

CHAPTER 1

General features of interpretation and report writing

There are features in the interpretation of lung function tests and report writing that are common to most tests of lung function. This chapter explores these general features.

General features of interpretation

The general features of interpretation are (1)

- 1 assessing test validity;
- 2 assessing the adequacy of reference values for the particular subject;
- 3 determining normality or abnormality using upper and/or lower limits of normal;
- 4 classifying detected abnormalities based on known patterns of disease;
- 5 determining the severity of an abnormality;
- 6 comparing current and previous results to identify significant changes over time;
- 7 attempting to address clinical question(s) mentioned in the referral.

Assessing test validity

- Interpretation of results should begin with a review of test quality. Good test quality is important as suboptimal quality tests may impact negatively on the interpretation of results and hence on clinical decision making. Information regarding indicators of test quality is provided in the test-specific chapters and in Chapter 7.



2 Chapter 1

- The identification of suboptimal quality results can be gleaned from examination of the raw test data, technical comments provided by the test operator or a combination of both.
- When a suboptimal quality test is obtained, a cautionary statement identifying the magnitude and direction of the impact of the suboptimal quality results should be included in the report. For example:

Results should be interpreted with caution as test performance for spirometry was suboptimal due to coughing at end expiration, and may result in potential underestimation of forced vital capacity (FVC).

Assessing the adequacy of reference values for the particular subject

- Lung function results are interpreted by comparing the obtained results to a known reference range.
- The reference range/equations chosen need to reflect the population(s) tested and the test methods used in the laboratory (1).
- The reference range used for each test, as well as the limits of the variables (e.g. age, height, weight) of the reference equations, should be known to those reporting.

If reference values are extrapolated beyond the limits of the variables (for example, a subject's age is 85 years, but the age range of the reference set used is 8–80 years), then a cautionary statement should be included as there is uncertainty regarding the validity of the reference data. For example: *Reference values for spirometry have been extrapolated for age and should be used with caution.*

- Lung function may be affected by race. Clear differences between Caucasian and African-American populations in the United States have been shown (2). Ideally, the subject's race (or the race they identify within the case of mixed race) should be taken into account in selecting appropriate reference sets. There are, however, practical issues in identifying and using appropriate reference sets for multiple races, and appropriate reference sets for some tests do not exist.

The Global Lung Initiative has published a multiethnic set of spirometry reference values (3), which goes some way to addressing the issue of race in reference values. At the time of writing, the Global Lung Initiative is working towards race-specific reference values for T_LCO also.

A useful, but less than ideal solution for this problem, is the application of a race correction factor (e.g. 0.88 for FEV_1 and FVC (forced vital capacity)) to

General features of interpretation and report writing 3

Caucasian reference values when testing non-Caucasian subjects (1). This method is by no means ideal and when a correction factor is applied, a cautionary statement should be used to inform the reader that the reference values have been adjusted for race. For example: *Reference values have been adjusted for race and should be used with caution.*

Determining normality or abnormality using upper and/or lower limits of normal

The normal range:

- The normal range is defined by the range in which there is confidence for inclusion of 95% of the normal population.
- The 95% confidence limits are determined using the mean predicted value (MPV) calculated from the reference equations and the residual standard deviation (RSD) that describes the amount of scatter or variation around the MPV.
- The upper limits of normal (ULN) and lower limits of normal (LLN) can be calculated using the MPV and the RSD as follows:
 - For parameters that may have an abnormally high or low result (e.g. haemoglobin), the upper and lower 95% confidence limits are given by
 - ULN: $MPV + 1.96RSD$
 - LLN: $MPV - 1.96RSD$
 - The limits are set at the 2.5th and 97.5th percentiles (5% in total lie outside the normal range)
 - For parameters where it is possible to have only abnormally low results (e.g. FEV_1 , FVC), the lower 95% confidence limit is given by
 - LLN: $MPV - 1.64RSD$
 - The lower limit is set at the 5% percentile (5% lie below the normal range)
 - For parameters where it is possible to have only abnormally high results (e.g. RV (residual volume):TLC (total lung capacity) ratio), the upper 95% confidence limit is given by
 - ULN: $MPV + 1.64RSD$
 - The upper limit is set at the 95th percentile (5% lie above the normal range)
- A z-score expresses the number of standard deviations a measured result is from the mean and is calculated $(\text{measured value} - MPV)/RSD$. z-score values below the MPV are recorded as a negative number and values above the MPV as a positive number.



4 Chapter 1

Using the 95% confidence limits to set the upper and/or lower limits of normal.

- Parameters that may have an abnormally low or high result: an abnormal result can be identified by a z-score either less than -1.96 or greater than $+1.96$, respectively.
- Parameters with only abnormally low results: an abnormal result can be identified by a z-score less than -1.64 .
- Parameters with only abnormally high results: an abnormal result can be identified by a z-score greater than $+1.64$.

Determining normality or abnormality

- Limit the number of parameters used in the interpretation of lung function. The more parameters that are included in the test analysis, the more likelihood there is of returning an abnormal finding.
- When results are within normal limits, they should be reported as being *within normal limits* rather than being normal. There may be lung disease present that has not as yet forced any parameters of lung function outside the normal limits.
- When a result is abnormal, it is described as being reduced if it is below the LLN or elevated if it is above the upper limit of normal.
- Borderline results require careful consideration in interpretation and it is acceptable to describe a result as borderline.
- As the normal range is defined as the range in which there is confidence that 95% of the normal population will be included, 5% of the normal population will have an abnormal finding. This is a particularly important consideration when lung function is being tested in a general population in the absence of symptoms (e.g. pre-employment medicals, epidemiological surveys). In a doctor referred population dictated by specific symptoms, an abnormal finding is more likely to be a true abnormal finding.

Classifying detected abnormality based on known patterns of disease

- When an abnormality is identified, the pattern of abnormality should be identified.
- Ensure that information is used from all the tests performed to inform the overall interpretation of the result. For example, when spirometry and static lung volumes are performed, they should be used together to determine the pattern of abnormality as they both measure aspects of ventilatory function.
- Lung function is rarely, if ever, used as a diagnostic tool in isolation. Lung function results are usually incorporated into the larger clinical

General features of interpretation and report writing 5

picture (patient history, imaging, blood tests, biopsies, etc.) to assist with making a diagnosis. Suggesting a specific diagnosis based only on abnormalities of lung function is unwise as a pattern of abnormality seen in lung function results may include multiple diseases/disorders. For example:

- An obstructive ventilatory defect may be present in asthma, chronic bronchitis, emphysema, cystic fibrosis, bronchiectasis or other disorders of the airways. Differentiation cannot be made between these disorders with spirometry alone. Hence, we can only describe a pattern and not specify a diagnosis.

In the report, describe the patterns of abnormality rather than suggest specific diagnoses. For example:

- The clinical notes in a referral state, *Chronic obstructive pulmonary disease (COPD)? Extensive smoking history*. The results show an obstructive pattern with no significant response to inhaled bronchodilator. This might be reported as *There is an obstructive ventilatory defect with no significant bronchodilator response. The result is consistent with the spirometric definition of COPD.*

Determining the severity of an abnormality

- Severity scales for grading abnormal tests of spirometry and carbon monoxide transfer factor are available (Tables 1.1–1.3)(1, 4). These scales are based on arbitrary cut-offs and do not reflect functional status. For example:

An FEV_1 of 62% MPV constitutes moderate lung disease (Table 1.1). This level of lung function may impact significantly on the functional status for one person, but not for another.

- Severity scales are not recommended due to the arbitrary nature of the cut-offs. Instead, an abnormal finding should simply be called an abnormal finding. It is recommended that those who do wish to grade severity in their practice use available published scales (1, 4) for consistency (Tables 1.1–1.3). A cautionary statement should be made when using arbitrary severity scales. For example:

The severity scale used is arbitrary and is not necessarily representative of functional status.

Comparing current and previous results to identify significant changes over time

Once the first result of a subject is recorded, progress is monitored by comparing current results to previous results. That is, the subject becomes their own control.

6 Chapter 1

Table 1.1 Severity scale for any spirometric abnormality (1).

Adapted and reproduced with permission of the European Respiratory Society: *Eur Respir J* November 2005 26:948-968; doi:10.1183/09031936.05.00035205

Severity classification	FEV ₁ % MPV
Mild	>70
Moderate	60–69
Moderately severe	50–59
Severe	35–49
Very severe	<35

Table 1.2 Severity scale for obstruction on spirometry using FEV₁ z-score (4).

^aLLN, lower limit of normal.
Reproduced with permission of the European Respiratory Society: *Eur Respir J* erj00863-2013; doi:10.1183/09031936.00086313

Severity classification	FEV ₁ /(F)VC < LLN ^a and FEV ₁ z-score
Mild	>–2
Moderate	Between –2.5 and –2
Moderately severe	Between –3 and –2.5
Severe	Between –4 and –3
Very severe	<–4

Table 1.3 Severity scale for an abnormally low carbon monoxide transfer factor (1).

^aLLN, lower limit of normal.
Adapted and reproduced with permission of the European Respiratory Society: *Eur Respir J* November 2005 26:948-968; doi:10.1183/09031936.05.00035205

Severity classification	T _L CO % MPV
Mild	>60% and <LLN ^a
Moderate	40–60%
Severe	<40

- The coefficient of repeatability (CR) can be used to determine whether a change seen over time is a clinically significant change or simply due to measurement variability.

— From studies collecting repeated measures of lung function in healthy subjects over time, the CR is calculated as two times the standard deviation of the differences between measures. In total, 95% of observations (the difference in value of a parameter across two occasions) should fall within this range and are considered to be within the variability of the measurement. A result (the difference in value of a parameter across two occasions) that exceeds the CR is



General features of interpretation and report writing 7

likely to be clinically significant as it is outside of the variability of the measurement.

— At the time of writing, there are few data available in the literature to establish the CR for parameters of lung function such as FEV₁, FVC and T_LCO. Laboratory biological controls can be used to track the CR of chosen tracked parameters to set limits for significant change (1).

- Limit the number of parameters used for monitoring changes to FEV₁, (F)VC and T_LCO. Other parameters may be monitored for change, but there is increased potential for false-positive results as the likelihood of identifying a change in time increases with the increasing number of parameters monitored.
- Where measurements are made at baseline and post-bronchodilator, comparisons over time should be made between post-bronchodilator results because
 - comparisons between ‘best’ results should be made,
 - baseline conditions may vary (for example, no recent short-acting bronchodilator on one visit, but short-acting bronchodilator within 4 h on another visit).
- When making comparisons to previous results, look at the most recent previous test *AND* test results further back.
 - Sometimes, there may have been no significant change from the immediate previous result, but over the prior 6–12 months there has been a significant change (improvement or decline) in FEV₁, (F)VC or T_LCO.
 - For subjects who are tested infrequently (i.e. years between visits), it may also be necessary to take into account changes due to normal lung ageing. We know, for example, that FEV₁ and FVC decline as age increases once we have reached peak lung function somewhere between 20 and 25 years of age. Studies suggest that in healthy individuals, a loss in volume of up to 30 mL in FEV₁ and FVC per year is possible (5–7).
- Monitoring changes in lung function over multiple visits, rather than just two, and plotting data will also assist with identifying changes that are real (Figure 1.1).
- A change that is within the CR should be documented as ‘no significant change’ rather than ‘no change’ as there may be change, which cannot be differentiated from the variability of the measurement.
- The referring physician, who knows the timing of interventions, is the most appropriate person to interpret change over time. For example:
A finding of no significant change over a specific time is clinically important in the case of a subject with asthma who is having their inhaled

8 Chapter 1

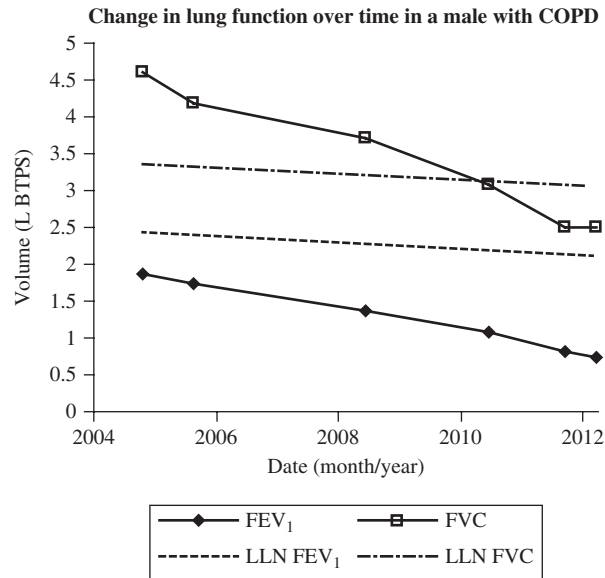


Figure 1.1 In this case, both FEV₁ and FVC appear to be changing more rapidly than expected with normal lung ageing, suggesting a significant change in ventilatory function over time. This case also illustrates no significant change in FVC between the two most recent visits, but a significant fall in FVC between the last and third last visit.

corticosteroid dose back-titrated. The reporter may not be aware of changes in management of disease.

Answering clinical question(s) raised in the referral

- Using the clinical notes, provide some clinical context to the interpretation or recommend further investigations to assist with answering the clinical question. This last step is difficult when the reporter is not the referring physician and has limited or no clinical information to aid in the interpretation. Usually the referring physician is the best individual to form the clinical context based on the 'technical' interpretation and the available clinical information.

General features of report writing

The report that accompanies the lung function results needs to be concise, informative and, where possible, address the clinical question (1). For example, consider the following case:

General features of interpretation and report writing 9

Gender:	Female				
Age (yr):	42	Weight (kg):	60		
Height (cm):	158	Race:	Caucasian		
Clinical notes:	Asthma. For review.				
	Normal range	Baseline	z-score	Post-BD	Change (%)
Spirometry					
FEV ₁ (L)	>2.26	2.74	-0.21	2.85	+4
FVC (L)	>2.80	3.55	+0.24	3.54	0
FEV ₁ /FVC (%)	>72	77	-0.79	81	
Technical comment:	Test performance was good. No bronchodilator use in last 12 h.				

Without worrying too much about how the interpretation was arrived at, a report could be written a number of ways. For example:

- 1 NAD (no abnormality detected).
- 2 Spirometry is normal.
- 3 Spirometry is within normal limits.
- 4 The test is of good quality. Spirometry is within normal limits.
- 5 The test is of good quality. Baseline spirometry is within normal limits. There is no response to inhaled bronchodilator.
- 6 The test is of good quality. Baseline spirometry is within normal limits. There is no response to inhaled bronchodilator. Asthma appears to be currently well controlled, though clinical correlation is required.

Although each example is concise and correct, with increasing number, each example provides more relevant information than those before. The more relevant the information that is given to the referring doctor, the better placed he or she is to make decisions regarding clinical care.

Technical interpretation versus clinical context

Report writing consists of two aspects: technical interpretation and clinical context.

Technical interpretation:

- Can generally be performed without knowledge of the clinical history of the subject.
- Notes the quality of the test performance and the effect of suboptimal quality tests on interpretation.
- Notes appropriateness of reference values used (where necessary).
- Identifies emerging patterns of normality or abnormality.



10 Chapter 1

- See report 5 in the example earlier for a written report with a technical interpretation only.
- Note: The technical interpretation is not to be confused with the technical comment, which is provided by the test operator at the time of the test and addresses any technical issues that may affect the quality and interpretation of the result.

The clinical context:

- Includes the technical interpretation.
- Is reliant on considerable clinical information about the subject being available to the reporter.
- Should be attempted in order to address the clinical question posed in the referral. However, without the necessary clinical information, specific diagnoses based on lung function tests alone should be avoided.
- Is best provided by the referring physician who has the relevant clinical information available. In this case, it is assumed that the referring doctor also has the skills and knowledge to interpret lung function results.
- Report 6 in the example earlier addresses the clinical context considering the clinical notes that have been provided.

Subjectivity

Interpretation of lung function has an element of subjectivity associated with it. Subjectivity in interpretation and report writing may impact the clinical management and care of a subject. The challenge, therefore, is to keep the degree of subjectivity in interpretation and report writing to a minimum.

Reasons for subjectivity include the following:

- The personal opinions and beliefs of the individual writing the report.
- Diversity in the literature for interpretation strategies for some tests.
- Lack of data in the literature for interpretation of particular tests or parameters within tests.
- Knowledge of the clinical background of the subject.

Strategies for reducing subjectivity include the following:

- Utilisation of published interpretation strategies where available.
- Within institutions, agreeing on interpretation strategies for tests lacking published guides or with diverse strategies in published guides.
- Requiring that all reporting personnel within an institution utilise a single, standardised lung function interpretation strategy and use similar reporting phrases – particularly for the technical interpretation.

Where possible, existing interpretation guidelines have been followed for the tests described in this book. However, there are instances where



definitions and interpretative strategies do not exist, and in these cases expert opinion has been used to create a strategy.

Summary

- 1 The written report should provide concise, clear and useful information regarding the test results (1).
 - Keep subjectivity to a minimum.
- 2 Reports should include two components, ideally:
 - a. A technical interpretation – notes quality and inadequacies of reference equations, where applicable. Use known patterns of abnormality to classify any observed abnormalities.
 - b. The clinical context – using the clinical notes, provide some clinical context to the technical interpretation or recommend further investigations to assist with answering the clinical question.

References

- 1 Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, et al. Interpretative strategies for lung function tests. *Eur Respir J*. 2005 Nov; 26(5):948–68.
- 2 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 Jan; 159(1):179–87.
- 3 Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, 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 Dec; 40(6):1324–43.
- 4 Quanjer PH, Pretto JJ, Brazzale DJ, Boros PW. Grading the severity of airways obstruction: new wine in new bottles. *Eur Respir J*. 2013 Aug; 43(2):505–12.
- 5 Xu X, Laird N, Dockery DW, Schouten JP, Rijcken B, Weiss ST. Age, period, and cohort effects on pulmonary function in a 24-year longitudinal study. *Am J Epidemiol*. 1995 Mar; 141(6):554–66.
- 6 Speizer FE, Tager IB. Epidemiology of chronic mucus hypersecretion and obstructive airways disease. *Epidemiol Rev*. 1979; 1:124–42.
- 7 Fletcher C, Peto R. The natural history of chronic airflow obstruction. *Br Med J*. 1977 Jun; 1(6077):1645–8.

