocols are included in the 2019 ATS/ERS spirometry standard [5]. The concept of a response to bronchodilators must not be confused with “reversibility” of airflow obstruction, which is a qualitative term reflecting the normalisation of FEV₁/FVC (and hence airflow obstruction) after bronchodilator administration [60]. Here we address how to interpret acute changes in lung function after bronchodilator administration and do not consider how BDR can be used to make diagnostic or clinical decisions.

Expressing the results of a BDR test

Interpretation of BDR can employ two approaches: 1) the upper limit of the changes expected in a healthy population or 2) a threshold at which a clinically meaningful event occurs. The upper limit of the changes expected in a healthy population may not be clinically relevant [61]. Although data are limited for clinically meaningful thresholds across a range of diseases and age groups, there is evidence related to survival to support a threshold-based approach [27, 57, 59, 62, 63]. In over 4000 patients referred for BDR in a hospital laboratory, those with BDR >8% of predicted FEV₁ had a lower subsequent mortality than those with BDR below this threshold [62]. Thus, a threshold approach that is supported by both methods (*i.e.* the percentage of predicted value threshold) should be used until further data are available [27].

Established methods to assess the change in FEV₁ and FVC after administration of a bronchodilator include: 1) an absolute change from the initial value, 2) a relative change related to the initial value or 3) a change related to the individual’s predicted value, or a combination of these options. The combination of an absolute and relative change (percentage change) in FEV₁ and FVC from baseline as evidence of BDR was recommended in the 2005 ATS/ERS interpretation statement (*i.e.* >200 mL AND ≥12% increase in FEV₁ and/or FVC) [3]. The major limitation to this approach is that the absolute and relative changes in FEV₁ and FVC are inversely proportional to baseline lung function, and are associated with height, age and sex in both health and disease [57, 59, 62–64]. The use of approaches 1) and 2) to define BDR are no longer recommended.

We recommend reporting the change in FEV₁ or FVC as the increase relative to the predicted value, which minimises sex and height difference in assessing BDR [57, 59, 62]. Two studies of collated epidemiological data in healthy adults reported the upper limit (95% percentile) of the range of bronchodilator response in healthy individuals to be 11.6% and 10.1% of predicted for FEV₁ and 10.2% and 9.6% of predicted for FVC [59, 62]. Similar changes of 8.5% for FEV_{0.75} in young children have been reported [65]. BDR in FVC, rather than FEV₁, has been shown to better reflect the physiological processes of air trapping [66–70]. Based on these considerations, it is recommended that BDR be classified as a change of >10% relative to the predicted value for FEV₁ or FVC (see box 1 for example calculation). This approach avoids misinterpretation due to the magnitude of the baseline lung function level. Over-reliance on strict cut-offs for BDR should be avoided as these cut-offs are prone to the same

BOX 1 Determination of a bronchodilator response

$$\text{Bronchodilator response} = \frac{(\text{post-bronchodilator value (L)} - \text{pre-bronchodilator value (L)}) \times 100}{\text{predicted value (L)}^\#}$$

A change of >10% is considered a significant bronchodilator response.

#: predicted value should be determined using the appropriate Global Lung Function Initiative (GLI) spirometry equation.

For example, a 50-year-old male, height 170 cm, has a pre-bronchodilator forced expiratory volume in 1 s (FEV₁) of 2.0 L and a post-bronchodilator FEV₁ of 2.4 L. The predicted FEV₁ is 3.32 L (GLI 2012 “other” equation).

$$\text{Bronchodilator response} = \frac{(2.4 - 2.0) \times 100}{3.32} = 12.1\%$$

Therefore, their bronchodilator response is reported as an increase of 12.1% of their predicted FEV₁ and classified as a significant response.

limitations as for limits of normal. Importantly, this is not equivalent to a 10% change between pre- and post-bronchodilator measurements.

Changes in forced expiratory flows (*e.g.* peak expiratory flow (PEF) or forced expiratory flow at 25–75% of FVC (FEF_{25–75%})) are highly variable and significantly influenced by changes in FVC, such that pre- and post-bronchodilator measurements are not comparable [3].

Future directions

The recommended BDR threshold balances the available data and consistency across age groups. There were limited data in children and young adults to inform recommendations; further evidence is needed to validate this approach in children. Future research is also needed to understand the impact of bronchodilator protocols (*e.g.* delivered dose) on results. The ability of an acute response to bronchodilators to predict future clinical status other than survival is unclear and BDR does not accurately differentiate between types of airway diseases [71–73]. Further evidence is needed to support anchor-based approaches associated with outcomes other than survival. Finally, there are limited data regarding changes in pulmonary function indices derived from lung volumes, gas transfer and airway resistance following bronchodilator administration.

Natural changes in lung function over time

There are limited data to support a single recommendation for interpreting PFT reproducibility. Two distinct approaches were identified to express natural changes in lung function: conditional change scores for children and FEV₁Q for adults.

The interpretation of a series of lung function measurements and identifying meaningful changes in lung function over time are often used to guide clinical decisions. Ideally, an individual’s pre-disease measure of lung function or baseline should be used as a reference. Comparison with the rate decline observed in a group of healthy individuals can help to determine if rate of decline is greater than what can be expected in health. Accelerated lung function decline, irrespective of baseline lung function, is associated with poor clinical outcomes [74, 75]. Interpretation of serial measurements relies on accurate limits of reproducibility of a PFT index, including the natural changes over time and the changes that would be considered outside the normal biological variability over both short and long periods of time.

Reproducibility

Previous recommendations define a meaningful change as one greater than the biological variability (and measurement error) of a test. An absolute change in FEV₁ (*e.g.* 100 mL) or the relative change from a previous assessment (*e.g.* a 10% change in FEV₁ from baseline in healthy individuals) has historically been used to indicate clinically meaningful changes. However, changes over time have been demonstrated to be dependent on age, sex, baseline lung function and disease severity, limiting the generalisability of these approaches [76, 77]. Furthermore, these limits were derived from population data in healthy individuals and do not necessarily reflect clinically meaningful outcomes for a specific disease or condition [78].

A visual representation of serial measurements (*e.g.* a trend graph) may be included as part of a PFT report. A decline in lung function observed from multiple measurements over time is more likely to reflect a real change in lung function than two measurements alone.

Considerations in children

Lung function measurements in children are more variable than in adults. This is due to both the physiology of the chest wall muscles as well as cognitive development, which may influence test quality and biological variability. Interpretation of serial measurements during periods of rapid growth and development (*e.g.* adolescence and early adulthood) requires special attention to avoid misinterpreting the normal plateau of lung growth. Examination of absolute measures should be used to verify “decline” in this period. Generally, limits of reproducibility applied in children are extrapolated from studies in adults and do not consider the unique developmental aspects of childhood, including how somatic and lung growth are not always synchronous. We identified one recently published study that demonstrates conditional change scores can be used to identify changes in lung function greater than what can be expected in healthy children and young people [77]. The conditional change scores adjust for longitudinal changes in FEV₁ z-score and conditions on the initial FEV₁ value (see box 2). This concept has yet to be validated, extended to adults or applied to other lung function indices, but may be a reasonable tool to facilitate interpretation.

Considerations in adults

In adults over the age of 25 years, FEV₁ typically declines in healthy non-smokers by 30 mL per year [79, 80]; however, this does not necessarily translate into a threshold of change that can be expected within an individual between two repeated measurements. In occupational medicine, where repeated measurements are made annually (or further apart), a 15% threshold has been proposed as a change outside the biological variability of the test and considered clinically relevant [80]. These limits would not necessarily apply to an individual with a chronic progressive lung disease where the follow-up interval is shorter. Individualised approaches that consider the test quality, time interval between tests, an individual’s baseline lung function as well as the clinical findings at the time of measurement are needed for accurate interpretation.

An alternative approach is FEV₁Q, *i.e.* FEV₁ divided by the sex-specific 1st percentile values of the absolute FEV₁ values found in adults with lung disease (0.4 L for women and 0.5 L for men) [81]. FEV₁Q expresses FEV₁ in relation to a “bottom line” required for survival, rather than how far an individual’s result was from their predicted value. Under normal circumstances 1 unit of FEV₁Q is lost approximately every 18 years, and about every 10 years in smokers and the elderly (see box 3). Over a short interval, or even annually, FEV₁Q should remain stable; changes in FEV₁Q may indicate a precipitous change in lung function and can be used as an alternative approach to gauge meaningful changes over time in adults. FEV₁Q is not appropriate for children and adolescents.

Further research

There is a paucity of data describing natural variability in lung function indices within an individual over time across all ages, PFTs and disease groups [82]. Future work is urgently needed to identify a minimum clinically important difference for each lung function test and index that is anchored to disease-specific outcomes. Further research addressing the short-term (months), annual and long-term (years) changes in healthy individuals is urgently needed. Disease-specific anchor-based approaches that link to clinically

BOX 2 Calculation of a conditional change score

The change score is defined as:

$$\frac{zFEV_{1t2} - (r \times zFEV_{1t1})}{\sqrt{1-r^2}}$$

where zFEV₁ at t₁ and t₂ are the observed forced expiratory volume in 1 s (FEV₁) z-scores at the initial and second time-point, and *r* is defined as 0.642–0.04×time (years)+0.020×age (years) at t₁. Changes within ±1.96 change scores are considered within the normal limits.

For example, a 14-year-old male (170 cm) with a lung function drop from –0.78 z-score (90.6% predicted) to –1.60 z-score (80.6% predicted) within 3 months (*r*=0.912) has a corresponding change score of –2.17, which is outside the limits of normal. The same drop over a period of 4 years (*r*=0.762) corresponds to a change score of –1.55, which is within the limits of normal variability.

BOX 3 Calculation of FEV₁Q in adults

FEV₁Q is the observed forced expiratory volume in 1 s (FEV₁) in litres divided by the sex-specific 1st percentile of the FEV₁ distribution found in adult subjects with lung disease; these percentiles are 0.5 L for males and 0.4 L for females. The index approximates the number of turnovers remaining of a lower survivable limit of FEV₁.

For example, a 70-year-old woman with an FEV₁ of 0.9 L would have an FEV₁Q of 0.9 L/0.4 L or 2.25. Values closer to 1 indicate a greater risk of death.

meaningful end-points are strongly recommended to define appropriate thresholds for clinical interpretation.

Severity of lung function impairment

A three-level system to assess the severity of lung function impairment using z-score values should be used; z-scores > -1.645 are normal, z-scores between -1.65 and -2.5 are mild, z-scores between -2.51 and -4 are moderate, and z-scores < -4.1 are severe.

The magnitude of lung function deviation from what is expected of healthy individuals, having accounted for age-dependent variability, can be used to determine the association with objective outcomes such as quality of life or mortality [83–87]. The association between lung function reported as z-scores with all-cause mortality in patients for FEV₁, FVC and D_{LCO} is shown in figure 6 [88]. As lung function impairment is a continuum, setting multiple fixed boundaries to define grades is in some sense artificial and may imply tiered differences that are unfounded.

The previously recommended severity levels for airflow obstruction used percent predicted FEV₁ with five levels using cut values of 70%, 60%, 50% and 35% [3]. The use of percent predicted does not give

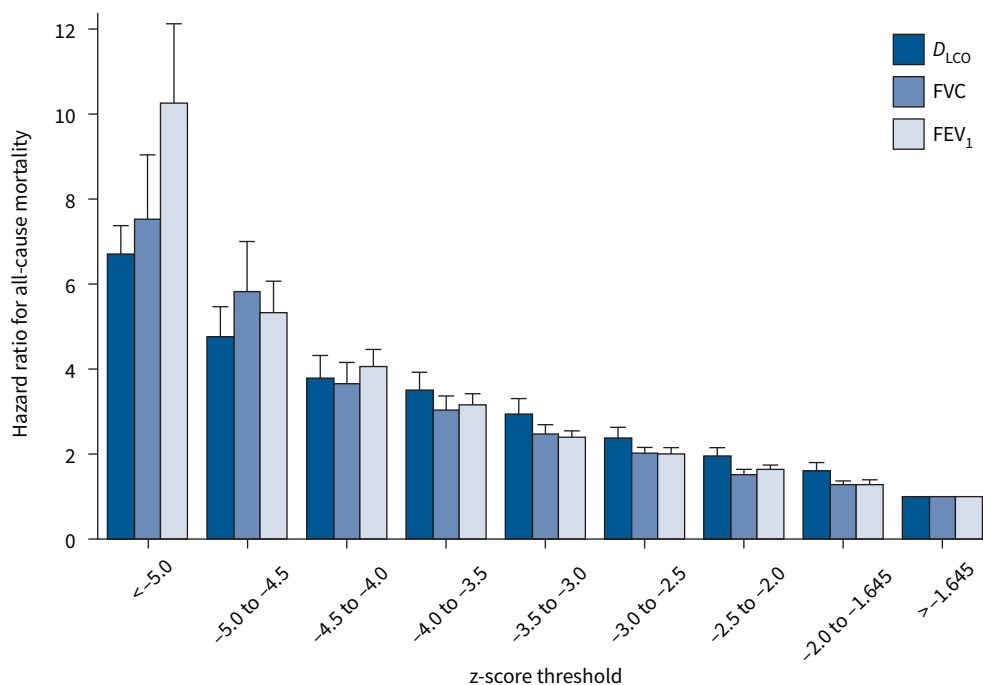


FIGURE 6 Plots of hazard ratios with 95% confidence limit for all-cause mortality for diffusing capacity of the lung for carbon monoxide (D_{LCO}), forced vital capacity (FVC) and forced expiratory volume in 1 s (FEV₁) for bins of z-score values. For FVC and FEV₁, this was from 27021 participants between the ages of 20 and 97 years (comprising 13899 from a population survey, 1094 individuals with chronic obstructive pulmonary disease (COPD) [98] and 12028 individuals seen in the clinic); for D_{LCO}, this was from 13829 clinic patients [88]. Hazard ratios were derived from Cox proportional hazards regression stratified for age and sex. The comparator for mortality was all participants with z-scores above the lower limit of normal (-1.645) who were assigned HR=1.

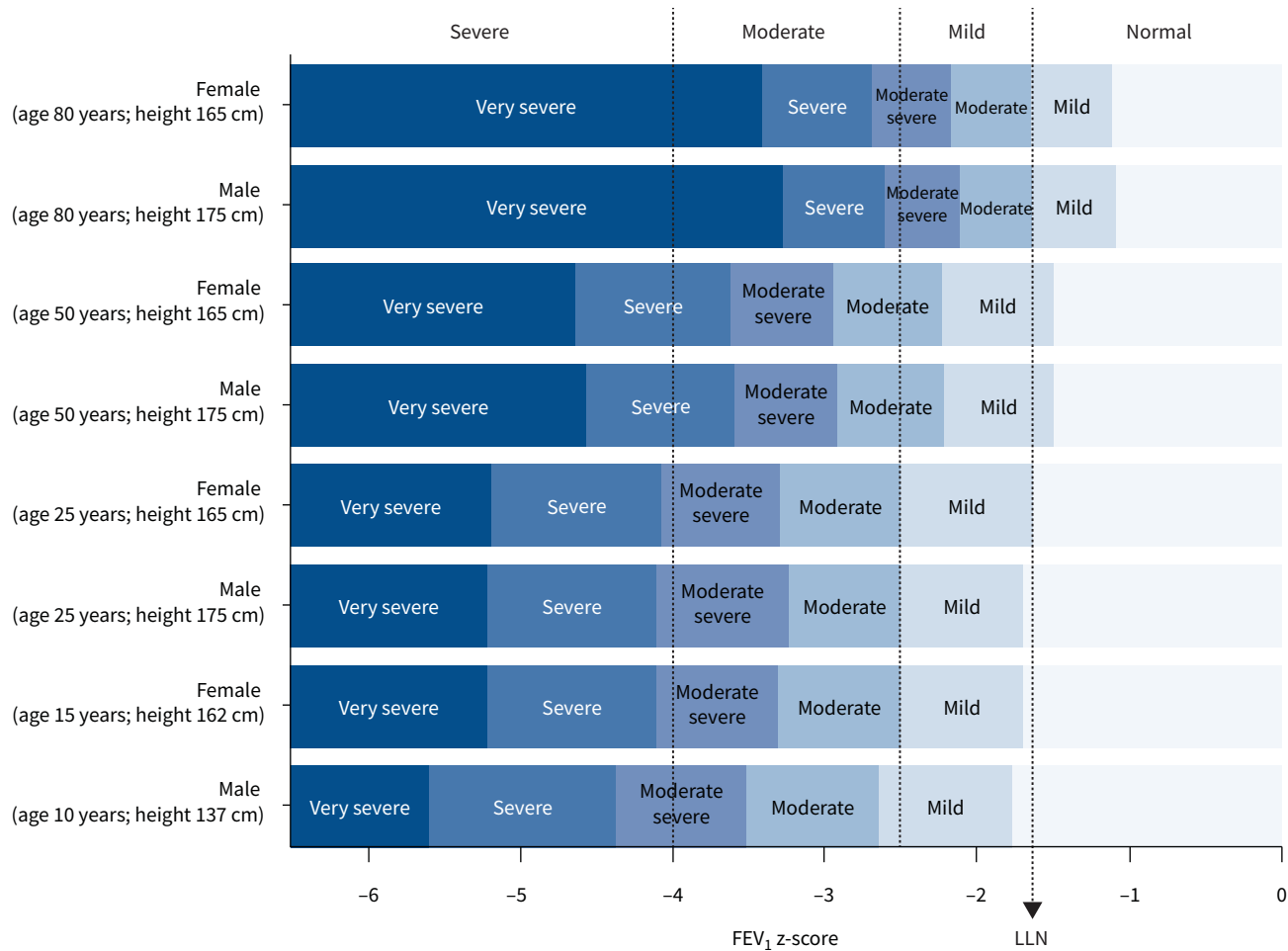


FIGURE 7 A plot of the old 2005 American Thoracic Society/European Respiratory Society Task Force [3] recommended thresholds for degree of lung function reduction grading of airflow obstruction using 70%, 60%, 50% and 35% of predicted forced expiratory volume in 1 s (FEV_1) for eight individuals with the FEV_1 cut-points expressed as z-score values on the abscissa scale. The lower limit of normal (LLN) at the 5th percentile (-1.645 z-score) is shown as a dashed arrow and the two new proposed cut levels of z-scores of -4 and -2.5 are shown as dashed lines with the new gradings above.

uniform gradations across age [53, 89]. To account for an individual's sex, height, age and ethnic background, the previous severity scale for airflow obstruction was adapted for z-scores with cut values of -2 , -2.5 , -3 and -4 [88, 90]. z-score cut levels between -1.65 and -2.5 have little difference in risk of death and were therefore merged into a "mild" group (figure 7). Individuals with z-scores between -2.51 and -4 exhibit a moderate risk of mortality and these categories were therefore merged into the category called "moderate". The proposed three-scale system reduces the previous two lower categories into one for mild impairment and extends the moderate levels to improve the fit for gradation of mortality risk [88].

Importantly, the severity of lung function impairment is not necessarily equivalent to disease severity, which encompasses quality of life, functional impairment, imaging, *etc.* Disease severity will be influenced by many other possible clinical features not related to lung function impairment such as anaemia, neuromuscular weakness or drug side-effects, to mention just a few. There are numerous questionnaires designed and validated to assess the severity of symptoms and impairment [91–94]; these are outside the scope of this work. In addition, the association between the proposed gradations and survival in children has not been evaluated.

Rationale for z-scores

z-scores express how far an observed lung function value is from the predicted value after accounting for sex, age, height and ancestral grouping, expressed in standard deviations. This is the method recommended for

determining the limit of normality and for stating the degree of lung function impairment. Percentile values are easily derived from z-scores, and explicitly indicate the probability a healthy individual would have a result below this level and where the individual's result lies in relation to the healthy population. Percentile values are useful in assessing results around the normal range but are less useful for extreme values.

T-scores are similar to z-scores but are expressed in the number of standard deviations an observation is below a maximum predicted value achieved during early adulthood for an individual of the same sex, height and ancestral grouping [95]. However, T-scores assume that population level maximum lung function can be maintained throughout adulthood. Furthermore, T-scores cannot be applied to children and young adults.

Assessing severity of impairment using z-scores is more consistent across age and sex than percent predicted [88, 90]. Figure 7 shows the previously recommended categories for airflow obstruction using percent predicted (*i.e.* 70%, 60%, 50% and 35% defining mild, moderate, moderately severe, severe and very severe) for eight different people at their respective z-score values. Older age has the greatest differences in interpretation between percent predicted and z-score cut-points such that the 80-year-old individual is deemed to have a mild impairment using percent predicted thresholds when their lung function is within the normal range using z-scores. Figure 7 shows that percent predicted creates problems in equitable grading with mild impairment, but z-scores have problems with respect to severe grading in older subjects as many older individuals will be classified as severe.

Other approaches

In adults, FEV₁Q has been found to be better than z-scores, percent predicted and FEV₁ standardised by powers of height (*e.g.* FEV₁·Ht⁻² and FEV₁·Ht⁻³) in predicting survival [81, 96, 97], chronic obstructive pulmonary disease (COPD) exacerbations [98] and adverse health outcomes [99]. There is also evidence that the FEV₁Q approach may be more useful to differentiate lung function impairment within the “severe” group and in older adults [81, 96, 97], but FEV₁Q has not been adequately explored in children and adolescents.

Considerations in the elderly

Reference equations for lung function indices represent the range of values expected in healthy individuals of the same sex, height and age. The number of healthy individuals over age 80 years in reference cohorts is smaller and may represent a selected population of survivors. In older individuals, interpreting lung function as an absolute measure, such as FEV₁Q, may be more meaningful than using reference equations. There is evidence that extrapolating predicted values from a younger age may address some of these issues [32, 33]. Nonetheless, interpretation at the extremes of the age and/or height ranges has greater uncertainty and requires careful consideration.

Future directions

Assessing the severity of lung function reduction should be linked to important clinical outcomes (survival, exacerbations, admissions, symptoms, imaging, *etc.*) which may be disease-specific. FEV₁Q and other reference-free indices should be explored in this way. FEV₁Q highlights that survival better relates to how far the FEV₁ is above a “survivable bottom line” rather than how far it has dropped from a predicted value. Simpler grading with fewer tiers as proposed should be investigated for a broader range of lung function indices and for different diseases in both children and adults.

Classification of physiological impairments by PFTs

The interpretation of PFTs should focus on values of airflow, lung volume and gas transfer measurements to recognise patterns of altered physiology. PFTs alone should not be used to diagnose a specific pathological condition.

PFT interpretations should be clear, concise and informative to help understand whether the observed result is normal and, if not, what type of physiological impairment is likely involved. In addition, repeated assessment of PFTs is important to detect clinically meaningful deviations from an individual's previous results. Here, we will review the interpretation of measurements made by spirometry, lung volumes and D_{LCO} as they relate to underlying pathophysiology.

Routine PFTs address three functional properties of the lungs: 1) airflow (inspiratory and expiratory), 2) lung volumes and capacities (TLC, residual volume (RV) and FRC), and 3) alveolar–capillary gas transfer (measurement of carbon monoxide uptake over time), expressed as the transfer capacity of the lung for carbon monoxide (T_{LCO}), also known as the diffusing capacity of the lung for carbon monoxide

(D_{LCO}). Abnormalities in these three functional properties are conventionally classified as obstructive ventilatory, restrictive ventilatory and gas transfer limitations or impairments (table 4).

Ventilatory impairments defined by spirometry

Airflow limitation and airflow obstruction

Expiratory airflow is generally assessed by spirometry, with the most important indices being FEV₁, FVC and the FEV₁/FVC ratio. In normal lungs, airflow is determined by the magnitude of expiratory driving pressure (expiratory muscles and elastic recoil) and the size and viscoelastic properties of the lungs and airways. Maximal airflow is generally assessed spirometrically and may be limited by different diseases that lead to different outcomes: 1) impaired expiratory muscle function (weakness or poor effort; neuromuscular ventilatory impairment), reduced elastic recoil or reduced chest wall expansion which reduce PEF, FEV₁ and FVC, with a variable FEV₁/FVC ratio; 2) physical obstruction of a central airway (*i.e.* outside of lung parenchyma), which can affect the trachea/major bronchi and leads to a disproportionate reduction in PEF compared to FEV₁ with variable FEV₁/FVC ratio; and 3) intrapulmonary airflow obstruction produced by premature airway collapse, bronchoconstriction or airway inflammation/wall thickening/oedema leading to airway narrowing. These obstructed airways reduce PEF and FEV₁ to a much greater extent than any reduction in FVC, so the FEV₁/FVC ratio is characteristically low [100–102].

While we recognise the normal physiological events involved in expiratory “airflow limitation”, we use the term “airflow obstruction” to refer to pathological reduction in airflow from the lungs that leads to a reduced FEV₁/FVC ratio.

An obstructive ventilatory impairment is defined by FEV₁/FVC (or VC) below the LLN, which is defined as the 5th percentile of a normal population (figure 8 and table 5). This spirometric definition of airflow obstruction is consistent with the 1991 ATS [103] and 2005 ATS/ERS [3] recommendations; however, it contrasts with the definitions suggested by the Global Initiative for Chronic Obstructive Lung Disease (GOLD) [104] and the ATS/ERS [105] guidelines on COPD, which use a fixed FEV₁/FVC value of 0.7 to define an obstructive ventilatory impairment.

The earliest changes associated with respiratory diseases that produce airflow obstruction are thought to occur in the smaller, more distal airways [106]. Since the total cross-sectional area of the small airways is very large, they offer little resistance to airflow at high lung volumes and impairment limited to these airways has little impact on maximal airflow as measured by FEV₁ [107]. However, as exhalation proceeds during a maximal forced exhalation manoeuvre, these smaller airways decrease in calibre, with a marked increase in resistance, which can reduce expiratory flow substantially at lower lung volumes. In addition, loss of elastic recoil with emphysematous changes in lung parenchyma also contribute to reduction in maximal expiratory flow [108]. This results in a slowing of flow in the terminal portion of the spirogram, even when the initial part of the spirogram is barely affected [100–102]. These reductions in late- or mid-expiratory flow are best appreciated by examination of the flow–volume loop, where a characteristic

TABLE 4 Functional classification of common impairments assessed by conventional pulmonary function tests and their pathophysiological determinants

Obstructive ventilatory impairments[#]	Narrowing of the airways in the lung by physical obstruction or by dynamic airway collapsing. More proximal airway properties determine airflow resistance at large lung volumes and drive the FEV ₁ /FVC measurement; more distal airway properties determine airflow resistance at small lung volumes and drive flow measurements later in a maximal exhalation. Because airway obstruction impairs lung emptying, it is often accompanied by air trapping and hyperinflation that may reduce the FVC but is more directly assessed by the RV measurement.
Restrictive ventilatory impairments[#]	Reduction in the size of the lung. This may reflect lung parenchymal or an inability to fully inhale due to extrapulmonary factors (<i>e.g.</i> weakness, chest wall abnormalities and obesity). Lung restriction reduces FEV ₁ , FVC (but not the FEV ₁ /FVC ratio) and TLC.
Gas transfer impairments	Reduction in transport of gas (carbon monoxide transfer as a surrogate for oxygen) between the alveolar spaces and alveolar capillary blood. This may be due to a reduction in alveolar surface area, abnormal alveolar–capillary membrane properties or reduced pulmonary capillary blood (haemoglobin) volume. Impaired gas transfer is generally assessed by analysis of carbon monoxide uptake during a breath-hold (D_{LCO}). Some conditions can lead to an increase in gas transfer.

FEV₁: forced expiratory volume in 1 s; FVC: forced vital capacity; RV: residual volume; TLC: total lung capacity; D_{LCO} : diffusing capacity of the lung for carbon monoxide. [#]: many authorities also use the term “ventilatory impairments” to group obstructive and restrictive impairments.

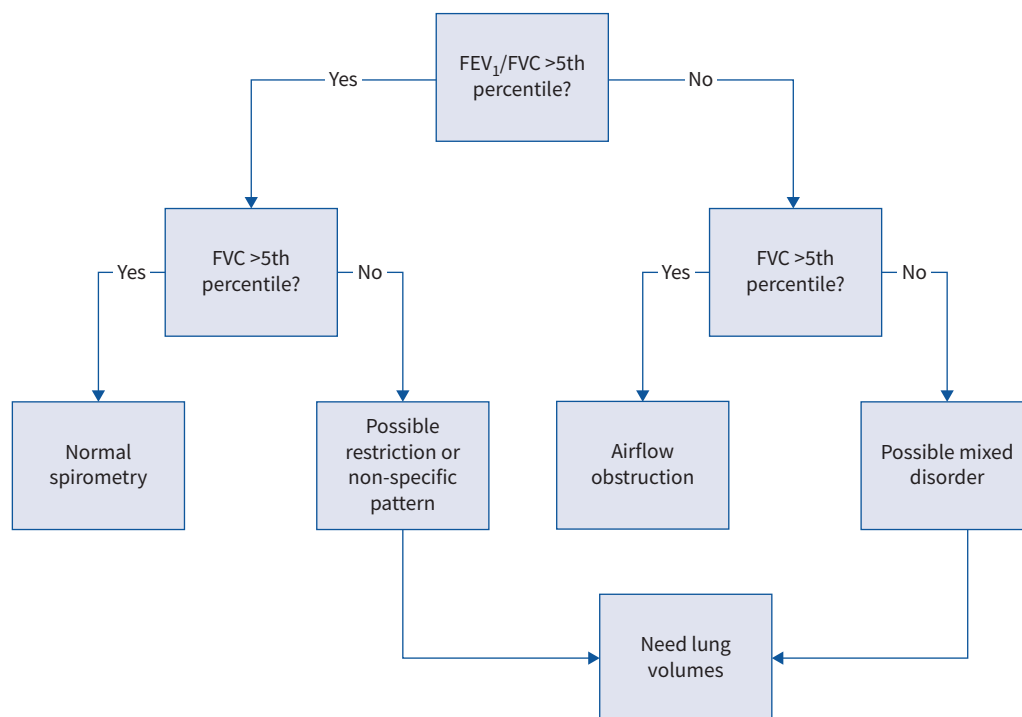


FIGURE 8 Approach to interpretation of spirometry. Beginning with the forced expiratory volume in 1 s (FEV_1)/forced vital capacity (FVC) ratio, determine whether obstruction is present based on whether the ratio is low (right side of figure). If obstruction is present, then assess the FVC to determine whether there is simply obstruction or whether there may be concomitant restriction (“mixed disorder”). Measurement of total lung capacity (TLC) will define restriction, so if TLC is normal, then there is only obstruction, but if TLC is low, then there is concomitant restriction. If FEV_1/FVC is normal, signifying no obstruction (left side of figure), then once again assess FVC. If FVC is normal, then spirometry is normal, but if FVC is low, then there may be possible restriction. This must be determined by measurement of TLC. If TLC is low, then spirometry is consistent with restriction. If restriction is ruled out by a normal TLC, then the pattern of impairment of low FVC with normal FEV_1/FVC has been termed the possible restriction or “non-specific” pattern, which may include diseases causing obstruction or restriction. Restriction presenting as the non-specific pattern is often caused by a chest wall or neuromuscular disorder.

concave shape is thought to reflect small airway dysfunction (figure 9b and c) compared to the normal flow–volume curve (figure 9a).

A number of attempts have been made to quantify these small airway impairments, especially when FEV_1 and FEV_1/FVC are normal (“isolated small airway dysfunction”) [109]. A common approach is to measure the average flow between 25% and 75% of exhaled FVC ($FEF_{25-75\%}$); however, mid-range flow measurements during a forced exhalation are highly variable, poorly reproducible and not specific for small airway disease in individuals [110]. Furthermore, mid-range flow measurements usually do not add to clinical decision making beyond information contributed by FEV_1 , FVC and FEV_1/FVC [111]. There is insufficient evidence to support the use of spirometry to identify small airway dysfunction [112]. There has been recent interest in FEV_3/FVC [113] or FEV_3/FEV_6 [114] providing more sensitive indication of airflow obstruction in adults when FEV_1/FVC is still in the normal range. Other tests such as oscillometry, multiple-breath washout and imaging may also provide evidence of airflow obstruction when FEV_1/FVC is normal [115].

Dysanapsis and other patterns of impairment in FEV_1 , FVC and FEV_1/FVC

For healthy individuals, the meaning of a low FEV_1/FVC ratio accompanied by FEV_1 within the normal range is unclear. This pattern may be due to “dysanaptic” or unequal growth of the airways and lung parenchyma [116]. While this pattern has been thought to be a normal physiological variant [103], new data suggest that it may be associated with the propensity for obstructive lung disease [117, 118]. Factors

TABLE 5 Classification of ventilatory impairments defined by spirometry (reduced or elevated results are defined by the lower and upper limits of normal, respectively)

	FEV ₁	FVC	FEV ₁ /FVC	Comments
Obstructive impairments	Normal/↓	Normal	↓	
Restrictive impairments	↓	↓	Normal/↑	TLC reduced to confirm
Non-specific pattern [121]	↓	↓	Normal	TLC normal (additional testing may be helpful, e.g. bronchodilator response, R _{aw}); when TLC is not available, this pattern has been described in population-based studies as preserved ratio impaired spirometry (PRISm) in current and former smokers [122]
Muscle weakness	↓	↓	Normal	Lack of sharp PEF
Suboptimal effort	↓	↓	Normal	Lack of sharp PEF
Mixed disorder	↓	↓	↓	Need lung volumes to confirm
Dysanapsis [118]	Normal	Normal/↑	↓	May be normal variant

FEV₁: forced expiratory volume in 1 s; FVC: forced vital capacity; TLC: total lung capacity; R_{aw}: airway resistance; PEF: peak expiratory flow.

associated with this pattern in healthy people included male sex, younger age and taller stature, with higher FVC above predicted and higher terminal flows as seen by FEF_{75%} [119]. A high FVC with a low RV can be seen in this instance (normal FEV₁ but low FEV₁/FVC). 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 BDR, D_{LCO}, gas exchange evaluation and measurement of muscle strength or exercise testing.

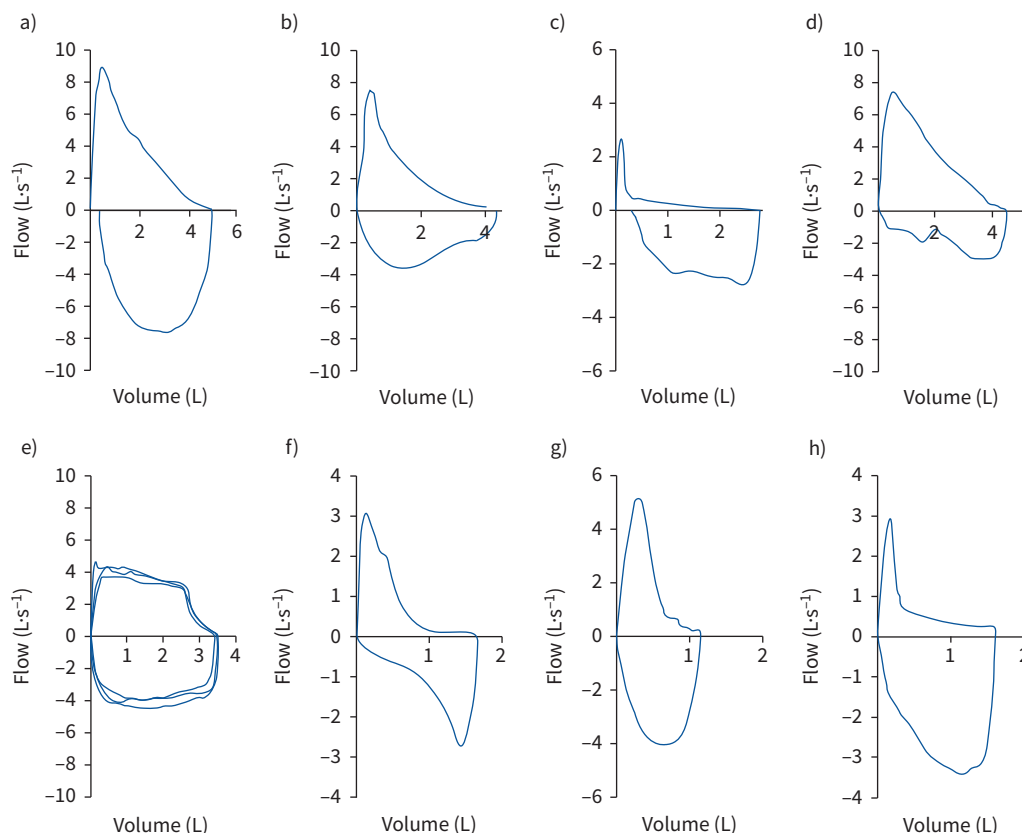


FIGURE 9 Examples of typical flow–volume loop configurations for a) normal, b) mild–moderate obstruction, c) severe obstruction, d) variable extrathoracic obstruction, e) fixed large/central airway obstruction, f) unilateral mainstem bronchial obstruction [179], g) restriction and h) mixed disorder.

The “non-specific” pattern: a low FEV₁ and FVC with normal FEV₁/FVC

The pattern of reduced FVC and/or FEV₁, normal FEV₁/FVC and normal TLC has been termed the “non-specific” pattern. This pattern was described in the 2005 ATS/ERS interpretation statement and was thought to relate to airflow occlusion/collapse [3], but we now know that this interpretation was too simple. Indeed, this pattern can reflect reduced effort, a restrictive ventilatory impairment or be an early consequence of small airway disease with air trapping and/or emphysema [120, 121]. However, measurement of a low TLC is necessary to confirm restriction.

In the setting of reduced effort, the non-specific pattern reflects the failure of the individual to inhale or exhale completely, resulting in a “falsely low” FEV₁ and FVC. It may also occur when the flow is so reduced that the subject cannot exhale long enough to empty the lungs to RV. In this circumstance, the flow–volume curve should appear concave downward towards the end of the manoeuvre. In this case the volume–time curve can also be informative and may help to differentiate between glottis closure and sudden interruption of the expiration due to poor effort, and even other causes.

The non-specific pattern may be an early indicator of a restrictive process in which FVC reduction is not yet accompanied by a reduction in RV. A low TLC under these circumstances would confirm restriction. In contrast, in early obstruction, small airway collapse can reduce FVC and increase RV before the FEV₁/FVC ratio falls. Three-year follow-up of the non-specific pattern has demonstrated continued non-specific pattern in two-thirds of people, with the other one-third having been diagnosed with overt obstructive or restrictive disease. In current and former smokers when TLC is not available (typically in population-based studies), the non-specific pattern has been labelled “preserved ratio impaired spirometry (PRISm)”, which in follow-up has been shown to be associated with both more typical restrictive or obstructive patterns [122–124]. As with any pattern involving a low FVC, TLC should be measured to confirm restriction, as clinically indicated

When the non-specific pattern is observed in an individual performing a maximal, sustained effort, it may be useful to repeat spirometry after treatment with an inhaled bronchodilator. Significant improvement in FEV₁, FVC or both would suggest the presence of some degree of bronchial responsiveness. Another approach is to compare FVC to an untimed slow vital capacity (SVC). If SVC is significantly larger than FVC (>100 mL [125]), it implies that airway collapse is occurring during the forced exhalation [126].

Alternative spirometric indices and supplementary tests assessing ventilatory impairments

The use of VC (*i.e.* the largest VC of SVC and FVC) in place of FVC in the ratio (*i.e.* FEV₁/VC) was recommended in the 2005 ATS/ERS interpretation document [3]. Using VC in this ratio for identifying obstruction may be more sensitive but not as specific compared to FEV₁/FVC [127]. The recording of FVC is easier to standardise because there are many ways to record VC, some using different equipment, and VC is very dependent on the preceding flow and volume histories [128]. In health, FVC does not differ significantly from VC [10]. The use of FVC for the FEV₁/FVC ratio should be used as they both should come from forced expiratory manoeuvres using the same equipment and there are robust reference equations for FEV₁/FVC but not for FEV₁/VC. Using the previously recommended FEV₁/VC to diagnose airflow obstruction will increase the uncertainty about the validity of the diagnosis especially in the older population.

In adults, FEV₆ may be substituted for FVC and appears accurate in diagnosing obstruction [129–133], but this only applies if the appropriate LLNs [134] for FEV₁/FEV₆ are used (GLI equations do not include FEV₆). FEV₂ or FEV₃ have also been shown to be useful surrogates for the estimation of FVC in terms of providing an accurate diagnosis of obstruction [135].

Another measure of an obstructive ventilatory impairment derived from spirometry is inspiratory capacity (IC). A reduction in IC usually reflects an elevated FRC due to air trapping. IC, when expressed relative to TLC, correlates closely with acute exacerbations and survival in individuals with COPD, and reduction in IC during exercise is an important determinant of dyspnoea and exercise intolerance [136].

Multiple other indices derived from analysis of the forced expiratory manoeuvre, such as measures of the slope or curvature of the flow–volume loop, have been identified [137]. In the future, techniques using artificial intelligence/machine learning (AIML) of the expiratory flow–volume loop may offer more accurate assessments of small airway function [138].

In people with early manifestations of lung disease, and especially in children, spirometry values can be normal even in those with confirmed disease. Other measurements of airway function may supplement

spirometry in assessing ventilatory impairments. Airway resistance (R_{aw}) measured by body plethysmography, and its volume-related measures of specific R_{aw} (sR_{aw}) or specific airway conductance (sG_{aw}), are not commonly used to identify airflow obstruction. They are more sensitive for detecting narrowing of extrathoracic or large central intrathoracic airways than of more peripheral intrathoracic airways. However, measurements of respiratory system resistance by the non-invasive techniques of oscillometry, which require only tidal breathing, may be useful in individuals who are unable to perform a maximal forced expiratory manoeuvre, including very young children [139–142].

Central and upper airway obstruction

Central airway obstruction and upper airway obstruction occur in the airways outside lung parenchyma. These may occur in the intrathoracic airways (intrathoracic trachea and main bronchi) or extrathoracic airways (pharynx, larynx and extrathoracic portion of the trachea). These conditions in the early stages may not lead to a decrease in FEV_1 and/or FVC, but PEF can be severely reduced. The indices presented in table 6 may help to distinguish intrathoracic from extrathoracic airway obstruction. Therefore, an increased ratio of FEV_1 (mL) to PEF ($L \cdot \text{min}^{-1}$) can alert the clinician to the need for an inspiratory and expiratory flow–volume loop [143]. An FEV_1/PEF ratio $>8 \text{ mL} \cdot \text{L}^{-1} \cdot \text{min}^{-1}$ in adults suggests the presence of central or upper airway obstruction [144]. Poor initial effort can also affect this ratio. Importantly, a progressively severe upper airway obstruction will ultimately reduce FEV_1 and the FEV_1/FVC (VC) ratio.

Examination of the expiratory flow–volume loop can be very helpful in assessing an upper airway obstruction. When a forced expiratory effort is acceptable, the *repeatable* pattern of a plateau of forced inspiratory flow in the presence of relatively normal expiratory flow suggests variable extrathoracic upper airway obstruction (figure 9d). Conversely, the pattern of a repeatable plateau in forced expiratory flow with relatively normal inspiratory flow suggests variable, intrathoracic central airway obstruction. The pattern of a repeatable plateau in both forced inspiratory and expiratory flows suggests fixed central or upper airway obstruction (figure 9e). With unilateral mainstem 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 (figure 9f). In this instance, during forced expiration, flow initially diminishes during forced expiration as the rapidly emptying regions of the lung empty, but then plateaus in the mid-portion of the expiratory loop as the slower emptying regions now dominate expiratory flow. Another pattern of flow oscillations (sawtooth pattern) may be occasionally observed on either the inspiratory or expiratory phase and likely represents a mechanical instability of the airway wall. The absence of classic spirometric patterns for central airway obstruction does not accurately predict the absence of pathology [145]. As a result, clinicians need to maintain a high degree of suspicion for this problem and refer suspected cases for direct endoscopic inspection or imaging of the airways.

Ventilatory impairments defined by lung volume measurements

Spirometry can only suggest a restrictive pattern and lung volume measurements are necessary to confirm this. Lung volume measurements start with determinations of FRC by gas wash-in/washout analyses or body plethysmography. Thereafter, expiration to RV and inspiration to TLC define fractional lung volumes.

Typically, measurement of TLC and fractional lung volumes discussed in this section add little to spirometric measurements in identifying an obstructive ventilatory impairment; however, these measurements may be helpful in the setting of borderline or atypical spirometric patterns [146–149]. An increase in RV or RV/TLC above the 95th percentile may indicate hyperinflation or air trapping due to the presence of airway obstruction [102]. Indeed, one of the earliest manifestations of small airway disease is

TABLE 6 Lung function indices capable of differentiating extrathoracic from intrathoracic obstruction in adults [142–144, 155]

	Extrathoracic obstruction		Intrathoracic obstruction
	Fixed	Variable	
PEF	Decreased	Normal or decreased	Decreased
$FIF_{50\%}$	Decreased	Decreased	Normal or decreased
$FIF_{50\%}/FEF_{50\%}$	~ 1	< 1	> 1

PEF: peak expiratory flow; $FIF_{50\%}$: forced inspiratory flow at 50% FVC; $FEF_{50\%}$: forced expiratory flow at 50% FVC; FVC: forced vital capacity.

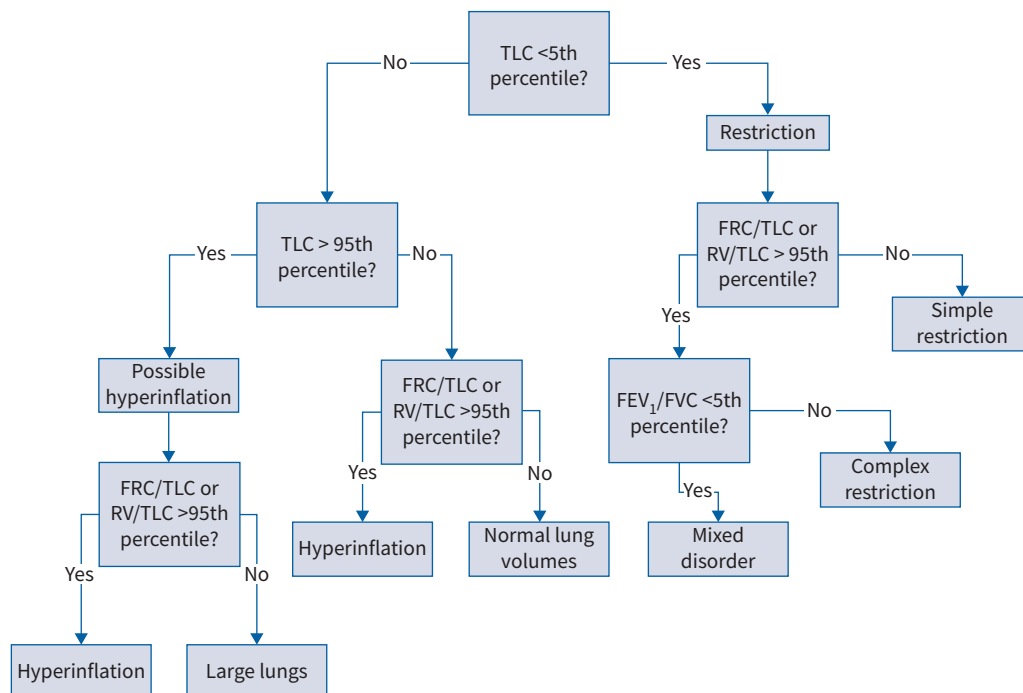


FIGURE 10 Approach to interpretation of lung volumes. Beginning with total lung capacity (TLC), determine whether restriction is present based on whether TLC is low (right side of figure). If restriction is present, then assess the relative size of functional residual capacity (FRC) or residual volume (RV) to the TLC: if FRC/TLC or RV/TLC are elevated, then determine if there is airflow obstruction based on the forced expiratory volume in 1 s (FEV_1)/forced vital capacity (FVC) ratio. If obstruction is present, then this is a mixed disorder, but if not present, then it may be a form of “complex restriction”, which implicates more than one process occurring to lower FVC out of proportion to the reduction in TLC. Obesity or neuromuscular disease are common causes of complex restriction. If the FRC/TLC and RV/TLC ratios are normal in the setting of reduced TLC, then simple restriction is present. If TLC is normal, ruling out restriction (left side of figure), then the next step is to determine if lung volumes are overall normal, or the individual has large lungs, or they may be hyperinflated, by following the pathways indicated. Note that hyperinflation may occur with TLC, FRC and RV, or may occur with FRC or RV alone; in the former situation, the rise in TLC indicates loss of elastic recoil, so is likely due to emphysema, whereas in the latter situation, the increase in FRC or RV without increase in TLC may be seen in chronic bronchitis or asthma.

an increase in RV or RV/TLC due to premature airway closure and air trapping. With progression, lung hyperinflation and air trapping are reflected by increases in FRC or FRC/TLC and often in TLC. An increased FRC/TLC indicates a reduced IC, which is a hallmark of COPD and closely associated with reduced exercise tolerance and dyspnoea [150]. Note that an increased RV/TLC may also be seen with muscle weakness or suboptimal effort and in some restrictive processes when TLC is reduced proportionally more than RV (table 5) [151, 152].

Restrictive impairments

A reduction in lung volumes defines a restrictive ventilatory impairment and is classically characterised by a reduction in TLC below the LLN (5th percentile) (figure 10 and table 7). A typical example is shown in figure 9g. The presence of a restrictive impairment may be suspected from spirometry alone when FVC is reduced, FEV_1/FVC is normal or increased and the flow–volume curve shows a convex pattern (reflecting high elastic recoil). However, a reduced FVC by itself does not prove a restrictive ventilatory impairment. Indeed, it is associated with a low TLC less than half the time [153, 154]. Conversely, in adults, a normal FVC and FEV_1/FVC are highly reliable at ruling out restriction as measured by low TLC [153]. Note that a high PEF with normal FEV_1 may be seen in early interstitial lung disease before restriction limits FVC [155].

In most restrictive disease processes, FEV_1 , FVC and TLC are typically reduced in roughly the same proportion; this pattern is known as “simple restriction”. However, some individuals present with a

TABLE 7 Classification of ventilatory impairments defined by lung volumes

	TLC	FRC	RV	FRC/TLC	RV/TLC	Comments
Large lungs	↑	↑	↑	Normal	Normal	Normal variant above ULN
Obstruction	Normal/↑	Normal/↑	↑	Normal/↑	↑	Hyperinflation if FRC/TLC and RV/TLC elevated; gas trapping if only RV/TLC elevated (e.g. COPD)
Simple restriction	↓	↓	↓	Normal	Normal	e.g. ILD
Complex restriction [156]	↓	↓	Normal/↑	Normal	↑	When FEV ₁ /FVC is normal, complex refers to the process contributing to a restrictive process that disproportionately reduces FVC relative to TLC (e.g. small airway disease with gas trapping and obesity)
Mixed disorder	↓	Normal/↓	Normal/↑	Normal/↑	Normal/↑	Typically, FEV ₁ /FVC is reduced (e.g. combined ILD and COPD)
Muscle weakness	↓	Normal/↓	↑	↑	↑	When effort appears sufficient; TLC is reduced especially with diaphragm weakness; RV is increased especially with expiratory muscle weakness
Suboptimal effort	↓	Normal	↑	↑	↑	Especially when effort appears insufficient
Obesity	Normal/↓	↓	Normal/↑	Normal/↓	Normal/↑	ERV low; reduced TLC at very high BMI (>40 kg·m ⁻²) [37]

TLC: total lung capacity; FRC: functional residual capacity; RV: residual volume; ULN: upper limit of normal; COPD: chronic obstructive pulmonary disease; ILD: interstitial lung disease; ERV: expiratory reserve volume; BMI: body mass index.

reduction in FVC that is out of proportion to the reduction in TLC, indicating a disproportionately elevated RV. This pattern is termed “complex restriction” and is associated with processes that impair lung emptying, such as neuromuscular disease, chest wall restriction or occult obstruction with gas trapping. When associated with a low FEV₁/FVC ratio, it is termed a “mixed” disorder, indicating the presence of both significant airflow obstruction and restriction [156].

Obstructive impairments

Obstructive ventilatory impairments are generally assessed with spirometric measurements of expiratory airflow. As noted earlier, however, there are specific lung volume patterns associated with airflow obstruction that generally reflect hyperinflation/air trapping. These patterns involve reduced VC, IC and FVC with increased FRC and RV. Obstructive diseases, because they interfere with intrapulmonary gas mixing, may also have important effects on gas dilution or washout techniques to measure FRC, alveolar volume (V_A) and TLC. In these conditions, TLC assessed by gas dilution techniques will be low since only communicating gas volume is measured. In the presence of airway disease, a low TLC from a single-breath test (such as V_A from the D_{LCO}) should not be interpreted as demonstrating restriction, since such measurements systematically underestimate TLC. The same is true of measuring lung volumes by multiple-breath helium dilution or nitrogen washout [157]. The degree of underestimation of lung volume increases as airflow obstruction and regional maldistribution of gas worsen. In the presence of severe airflow obstruction, TLC can be underestimated by a gas dilution method by as much as 3 L, greatly increasing the risk of misclassification of the type of physiological phenotype [158–160]. A method of adjusting the single-breath V_A for the effect of airflow obstruction has been published but needs further validation [125, 161]. In the case of severe airflow obstruction, lung volume may be overestimated by body plethysmography, possibly due to heterogenous time constants (resulting in underestimation of alveolar pressure by mouth pressure) and increased extrathoracic airway compliance [160].

Mixed ventilatory impairments

A mixed ventilatory impairment is characterised by the coexistence of obstruction and restriction, and is defined physiologically when both FEV₁/FVC and TLC are below the LLN (5th percentile). Since FVC may be equally reduced in either obstruction or restriction, the presence of a restrictive component in an obstructed individual cannot be inferred from simple measurements of FEV₁ and FVC. A typical example is presented in figure 9h. If FEV₁/FVC is low, FVC is below its LLN and there is no measurement of TLC by body plethysmography, it is possible that the reduction in FVC is due to an increased RV, but a superimposed restriction of lung volumes cannot be ruled out [162]. Conversely, when FEV₁/FVC is low and FVC is normal, a superimposed restriction of lung volumes can almost always be ruled out [153, 154]. Mixed obstruction and restriction commonly involves the combination of a pulmonary parenchymal disorder plus a non-pulmonary disorder, such as COPD plus congestive heart failure [163]. In cases where expiratory airflow obstruction and restriction are concomitantly present, the sensitivity of a reduced FEV₁/

TABLE 8 Summary of types of spirometrically defined and lung volume defined ventilatory impairments

Ventilatory impairments	Patterns
Obstruction	<ul style="list-style-type: none"> • FEV₁/FVC <5th percentile • Decrease in flow at low lung volume may reflect small airway disease in individuals [100, 101, 108] • Concomitant decrease in FEV₁ and FVC most commonly due to poor effort but may reflect airflow obstruction or a restrictive pattern; recommend lung volumes • Measurement of absolute lung volumes may assist in diagnosis and assessment of hyperinflation [108] • Measurement of airflow resistance may assist in diagnosis [139]
Restriction	<ul style="list-style-type: none"> • TLC <5th percentile • Reduced FVC does not prove restrictive impairment but may be suggestive of restriction when FEV₁/FVC is normal or increased • Low TLC from single-breath test not reliable, especially with low FEV₁/FVC [125] • A normal FVC usually excludes restriction [153]
Mixed	<ul style="list-style-type: none"> • FEV₁/FVC and TLC both <5th percentile

FEV₁: forced expiratory volume in 1 s; FVC: forced vital capacity.

FVC or reduced TLC to identify one of these conditions is reduced. Table 8 shows a summary of spirometric and lung volume patterns with obstructive, restrictive and mixed ventilatory impairments.

Gas transfer impairments defined by D_{LCO}

Gas transfer is commonly assessed by measuring the uptake of carbon monoxide (as a surrogate for oxygen) by the lungs. In general, overall carbon monoxide uptake is determined by the alveolar–capillary membrane surface area and diffusion properties, the volume of capillary blood haemoglobin in contact with alveolar gas (V_c), and the reaction rate (θ) between haemoglobin and carbon monoxide. The importance of haemoglobin cannot be overemphasised and all interpretations must have the reference values adjusted for haemoglobin content.

The primary measurements are K_{CO} (the measured carbon monoxide concentration change over time) and V_A (the volume of gas containing carbon monoxide measured by the dilution of an inert tracer gas in the inspired volume). Their product ($D_{LCO}=K_{CO}\times V_A$) is the key index that is interpreted for gas transfer, with its pathophysiological importance previously reviewed [164, 165].

Interpreting a reduced D_{LCO} must be done with these concepts in mind. The normal range for D_{LCO} and V_A should be based on the 5th percentile and 95th percentile [6, 11]. In the setting of a normal V_A , K_{CO} also has 5th and 95th percentile values. However, because K_{CO} will rise in a non-linear fashion as lung volumes fall (smaller lung gas volumes mean more rapid carbon monoxide concentration changes due to an increasingly higher surface area/volume ratio), this “normal” range for K_{CO} progressively loses meaning as lung volumes decrease. This is why in the setting of low V_A , a so-called “normal” K_{CO} (often expressed as D_{LCO}/V_A) cannot “correct” for low lung volumes [154]. Defining an impaired K_{CO} in the setting of a low V_A has minimal evidence to inform interpreters and, in practice, becomes an empirical exercise often focusing on the observed K_{CO} percent predicted [166]. Figure 11 depicts a reasonable interpretation algorithm using D_{LCO} along with K_{CO} and V_A .

It is also useful to compare V_A to TLC measured by body plethysmography to determine whether test gas maldistribution may contribute to lowering the D_{LCO} (*i.e.* carbon monoxide uptake can only be determined for the regions in which the test gases distribute). The normal value for the ratio of V_A /TLC in adults is ~0.85–0.90 [166]. Values significantly below this suggest that gas mixing impairments are likely contributing to a low measured D_{LCO} . In the absence of plethysmographic lung volume data, the presence of a steep downward slope to the inert gas tracing during exhalation suggests the possibility of gas maldistribution. There are no ideal ways to adjust for these conditions and the interpreter can only note that the problem exists [167, 168].

The future of pulmonary function interpretation

Normal results from routine PFTs do not exclude physiological impairment, especially in mild disease and in children. Specialised PFTs, when used together with routine PFTs, may provide a more comprehensive

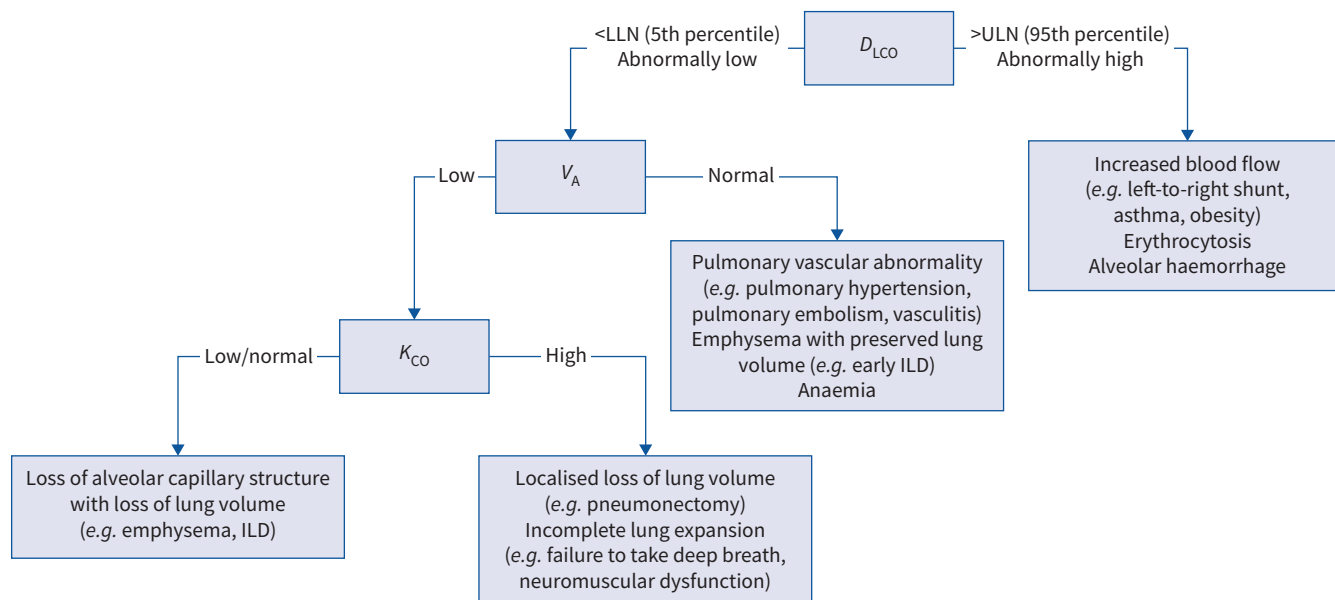


FIGURE 11 Approach to interpretation of diffusing capacity of the lung for carbon monoxide (D_{LCO}). First determine if D_{LCO} is low or high based on the lower and upper bounds defined by the 5th and 95th percentiles of the reference values. A high D_{LCO} is almost always due to increased pulmonary blood volume, as in a left-to-right shunt, increased haemoglobin, or free haemoglobin in any component of the airway, as in alveolar haemorrhage [180]. To further understand the cause of a low D_{LCO} , next examine its components, alveolar volume (V_A) and transfer coefficient of the lung for carbon monoxide (K_{CO}). If V_A is normal, then this is consistent with pulmonary vascular impairment, emphysema with preserved lung volume or anaemia. If V_A is low and K_{CO} is low or normal, then there is typically loss of alveolar capillary structure such as in emphysema or interstitial lung disease (ILD) with loss of lung volume. If V_A is low and K_{CO} is high, then there is a low lung volume state, either due to localised loss of lung volume, such as from lung resection, which may raise K_{CO} somewhat, or incomplete lung expansion, such as failure to fully inspire, which can increase K_{CO} substantially.

and multidimensional evaluation of lung function and may further improve interpretation. There is also rapid development of wearable devices that allow continuous monitoring of ventilatory indices during daily life (*i.e.* under natural physiological conditions) [169]. Together with applications that capture and interpret data, and integrated enterprise and cloud data repositories, wearable devices will provide novel solutions for personalised respiratory medicine, including tele-monitoring of respiratory function.

In the era of precision medicine and novel prediction tools, more sophisticated diagnostic models should be developed to more accurately identify early determinants of reductions in lung function. Longitudinal data across the life course are essential to identify opportunities for early intervention. There is exciting research in this field that will likely provide significant improvements, especially around the uncertainty of measurements. There are ongoing efforts devoted to the development of AIML approaches to both novel tests as well as currently standard tests. The updated interpretation standards may inform future AIML algorithms and ensure uncertainty is considered in the algorithm. Examples of uses in standard tests include AI analysis of the expiratory flow–volume pattern as noted earlier, along with measurements of inert gas washout and the carbon monoxide measurements through the D_{LCO} exhalation manoeuvre [170]. AIML-based software may also provide more accurate and standardised interpretations, and may serve as a powerful decision support tool to improve clinical practice [171, 172]. AIML may help to develop personalised, unbiased prediction of normal lung function. AIML may enhance the analysis of lung function data by identifying complex, multidimensional patterns associated with disease subtypes. While such algorithms may help to reduce any bias from poor quality data [172], AIML must use only good quality data in training to avoid introducing bias into any algorithms.

The widespread use of electronic health records [173] for data collected during the course of routine clinical practice and large clinical databases from multicentre randomised controlled trials offer unique data sources for training AIML algorithms. These algorithms may be combined with natural language processing, a set of methods which apply linguistics and ML to large corpora of clinical textual passages in order to extract structured information at a large scale. Using linguistics and computer science to process and understand text written in natural language has the potential to extract relevant information on a large

scale. Sharing and using individual data requires a robust and appropriate internationally recognised ethical, legal and information governance framework, which has yet to be established.

Conclusions

When interpreting PFT results, a clinician must interpret a particular result as within or outside the normal range for an individual of that age, sex, height and ethnic background based on reference equations, and consider how measures of lung function change over time. Interpretation of PFTs must take into account a level of uncertainty relating to 1) how representative the obtained result was of the individual's lung function at the time of testing, 2) how pre-test probability of disease may influence what is the appropriate threshold for each individual and 3) how valid for the individual is the reference population against which the test is being judged.

The requirements for obtaining a technically acceptable measurement have already been set out [4–7]. The quality of individual effort must therefore also be considered when assessing how representative the obtained result is of the individual's lung function. A poor quality result might be sufficient to answer a particular clinical question, such as if there is sufficient function to perform a lobectomy. However, a poor quality result should ideally be repeated before important decisions are made from the result. Some lung function indices are inherently more reproducible over time, such as FEV₁, FEV₆ and FVC, and will lead to more certainty in decision making than less reproducible tests.

There is clearly a level of uncertainty about the best choice of reference equations that considers an individual's sex, geographic and ancestral background. The GLI equations are the most generalisable suite of equations to date. Nonetheless, it remains unclear how to apply such a reference equation without introducing the possibility of bias. Clinicians must always take this increased uncertainty into consideration when making diagnoses and treatment recommendations.

It may also be reasonable to set clinical decision-making thresholds for a test based on clinical risk and observed clinical outcomes. A more comprehensive approach to interpretation (not simply relying on whether results are within or outside the normal range) is imperative for appropriate interpretation of lung function when pre-screening for employment, tracking the effects of exposure, disability assessment and risk assessment for therapies potentially toxic to the lungs. To date, no satisfactory outcome-based thresholds for lung function have been defined; therefore, careful consideration of the medical and exposure history of an individual is necessary when interpreting lung function results.

Importantly, clinicians should take time to explain PFT results to individuals and how these are used to guide decisions. A recent survey of people living with respiratory conditions found that more than half (59.4%) did not know what FEV₁ meant or what it represented for their condition [174]. People living with respiratory conditions, as well as those referred for PFTs, may want to know what their results mean for them.

Translation of these recommendations to clinical practice will require a paradigm shift whereby the idea of an absolute level of ideal lung function (*i.e.* the predicted value) is replaced in favour of a range of values that are observed in the majority of individuals without respiratory disease (*i.e.* z-scores or percentiles). Graphical displays as part of a PFT report can be helpful in communicating results. Interpretation of results should consider the inherent biological variability of the tests and the uncertainty of the test result. We anticipate that these interpretation recommendations will be considered in future disease-specific guidelines.

This document was endorsed by the ERS Executive Committee on 6 December 2021, by the American Thoracic Society Board of Directors on 8 December 2021, and by the Thoracic Society of Australia and New Zealand on 10 December 2021.

Conflict of interest: S. Stanojevic reports grants, support for travel from the European Respiratory Society and was the Co-Chair of the Global Lung Function Initiative for the European Respiratory Society; and additionally acted on the Pulmonary Function Testing Proficiency committee for the American Thoracic Society, all outside the submitted work. D.A. Kaminsky reports speaker fees from MGC Diagnostics, Inc. and honoraria from UpToDate, Inc., outside the submitted work. M.R. Miller has nothing to disclose. B. Thompson reports grants from the NHMRC; consulting fees from 4D Medical and Chiesi; lecture honoraria from 4D Medical, Chiesi and Mundipharma; and reports past academic work with the Global Lung Function Initiative for the European Respiratory Society, outside the submitted work. A. Aliverti reports patents on Forced Oscillation Technique from Philips and patents

on Opto-Electronic Plethysmography from BTS Bioengineering, outside the submitted work. I. Barjaktarevic has nothing to disclose. B.G. Cooper acted as Co-Chair of the Global Lung Function Initiative at the European Respiratory Society, outside the submitted work. B. Culver has nothing to disclose. E. Derom has nothing to disclose. G.L. Hall was former Co-Chair of the Global Lung Function Initiative and reports previous academic and leadership work with the Global Lung Function Initiative, both with the European Respiratory Society, outside the submitted work. T.S. Hallstrand reports research grants from the NIH (NHLBI, NIAID), outside the submitted work. J.D. Leuppi reports grants from the Swiss National Science Foundation (SNF 160072 and 185592) and Swiss Personalised Health Network (SPHN 2018DR108); and has also received unrestricted grants from AstraZeneca AG Switzerland, Boehringer Ingelheim GmbH Switzerland, GSK AG Switzerland and Novartis AG Switzerland, outside the submitted work. N. MacIntyre reports consulting fees from Vyaire, outside the submitted work. M. McCormack reports royalties for authorship for PFT chapters from UpToDate; consulting fees related to PFT quality and reading from Ardis, outside the submitted work. M. Rosenfeld has nothing to disclose. E.R. Swenson has nothing to disclose.

References

- 1 Miller MR, Crapo R, Hankinson J, *et al.* General considerations for lung function testing. *Eur Respir J* 2005; 26: 153–162.
- 2 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.
- 3 Pellegrino R, Viegi G, Brusasco V, *et al.* Interpretative strategies for lung function tests. *Eur Respir J* 2005; 26: 948–968.
- 4 Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.
- 5 Graham BL, Steenbruggen I, Miller MR, *et al.* Standardization of spirometry 2019 update. An official American Thoracic Society and European Respiratory Society technical statement. *Am J Respir Crit Care Med* 2019; 200: e70–e88.
- 6 Graham BL, Brusasco V, Burgos F, *et al.* 2017 ERS/ATS standards for single-breath carbon monoxide uptake in the lung. *Eur Respir J* 2017; 49: 1600016.
- 7 Graham BL, Brusasco V, Burgos F, *et al.* Executive Summary: 2017 ERS/ATS standards for single-breath carbon monoxide uptake in the lung. *Eur Respir J* 2017; 49: 16E0016.
- 8 Wanger J, Clausen JL, Coates A, *et al.* Standardisation of the measurement of lung volumes. *Eur Respir J* 2005; 26: 511–522.
- 9 Culver BH, Graham BL, Coates AL, *et al.* Recommendations for a standardized pulmonary function report. An official American Thoracic Society technical statement. *Am J Respir Crit Care Med* 2017; 196: 1463–1472.
- 10 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.
- 11 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.
- 12 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.
- 13 Pomeroy E, Stock JT, Wells JCK. Population history and ecology, in addition to climate, influence human stature and body proportions. *Sci Rep* 2021; 11: 274.
- 14 Goldsmith TH. Photoreceptor processes: some problems and perspectives. *J Exp Zool* 1975; 194: 89–101.
- 15 Quanjer PH, Stanojevic S, Stocks J, *et al.* Changes in the FEV₁/FVC ratio during childhood and adolescence: an intercontinental study. *Eur Respir J* 2010; 36: 1391–1399.
- 16 Fain SB, Altes TA, Panth SR, *et al.* Detection of age-dependent changes in healthy adult lungs with diffusion-weighted ³He MRI. *Acad Radiol* 2005; 12: 1385–1393.
- 17 Meiners S, Eickelberg O, Konigshoff M. Hallmarks of the ageing lung. *Eur Respir J* 2015; 45: 807–827.
- 18 LoMauro A, Aliverti A. Sex differences in respiratory function. *Breathe* 2018; 14: 131–140.
- 19 Bellemare F, Jeanneret A, Couture J. Sex differences in thoracic dimensions and configuration. *Am J Respir Crit Care Med* 2003; 168: 305–312.
- 20 Seaborn T, Simard M, Provost PR, *et al.* Sex hormone metabolism in lung development and maturation. *Trends Endocrinol Metab* 2010; 21: 729–738.
- 21 Townsend EA, Miller VM, Prakash YS. Sex differences and sex steroids in lung health and disease. *Endocr Rev* 2012; 33: 1–47.
- 22 Foer D, Rubins D, Almazan A, *et al.* Gender reference use in spirometry for transgender patients. *Ann Am Thorac Soc* 2021; 18: 537–540.
- 23 Haynes JM, Stumbo RW. The impact of using non-birth sex on the interpretation of spirometry data in subjects with air-flow obstruction. *Respir Care* 2018; 63: 215–218.
- 24 Collen J, Greenburg D, Holley A, *et al.* Racial discordance in spirometry comparing four commonly used reference equations to the National Health and Nutrition Examination Study III. *Respir Med* 2010; 104: 705–711.

- 25 Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999; 159: 179–187.
- 26 Burton RF, Nevill AM, Stewart AD, et al. Statistical approaches to relationships between sitting height and leg length in adults. *Ann Hum Biol* 2013; 40: 64–69.
- 27 Ioachimescu OC, Ramos JA, Hoffman M, et al. Assessing bronchodilator response by changes in per cent predicted forced expiratory volume in one second. *J Investig Med* 2021; 69: 1027–1034.
- 28 Whittaker AL, Sutton AJ, Beardsmore CS. Are ethnic differences in lung function explained by chest size? *Arch Dis Child Fetal Neonatal Ed* 2005; 90: F423–F428.
- 29 Quanjer PH, Kubota M, Kobayashi H, et al. Secular changes in relative leg length confound height-based spirometric reference values. *Chest* 2015; 147: 792–797.
- 30 Kiefer EM, Hankinson JL, Barr RG. Similar relation of age and height to lung function among Whites, African Americans, and Hispanics. *Am J Epidemiol* 2011; 173: 376–387.
- 31 Huprikar NA, Holley AB, Skabelund AJ, et al. A comparison of Global Lung Initiative 2012 with Third National Health and Nutrition Examination Survey spirometry reference values. Implications in defining obstruction. *Ann Am Thorac Soc* 2019; 16: 225–230.
- 32 Miller MR, Thinggaard M, Christensen K, et al. Best lung function equations for the very elderly selected by survival analysis. *Eur Respir J* 2014; 43: 1338–1346.
- 33 Linares-Perdomo O, Hegewald M, Collingridge DS, et al. Comparison of NHANES III and ERS/GLI 12 for airway obstruction classification and severity. *Eur Respir J* 2016; 48: 133–141.
- 34 Quanjer PH, Brazzale DJ, Boros PW, et al. Implications of adopting the Global Lungs Initiative 2012 all-age reference equations for spirometry. *Eur Respir J* 2013; 42: 1046–1054.
- 35 Hulo S, de Broucker V, Giovannelli J, et al. Global Lung Function Initiative reference equations better describe a middle-aged, healthy French population than the European Community for Steel and Coal values. *Eur Respir J* 2016; 48: 1779–1781.
- 36 Langhammer A, Johannessen A, Holmen TL, et al. Global Lung Function Initiative 2012 reference equations for spirometry in the Norwegian population. *Eur Respir J* 2016; 48: 1602–1611.
- 37 Jones RL, Nzekwu MM. The effects of body mass index on lung volumes. *Chest* 2006; 130: 827–833.
- 38 Littleton SW, Tulaimat A. The effects of obesity on lung volumes and oxygenation. *Respir Med* 2017; 124: 15–20.
- 39 Winck AD, Heinzmann-Filho JP, Soares RB, et al. Effects of obesity on lung volume and capacity in children and adolescents: a systematic review. *Rev Paul Pediatr* 2016; 34: 510–517.
- 40 Saliman JA, Benditt JO, Flum DR, et al. Pulmonary function in the morbidly obese. *Surg Obes Relat Dis* 2008; 4: 632–639.
- 41 Tan EK, Tan EL. Alterations in physiology and anatomy during pregnancy. *Best Pract Res Clin Obstet Gynaecol* 2013; 27: 791–802.
- 42 McAuliffe F, Kametas N, Costello J, et al. Respiratory function in singleton and twin pregnancy. *BJOG* 2002; 109: 765–769.
- 43 Kirkby J, Aurora P, Spencer H, et al. Stitching and switching: the impact of discontinuous lung function reference equations. *Eur Respir J* 2012; 39: 1256–1257.
- 44 Rosenfeld M, Pepe MS, Longton G, et al. Effect of choice of reference equation on analysis of pulmonary function in cystic fibrosis patients. *Pediatr Pulmonol* 2001; 31: 227–237.
- 45 Subbarao P, Lebecque P, Corey M, et al. Comparison of spirometric reference values. *Pediatr Pulmonol* 2004; 37: 515–522.
- 46 Quanjer PH, Stanojevic S. Do the Global Lung Function Initiative 2012 equations fit my population? *Eur Respir J* 2016; 48: 1782–1785.
- 47 National High Blood Pressure Education Program Working Group on High Blood Pressure in Children and Adolescents. The fourth report on the diagnosis, evaluation, and treatment of high blood pressure in children and adolescents. *Pediatrics* 2004; 114: 555–576.
- 48 Kuczumarski R, Ogden CL, Guo S, et al. CDC Growth Charts. Hyattsville, National Center for Health Statistics, 2000.
- 49 International Federation of Clinical Chemistry, Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values (EPTRV). The theory of reference values. Part 6. Presentation of observed values related to reference values. *Clin Chim Acta* 1983; 127: 441F–448F.
- 50 Neder JA, Berton DC, O'Donnell DE. The lung function laboratory to assist clinical decision-making in pulmonology: evolving challenges to an old issue. *Chest* 2020; 158: 1629–1643.
- 51 Swanney MP, Ruppel G, Enright PL, et al. Using the lower limit of normal for the FEV₁/FVC ratio reduces the misclassification of airway obstruction. *Thorax* 2008; 63: 1046–1051.
- 52 Miller MR, Quanjer PH, Swanney MP, et al. Interpreting lung function data using 80% predicted and fixed thresholds misclassifies more than 20% of patients. *Chest* 2011; 139: 52–59.
- 53 Ali SS, Elliott WH. Bile acids. XLVII. 12 α -hydroxylation of precursors of allo bile acids by rabbit liver microsomes. *Biochim Biophys Acta* 1975; 409: 249–257.

- 54 Agusti A, Noell G, Brugada J, *et al.* Lung function in early adulthood and health in later life: a transgenerational cohort analysis. *Lancet Respir Med* 2017; 5: 935–945.
- 55 Karmaus W, Mukherjee N, Janjanam VD, *et al.* Distinctive lung function trajectories from age 10 to 26 years in men and women and associated early life risk factors – a birth cohort study. *Respir Res* 2019; 20: 98.
- 56 Fletcher C, Peto R. The natural history of chronic airflow obstruction. *Br Med J* 1977; 1: 1645–1648.
- 57 Brand PL, Quanjer PH, Postma DS, *et al.* Interpretation of bronchodilator response in patients with obstructive airways disease. The Dutch Chronic Non-Specific Lung Disease (CNSLD) Study Group. *Thorax* 1992; 47: 429–436.
- 58 Chhabra SK. Acute bronchodilator response has limited value in differentiating bronchial asthma from COPD. *J Asthma* 2005; 42: 367–372.
- 59 Quanjer PH, Ruppel GL, Langhammer A, *et al.* Bronchodilator response in FVC is larger and more relevant than in FEV₁ in severe airflow obstruction. *Chest* 2017; 151: 1088–1098.
- 60 Barjaktarevic I, Kaner R, Buhr RG, *et al.* Bronchodilator responsiveness or reversibility in asthma and COPD – a need for clarity. *Int J Chron Obstruct Pulmon Dis* 2018; 13: 3511–3513.
- 61 Barjaktarevic IZ, Buhr RG, Wang X, *et al.* Clinical significance of bronchodilator responsiveness evaluated by forced vital capacity in COPD: SPIROMICS cohort analysis. *Int J Chron Obstruct Pulmon Dis* 2019; 14: 2927–2938.
- 62 Tan WC, Vollmer WM, Lamprecht B, *et al.* Worldwide patterns of bronchodilator responsiveness: results from the Burden of Obstructive Lung Disease study. *Thorax* 2012; 67: 718–726.
- 63 Ward H, Cooper BG, Miller MR. Improved criterion for assessing lung function reversibility. *Chest* 2015; 148: 877–886.
- 64 Koga T, Kamimura T, Oshita Y, *et al.* Determinants of bronchodilator responsiveness in patients with controlled asthma. *J Asthma* 2006; 43: 71–74.
- 65 Burity EF, Pereira CA, Jones MH, *et al.* Bronchodilator response cut-off points and FEV_{0.75} reference values for spirometry in preschoolers. *J Bras Pneumol* 2016; 42: 326–332.
- 66 Chen C, Jian W, Gao Y, *et al.* Early COPD patients with lung hyperinflation associated with poorer lung function but better bronchodilator responsiveness. *Int J Chron Obstruct Pulmon Dis* 2016; 11: 2519–2526.
- 67 Han MK, Wise R, Mumford J, *et al.* Prevalence and clinical correlates of bronchoreversibility in severe emphysema. *Eur Respir J* 2010; 35: 1048–1056.
- 68 Lee JS, Huh JW, Chae EJ, *et al.* Response patterns to bronchodilator and quantitative computed tomography in chronic obstructive pulmonary disease. *Clin Physiol Funct Imaging* 2012; 32: 12–18.
- 69 Rodriguez-Carballeira M, Heredia JL, Rue M, *et al.* The bronchodilator test in chronic obstructive pulmonary disease: interpretation methods. *Respir Med* 2007; 101: 34–42.
- 70 Walker PP, Calverley PM. The volumetric response to bronchodilators in stable chronic obstructive pulmonary disease. *COPD* 2008; 5: 147–152.
- 71 Ferrer Galvan M, Alvarez Gutierrez FJ, Romero Falcon A, *et al.* Is the bronchodilator test an useful tool to measure asthma control? *Respir Med* 2017; 126: 26–31.
- 72 Hanania NA, Sharafkhaneh A, Celli B, *et al.* Acute bronchodilator responsiveness and health outcomes in COPD patients in the UPLIFT trial. *Respir Res* 2011; 12: 6.
- 73 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.
- 74 Lange P, Celli B, Agusti A, *et al.* Lung-function trajectories leading to chronic obstructive pulmonary disease. *N Engl J Med* 2015; 373: 111–122.
- 75 Moore OA, Proudman SM, Goh N, *et al.* Quantifying change in pulmonary function as a prognostic marker in systemic sclerosis-related interstitial lung disease. *Clin Exp Rheumatol* 2015; 33: S111–S116.
- 76 Wang ML, Avashia BH, Petsonk EL. Interpreting periodic lung function tests in individuals: the relationship between 1- to 5-year and long-term FEV₁ changes. *Chest* 2006; 130: 493–499.
- 77 Stanojevic S, Filipow N, Ratjen F. Paediatric reproducibility limits for the forced expiratory volume in 1 s. *Thorax* 2020; 75: 891–896.
- 78 Donohue JF. Minimal clinically important differences in COPD lung function. *COPD* 2005; 2: 111–124.
- 79 Oelsner EC, Balte PP, Bhatt SP, *et al.* Lung function decline in former smokers and low-intensity current smokers: a secondary data analysis of the NHLBI Pooled Cohorts Study. *Lancet Respir Med* 2020; 8: 34–44.
- 80 Redlich CA, Tarlo SM, Hankinson JL, *et al.* Official American Thoracic Society technical standards: spirometry in the occupational setting. *Am J Respir Crit Care Med* 2014; 189: 983–993.
- 81 Miller MR, Pedersen OF. New concepts for expressing forced expiratory volume in 1 s arising from survival analysis. *Eur Respir J* 2010; 35: 873–882.
- 82 Jones PW, Beeh KM, Chapman KR, *et al.* Minimal clinically important differences in pharmacological trials. *Am J Respir Crit Care Med* 2014; 189: 250–255.
- 83 Kannel WB, Lew EA, Hubert HB, *et al.* The value of measuring vital capacity for prognostic purposes. *Trans Assoc Life Insur Med Dir Am* 1980; 64: 66–83.

- 84 Peto R, Speizer FE, Cochrane AL, *et al.* The relevance in adults of air-flow obstruction, but not of mucus hypersecretion, to mortality from chronic lung disease. Results from 20 years of prospective observation. *Am Rev Respir Dis* 1983; 128: 491–500.
- 85 Ferguson MK, Little L, Rizzo L, *et al.* Diffusing capacity predicts morbidity and mortality after pulmonary resection. *J Thorac Cardiovasc Surg* 1988; 96: 894–900.
- 86 Kannel WB, Hubert H, Lew EA. Vital capacity as a predictor of cardiovascular disease: the Framingham study. *Am Heart J* 1983; 105: 311–315.
- 87 Neas LM, Schwartz J. Pulmonary function levels as predictors of mortality in a national sample of US adults. *Am J Epidemiol* 1998; 147: 1011–1018.
- 88 Miller MR, Cooper BG. Reduction in T_{LCO} and survival in a clinical population. *Eur Respir J* 2021; 58: 2002046.
- 89 Stanojevic S, Wade A, Stocks J, *et al.* Reference ranges for spirometry across all ages: a new approach. *Am J Respir Crit Care Med* 2008; 177: 253–260.
- 90 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.
- 91 Jones PW, Quirk FH, Baveystock CM, *et al.* A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. *Am Rev Respir Dis* 1992; 145: 1321–1327.
- 92 Jones PW, Harding G, Berry P, *et al.* Development and first validation of the COPD Assessment Test. *Eur Respir J* 2009; 34: 648–654.
- 93 Trinick R, Southern KW, McNamara PS. Assessing the Liverpool Respiratory Symptom Questionnaire in children with cystic fibrosis. *Eur Respir J* 2012; 39: 899–905.
- 94 Cassidy RN, Roberts ME, Colby SM. Validation of a respiratory symptom questionnaire in adolescent smokers. *Tob Regul Sci* 2015; 1: 121–128.
- 95 Lee SW, Kim HK, Baek S, *et al.* Development of a spirometry T-score in the general population. *Int J Chron Obstruct Pulmon Dis* 2016; 11: 369–379.
- 96 Pedone C, Scarlata S, Scichilone N, *et al.* Alternative ways of expressing FEV₁ and mortality in elderly people with and without COPD. *Eur Respir J* 2013; 41: 800–805.
- 97 Huang TH, Hsiue TR, Lin SH, *et al.* Comparison of different staging methods for COPD in predicting outcomes. *Eur Respir J* 2018; 51: 1700577.
- 98 Miller MR, Pedersen OF, Dirksen A. A new staging strategy for chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis* 2007; 2: 657–663.
- 99 Hegendorfer E, Vaes B, Andreeva E, *et al.* Predictive value of different expressions of forced expiratory volume in 1 second (FEV₁) for adverse outcomes in a cohort of adults aged 80 and older. *J Am Med Dir Assoc* 2017; 18: 123–130.
- 100 Bates DV. *Respiratory Function in Disease*. Philadelphia, WB Saunders, 1989.
- 101 Wilson A. *Pulmonary Function Testing, Indications and Interpretations*. Orlando, Grune & Stratton, 1985.
- 102 Pride NB, Macklem PT. Lung mechanics in disease. *In*: Macklem PT, Mead J, eds. *Handbook of Physiology. The Respiratory System: Mechanics of Breathing*. Bethesda, American Physiological Society, 1986; pp. 659–692.
- 103 American Thoracic Society. Lung function testing: selection of reference values and interpretative strategies. *Am Rev Respir Dis* 1991; 144: 1202–1218.
- 104 Singh D, Agusti A, Anzueto A, *et al.* Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease: the GOLD science committee report 2019. *Eur Respir J* 2019; 53: 1900164.
- 105 Qaseem A, Wilt TJ, Weinberger SE, *et al.* Diagnosis and management of stable chronic obstructive pulmonary disease: a clinical practice guideline update from the American College of Physicians, American College of Chest Physicians, American Thoracic Society, and European Respiratory Society. *Ann Intern Med* 2011; 155: 179–191.
- 106 McDonough JE, Yuan R, Suzuki M, *et al.* Small-airway obstruction and emphysema in chronic obstructive pulmonary disease. *N Engl J Med* 2011; 365: 1567–1575.
- 107 Macklem PT, Mead J. Resistance of central and peripheral airways measured by a retrograde catheter. *J Appl Physiol* 1967; 22: 395–401.
- 108 Pride NB, Macklem PT. *Lung Mechanics in Disease*. Bethesda, American Physiological Society, 1986.
- 109 Bhatt SP, Bhakta NR, Wilson CG, *et al.* New spirometry indices for detecting mild airflow obstruction. *Sci Rep* 2018; 8: 17484.
- 110 Flenley DC. Chronic obstructive pulmonary disease. *Dis Mon* 1988; 34: 537–599.
- 111 Quanjer PH, Stanojevic S, Thompson BR. Spirometric thresholds and biased interpretation of test results. *Thorax* 2014; 69: 1146.
- 112 Pellegrino R, Brusasco V, Miller MR. Question everything. *Eur Respir J* 2014; 43: 947–948.
- 113 Morris ZQ, Coz A, Starosta D. An isolated reduction of the FEV₃/FVC ratio is an indicator of mild lung injury. *Chest* 2013; 144: 1117–1123.
- 114 Dilektasli AG, Porszasz J, Casaburi R, *et al.* A novel spirometric measure identifies mild COPD unidentified by standard criteria. *Chest* 2016; 150: 1080–1090.

- 115 Zimmermann SC, Tonga KO, Thamrin C. Dismantling airway disease with the use of new pulmonary function indices. *Eur Respir Rev* 2019; 28: 180122.
- 116 Hyatt RE. Forced exhalation. In: Macklem PT, Mead J, eds. *Handbook of Physiology. The Respiratory System: Mechanics of Breathing*. Bethesda, American Physiological Society, 1986; pp. 295–314.
- 117 Forno E, Weiner DJ, Mullen J, et al. Obesity and airway dysanapsis in children with and without asthma. *Am J Respir Crit Care Med* 2017; 195: 314–323.
- 118 Thompson BR. Dysanapsis – once believed to be a physiological curiosity – is now clinically important. *Am J Respir Crit Care Med* 2017; 195: 277–278.
- 119 Dos Santos Andreato L, Soares MR, Pereira CA. Reduced FEV₁/FVC and FEV₁ in the normal range as a physiological variant. *Respir Care* 2019; 64: 570–575.
- 120 Iyer VN, Schroeder DR, Parker KO, et al. The nonspecific pulmonary function test: longitudinal follow-up and outcomes. *Chest* 2011; 139: 878–886.
- 121 Hyatt RE, Cowl CT, Bjoraker JA, et al. Conditions associated with an abnormal nonspecific pattern of pulmonary function tests. *Chest* 2009; 135: 419–424.
- 122 Wan ES, Fortis S, Regan EA, et al. Longitudinal phenotypes and mortality in preserved ratio impaired spirometry in the COPDGene study. *Am J Respir Crit Care Med* 2018; 198: 1397–1405.
- 123 Fortis S, Comellas A, Kim V, et al. Low FVC/TLC in preserved ratio impaired spirometry (PRISm) is associated with features of and progression to obstructive lung disease. *Sci Rep* 2020; 10: 5169.
- 124 Marott JL, Ingebrigtsen TS, Colak Y, et al. Trajectory of preserved ratio impaired spirometry: natural history and long-term prognosis. *Am J Respir Crit Care Med* 2021; 204: 910–920.
- 125 Punjabi NM, Shade D, Wise RA. Correction of single-breath helium lung volumes in patients with airflow obstruction. *Chest* 1998; 114: 907–918.
- 126 Chan ED, Irvin CG. The detection of collapsible airways contributing to airflow limitation. *Chest* 1995; 107: 856–859.
- 127 Saint-Pierre M, Ladha J, Berton DC, et al. Is the slow vital capacity clinically useful to uncover airflow limitation in subjects with preserved FEV₁/FVC ratio? *Chest* 2019; 156: 497–506.
- 128 Brusasco V, Pellegrino R, Rodarte JR. Vital capacities in acute and chronic airway obstruction: dependence on flow and volume histories. *Eur Respir J* 1997; 10: 1316–1320.
- 129 Akpınar-Elci M, Fedan KB, Enright PL. FEV₆ as a surrogate for FVC in detecting airways obstruction and restriction in the workplace. *Eur Respir J* 2006; 27: 374–377.
- 130 Gleeson S, Mitchell B, Pasquarella C, et al. Comparison of FEV₆ and FVC for detection of airway obstruction in a community hospital pulmonary function laboratory. *Respir Med* 2006; 100: 1397–1401.
- 131 Vandevoorde J, Verbanck S, Schuermans D, et al. FEV₁/FEV₆ and FEV₆ as an alternative for FEV₁/FVC and FVC in the spirometric detection of airway obstruction and restriction. *Chest* 2005; 127: 1560–1564.
- 132 Vandevoorde J, Verbanck S, Schuermans D, et al. Obstructive and restrictive spirometric patterns: fixed cut-offs for FEV₁/FEV₆ and FEV₆. *Eur Respir J* 2006; 27: 378–383.
- 133 Vandevoorde J, Verbanck S, Schuermans D, et al. The role of FEV₆ in the detection of airway obstruction. *Respir Med* 2005; 99: 1465–1466.
- 134 Thompson EB, Anderson CU, Lippman ME. Serum-free growth of HTC cells containing glucocorticoid- and insulin-inducible tyrosine aminotransferase and cytoplasmic glucocorticoid receptors. *J Cell Physiol* 1975; 86: 403–411.
- 135 Ioachimescu OC, Venkateshiah SB, Kavuru MS, et al. Estimating FVC from FEV₂ and FEV₃: assessment of a surrogate spirometric parameter. *Chest* 2005; 128: 1274–1281.
- 136 Guenette JA, Chin RC, Cory JM, et al. Inspiratory capacity during exercise: measurement, analysis, and interpretation. *Pulm Med* 2013; 2013: 956081.
- 137 Hoesterey D, Das N, Janssens W, et al. Spirometric indices of early airflow impairment in individuals at risk of developing COPD: spirometry beyond FEV₁/FVC. *Respir Med* 2019; 156: 58–68.
- 138 Bodduluri S, Nakhmani A, Reinhardt JM, et al. Deep neural network analyses of spirometry for structural phenotyping of chronic obstructive pulmonary disease. *JCI Insight* 2020; 5: e132781.
- 139 Kaminsky DA. What does airway resistance tell us about lung function? *Respir Care* 2012; 57: 85–96.
- 140 Beydon N, Davis SD, Lombardi E, et al. An official American Thoracic Society/European Respiratory Society statement: pulmonary function testing in preschool children. *Am J Respir Crit Care Med* 2007; 175: 1304–1345.
- 141 Rosenfeld M, Allen J, Arets BH, et al. An official American Thoracic Society workshop report: optimal lung function tests for monitoring cystic fibrosis, bronchopulmonary dysplasia, and recurrent wheezing in children less than 6 years of age. *Ann Am Thorac Soc* 2013; 10: S1–S11.
- 142 Peterson-Carmichael S, Seddon PC, Cheifetz IM, et al. An official American Thoracic Society/European Respiratory Society workshop report: evaluation of respiratory mechanics and function in the pediatric and neonatal intensive care units. *Ann Am Thorac Soc* 2016; 13: S1–S11.
- 143 Empey DW. Assessment of upper airways obstruction. *Br Med J* 1972; 3: 503–505.

- 144 Miller MR, Pincock AC, Oates GD, *et al.* Upper airway obstruction due to goitre: detection, prevalence and results of surgical management. *Q J Med* 1990; 74: 177–188.
- 145 Modrykamien AM, Gudavalli R, McCarthy K, *et al.* Detection of upper airway obstruction with spirometry results and the flow-volume loop: a comparison of quantitative and visual inspection criteria. *Respir Care* 2009; 54: 474–479.
- 146 Arjomandi M, Zeng S, Barjaktarevic I, *et al.* Radiographic lung volumes predict progression to COPD in smokers with preserved spirometry in SPIROMICS. *Eur Respir J* 2019; 54: 1802214.
- 147 Arjomandi M, Zeng S, Geerts J, *et al.* Lung volumes identify an at-risk group in persons with prolonged secondhand tobacco smoke exposure but without overt airflow obstruction. *BMJ Open Respir Res* 2018; 5: e000284.
- 148 Sood N, Turcotte SE, Wasilewski NV, *et al.* Small-airway obstruction, dynamic hyperinflation, and gas trapping despite normal airway sensitivity to methacholine in adults with chronic cough. *J Appl Physiol* 2019; 126: 294–304.
- 149 Vaz Fragoso CA, Cain HC, Casaburi R, *et al.* Spirometry, static lung volumes, and diffusing capacity. *Respir Care* 2017; 62: 1137–1147.
- 150 O'Donnell DE, Elbehairy AF, Webb KA, *et al.* The link between reduced inspiratory capacity and exercise intolerance in chronic obstructive pulmonary disease. *Ann Am Thorac Soc* 2017; 14: S30–S39.
- 151 Chandrasoma B, Balfe D, Naik T, *et al.* Pulmonary function in patients with amyotrophic lateral sclerosis at disease onset. *Monaldi Arch Chest Dis* 2012; 77: 129–133.
- 152 Fauroux B, Khirani S. Neuromuscular disease and respiratory physiology in children: putting lung function into perspective. *Respirology* 2014; 19: 782–791.
- 153 Aaron SD, Dales RE, Cardinal P. How accurate is spirometry at predicting restrictive pulmonary impairment? *Chest* 1999; 115: 869–873.
- 154 Gladly CA, Aaron SD, Lunau M, *et al.* A spirometry-based algorithm to direct lung function testing in the pulmonary function laboratory. *Chest* 2003; 123: 1939–1946.
- 155 Tan CS, Tashkin DP. Supernormal maximal mid-expiratory flow rates in diffuse interstitial lung disease. *Respiration* 1981; 42: 200–208.
- 156 Clay RD, Iyer VN, Reddy DR, *et al.* The “complex restrictive” pulmonary function pattern: clinical and radiologic analysis of a common but previously undescribed restrictive pattern. *Chest* 2017; 152: 1258–1265.
- 157 Tantucci C, Bottone D, Borghesi A, *et al.* Methods for measuring lung volumes: is there a better one? *Respiration* 2016; 91: 273–280.
- 158 Ferris BG. Epidemiology standardization project (American Thoracic Society). *Am Rev Respir Dis* 1978; 118: 1–120.
- 159 Milite F, Lederer DJ, Weingarten JA, *et al.* Quantification of single-breath underestimation of lung volume in emphysema. *Respir Physiol Neurobiol* 2009; 165: 215–220.
- 160 Rodenstein DO, Stanescu DC. Reassessment of lung volume measurement by helium dilution and by body plethysmography in chronic air-flow obstruction. *Am Rev Respir Dis* 1982; 126: 1040–1044.
- 161 Stanescu DC, Rodenstein D, Cauberghs M, *et al.* Failure of body plethysmography in bronchial asthma. *J Appl Physiol Respir Environ Exerc Physiol* 1982; 52: 939–948.
- 162 Dykstra BJ, Scanlon PD, Kester MM, *et al.* Lung volumes in 4,774 patients with obstructive lung disease. *Chest* 1999; 115: 68–74.
- 163 Diaz-Guzman E, McCarthy K, Siu A, *et al.* Frequency and causes of combined obstruction and restriction identified in pulmonary function tests in adults. *Respir Care* 2010; 55: 310–316.
- 164 Hughes JM. The single breath transfer factor ($T_{L,CO}$) and the transfer coefficient (K_{CO}): a window onto the pulmonary microcirculation. *Clin Physiol Funct Imaging* 2003; 23: 63–71.
- 165 Hughes JM, Bates DV. Historical review: the carbon monoxide diffusing capacity (D_{LCO}) and its membrane (D_m) and red cell ($\Theta \cdot V_c$) components. *Respir Physiol Neurobiol* 2003; 138: 115–142.
- 166 Hughes JM, Pride NB. Examination of the carbon monoxide diffusing capacity (D_{LCO}) in relation to its K_{CO} and V_A components. *Am J Respir Crit Care Med* 2012; 186: 132–139.
- 167 Hughes JM, Pride NB. In defence of the carbon monoxide transfer coefficient $K_{CO} (T_L/V_A)$. *Eur Respir J* 2001; 17: 168–174.
- 168 Kanengiser LC, Rapoport DM, Epstein H, *et al.* Volume adjustment of mechanics and diffusion in interstitial lung disease. Lack of clinical relevance. *Chest* 1989; 96: 1036–1042.
- 169 Aliverti A. Wearable technology: role in respiratory health and disease. *Breathe* 2017; 13: e27–e36.
- 170 Huang YC, Macintyre NR. Real-time gas analysis improves the measurement of single-breath diffusing capacity. *Am Rev Respir Dis* 1992; 146: 946–950.
- 171 Topalovic M, Das N, Burgel PR, *et al.* Artificial intelligence outperforms pulmonologists in the interpretation of pulmonary function tests. *Eur Respir J* 2019; 53: 1801660.
- 172 Das N, Verstraete K, Stanojevic S, *et al.* Deep-learning algorithm helps to standardise ATS/ERS spirometric acceptability and usability criteria. *Eur Respir J* 2020; 56: 2000603.

- 173 McCormack MC, Bascom R, Brandt M, *et al.* Electronic health records and pulmonary function data: developing an interoperability roadmap. An official American Thoracic Society workshop report. *Ann Am Thorac Soc* 2021; 18: 1–11.
- 174 Johnson B, Steenbruggen I, Graham BL, *et al.* Improving spirometry testing by understanding patient preferences. *ERJ Open Res* 2020; 7: 00712–2020.
- 175 Busi LE, Sly PD. Validation of the GLI-2012 spirometry reference equations in Argentinian children. *Pediatr Pulmonol* 2018; 53: 204–208.
- 176 Hall GL, Thompson BR, Stanojevic S, *et al.* The Global Lung Initiative 2012 reference values reflect contemporary Australasian spirometry. *Respirology* 2012; 17: 1150–1151.
- 177 Smith SJ, Gray DM, MacGinty RP, *et al.* Choosing the better Global Lung Initiative 2012 equation in South African population groups. *Am J Respir Crit Care Med* 2020; 202: 1724–1727.
- 178 Sonnappa S, Lum S, Kirkby J, *et al.* Disparities in pulmonary function in healthy children across the Indian urban–rural continuum. *Am J Respir Crit Care Med* 2015; 191: 79–86.
- 179 Gascoigne AD, Corris PA, Dark JH, *et al.* The biphasic spirogram: a clue to unilateral narrowing of a mainstem bronchus. *Thorax* 1990; 45: 637–638.
- 180 Saydain G, Beck KC, Decker PA, *et al.* Clinical significance of elevated diffusing capacity. *Chest* 2004; 125: 446–452.