J Bras Pneumol. 2022;48(3):e20220182
https://dx.doi.org/10.36416/1806-3756/e20220182

Calculating the statistical limits of normal and Z-scores for pulmonary function tests

José Alberto Neder[1], Danilo Cortozi Berton[2], Denis E O'Donnell[1]

BACKGROUND

Interpretation of pulmonary function tests (PFTs) requires comparison with the range of expected values in order to discriminate the effects of disease from the normal variability observed in healthy individuals.[(1)] Modern PFT systems provide several different reference equations with automated computation of predicted values and limits of normal. The pulmonologist in charge of the PFT lab should understand how these variables are calculated in order to minimize the risk of overdiagnosis or underdiagnosis.

OVERVIEW

Given the extreme variability of lung function according to sex, age, and body dimensions (particularly height), some basic statistical concepts are applied to differentiate "normality" versus "abnormality". A pragmatic strategy assumes that a) pulmonary function variables measured in a population of interest are equally distributed around the mean; and b) there are more values closer to the mean than further away. Thus, a bell-shaped (Gaussian) curve emerges when we plot the distribution of the values (Figure 1). In a Gaussian distribution, a given percentile represents the value below which a certain percentage of scores fall.

In this context, if a variable has clinical meaning only when abnormally low (e.g., FEV_1), the lower limit of normal (LLN) is set at the value which corresponds to the lowest 5% of the reference population. The LLN can also be roughly estimated as the predicted value minus 1.645 standard deviations from the mean (Z-score; Figure 1A). In simpler words, it means that at the 5th percentile (corresponding to a Z-score of −1.645), there is a 5% chance that the results of a healthy individual will be at or below this level. Accepting a 5% false-positive rate is usually considered acceptable for most clinical applications of PFTs.[(2)] A different scenario emerges when values in both directions (i.e., too low or too high) are clinically relevant, as is the case for some "static" lung volumes which can be reduced in restrictive ventilatory defects or increased in obstructive ventilatory disorders. An acceptable strategy is to divide the 5% error on each end of the distribution using a Z-score of ± 1.96 (Figure 1B),[(3)] establishing the LLN and the upper limit of normal. In Figure 1, a 63-year-old man presents with a measured FEV_1 of 2.07 L (65% of the predicted value). This corresponds to a Z-score of −2.38, that is, below the calculated LLN of 2.41

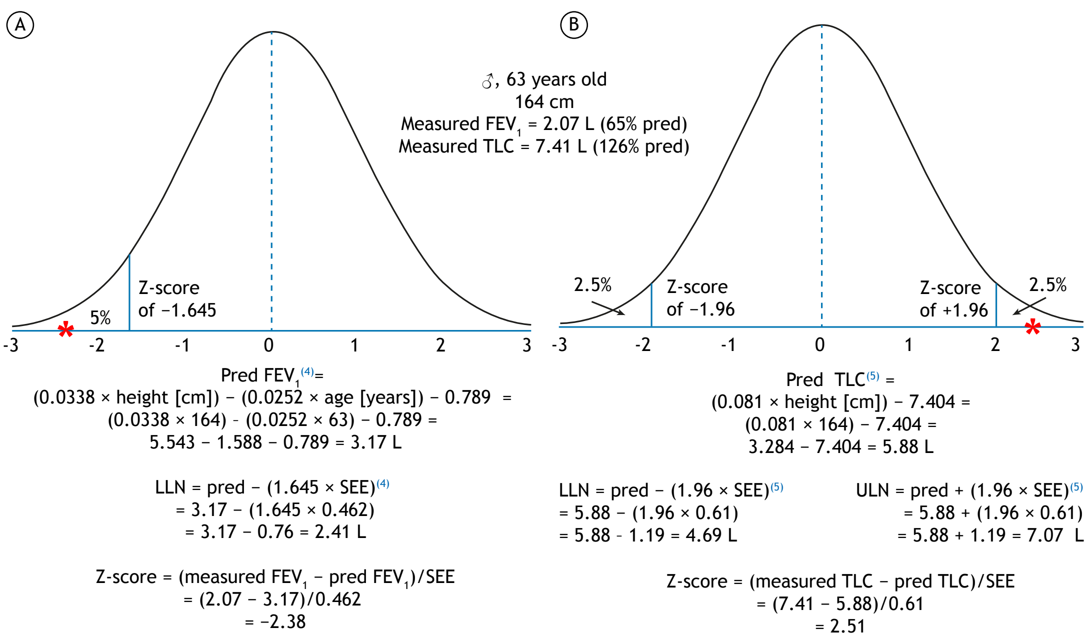

Figure 1. Calculation of key parameters (predicted value, limits of normal, Z-score) necessary for pulmonary function test interpretation: in A, a variable (e.g., FEV_1)[(4)] for which only abnormally low values have clinical meaning; in B, a variable (e.g., TLC)[(5)] for which both abnormally low or high values can be of clinical relevance. pred: predicted; LLN: lower limit of normal; ULN: upper limit of normal; SEE: standard error of the estimate (derived from the respective regression equations).

1. Pulmonary Function Laboratory and Respiratory Investigation Unit, Division of Respirology, Kingston Health Science Center & Queen's University, Kingston, ON, Canada.
2. Unidade de Fisiologia Pulmonar, Hospital de Clínicas de Porto Alegre, Universidade Federal do Rio Grande do Sul, Porto Alegre (RS) Brasil.

L, indicating an abnormally low value (Figure 1A). His measured TLC of 7.41 L, in turn, corresponds to 126% of the predicted value or a Z-score of +2.51, signaling thoracic hyperinflation (Figure 1B).

CLINICAL MESSAGE

The statistical limits of normal do not necessarily separate disease from health. It should also be recognized that there will always be some uncertainty in values close to (i.e., slightly below or above) the LLN or the upper limit of normal: clinical judgment is paramount. The 5th percentile used to define an abnormal test result can be changed depending on the pretest probability of disease, that is, it can either be increased (e.g., in heavy smokers with exertional dyspnea) or decreased (e.g., in asymptomatic non-smokers).

REFERENCES

1. Stanojevic S, Kaminsky DA, Miller M, Thompson B, Aliverti A, Barjaktarevic I, et al. ERS/ATS technical standard on interpretive strategies for routine lung function tests. Eur Respir J. 2021;2101499. https://doi.org/10.1183/13993003.01499-2021
2. Lung function testing: selection of reference values and interpretative strategies. American Thoracic Society. Am Rev Respir Dis. 1991;144(5):1202-1218. https://doi.org/10.1164/ajrccm/144.5.1202
3. Haynes JM, Kaminsky DA, Stanojevic S, Ruppel GL. Pulmonary Function Reference Equations: A Brief History to Explain All the Confusion. Respir Care. 2020;65(7):1030-1038. https://doi.org/10.4187/respcare.07188
4. Pereira CA, Sato T, Rodrigues SC. New reference values for forced spirometry in white adults in Brazil. J Bras Pneumol. 2007;33(4):397-406. https://doi.org/10.1590/S1806-37132007000400008
5. Lessa T, Pereira CAC, Soares MR. Reference equations for plethysmographic lung volumes in White adults in Brazil as derived by linear regression. J Bras Pneumol. 2021;47(1):e20200359. https://doi.org/10.36416/1806-3756/e20200359