

## Pruebas de función respiratoria: ¿interpretación o clasificación?

Luis Torre-Bouscoulet<sup>1</sup>

Dirección Médica, Instituto para el Desarrollo e Innovación en Fisiología Respiratoria, Ciudad de México, México

Las estrategias actuales para la interpretación de las pruebas de función respiratoria (PFR) son en realidad guías para clasificarlas<sup>1</sup>. Los verbos clasificar e interpretar no son intercambiables. Interpretar implica darle sentido a un resultado. La interpretación de los resultados de las PFR le corresponde al clínico, el cual, a la luz de la anamnesis, del examen físico y de otros estudios de diagnóstico, les dará sentido a los resultados funcionales. La interpretación implica integración de información. Por el contrario, clasificar los resultados de los estudios funcionales es algo simple, cuyo objetivo es estandarizar la forma en la que nos referimos a los patrones fisiológicos. Para clasificar de forma estandarizada solo basta adherirse a definiciones operacionales, las cuales tendrán limitaciones intrínsecas. A pesar de las claras diferencias entre clasificar e interpretar, es frecuente que ambas palabras las utilicemos como sinónimos y de ahí surgen la mayoría de las controversias clínicas de los resultados de las PFR.

Los patrones funcionales obtenidos de las PFR solo aumentan o disminuyen la probabilidad preprueba de un diagnóstico nosológico. Los algoritmos de clasificación se basan en modelos matemáticos. Los modelos matemáticos pueden ser determinísticos o probabilísticos. Ejemplo de los primeros son las leyes de la física clásica, en las cuales se conocen en su totalidad los factores que determinan una resultante. Un ejemplo es

la velocidad, que está determinada por la distancia y el tiempo. Por el contrario, en medicina clínica, los modelos suelen ser probabilísticos. Es decir, se construyen modelos mediante los cuales podemos conocer la probabilidad de que ocurra un fenómeno. Los modelos probabilísticos aumentan o disminuyen la certidumbre acerca de la existencia de un fenómeno. Digamos que si obtenemos de un sujeto una capacidad vital forzada (FVC) de 0.8 l, existe «alta» probabilidad (mayor grado de certidumbre) de que el sujeto tenga una enfermedad fibrosante; sin embargo, es posible que ese resultado sea debido a otras causas (debilidad muscular, atrapamiento de aire, etc.). Así, la probabilidad es una medida de la certidumbre y se mide en una escala de 0 a 1. La habilidad predictiva de los modelos matemáticos que se utilizan en fisiología respiratoria, aun incluyendo varios parámetros, oscilan entre 0.4 y 0.6. Lo que falta para alcanzar la unidad es la incertidumbre.

En fisiología clínica, el comparador es, usualmente, una ecuación de regresión múltiple (que toma en cuenta la edad, la estatura y el sexo) que nos permite estimar la función respiratoria que debería tener un individuo si este fuera sano. Los datos con los que se construyen esas ecuaciones (modelos) provienen de sujetos que son considerados sanos. Pero ¿qué es ser sano? Mientras más estrictos sean los criterios para catalogar a una persona como sana, más sesgo de

**Correspondencia:**

Luis Torre-Bouscoulet  
E-mail: luistorreb@gmail.com

Recibido: 31-05-2025

Aceptado: 19-06-2025

DOI: 10.24875/NCT.M25000039

Disponible en línea: 23-10-2025

Neumol Cir Torax. 2025;84(2):163-166

[www.revistancit.org.mx](http://www.revistancit.org.mx)

selección existirá; es decir, los valores funcionales que se obtienen de esos sujetos muy sanos serán inapropiadamente elevados para ser utilizados como comparador, lo que aumentará la tasa de falsos positivos<sup>2</sup>. Lo anterior es solo un ejemplo de las fuentes de variación que tienen las ecuaciones de referencia. En la actualidad se aceptan ecuaciones de referencia construidas con datos provenientes de miles de sujetos sanos de muchos países, por lo que se consideran ecuaciones globales<sup>3</sup>. Con la implementación de ecuaciones globales ha cambiado la prevalencia de anormalidades funcionales y de diagnósticos clínicos, observándose un incremento significativo en la frecuencia de enfermedades respiratorias, en especial en personas afroamericanas<sup>3,4</sup>.

Asumiendo que las ecuaciones de referencia para función pulmonar representan fielmente a los sujetos sanos de una población, se agrega otro problema que es la definición de normalidad. Hemos aceptado la definición de normalidad basada en frecuencias; sin embargo, lo frecuente no es sinónimo de lo normal, o quizás sí, si aceptamos que lo normal es lo frecuente; sin embargo, ser normal (frecuente) no implica estar sano. Un ejemplo es el índice de masa corporal (IMC). Por frecuencia, tendríamos que aceptar que el IMC normal, al menos en México, es el que va de 28 a 32 pero eso, aunque frecuente y por tanto normal, no es sano. No debemos confundir «normal» con «sano».

Para facilitar la definición de normal, usamos como referente la distribución de los datos. Una distribución estándar normal es aquella que tiene una media aritmética de cero y desviación estándar de uno. En esas condiciones, la media, la mediana y la moda serán idénticas. Se ha usado esa curva estándar como un referente para definir normalidad<sup>5</sup>. Volvamos al ejemplo de la FVC. Si una persona tiene una FVC de 4.9 l y la media del grupo de referencia es 5 l, existe alta probabilidad de que sea «normal», porque los 4.9 l se alejan muy poco de la media, 5 l. ¿Qué tan cerca de la media debe estar la medición obtenida del sujeto para que sigamos considerando que es «normal»? Lo que separa al valor medido (de una persona) de la media de referencia (del grupo) lo podríamos obtener en las unidades originales, por ejemplo, litros en el caso de la FVC; pero también lo podemos expresar en una unidad estandarizada de la dispersión de los datos que se conoce como desviación estándar. Así, el número de desviaciones estándar es una forma uniforme y comparable para saber cuánto se aleja el valor obtenido de una persona en relación con la media grupal. Un sinónimo de unidades de desviación

estándar es el puntaje Z o score-Z<sup>6</sup>. El puntaje Z es una forma estandarizada de expresar la dispersión de los datos<sup>6</sup>.

En una curva estándar normal, el 90% de las observaciones se encuentran entre -1.64 y +1.64 de puntaje Z, el 95% entre -1.96 y +1.96 y el 99% entre -2.58 y +2.58. Por convención, el límite inferior normal (LIN) se ha establecido en la percentila 5, que corresponde a -1.64 de puntaje Z. Note que el 5% de los sujetos sanos tendrán valores inferiores al LIN (más a la izquierda, valores más negativos). Así, tácitamente, la definición de normalidad basada en el LIN deja fuera al 5% de personas sanas y es, en consecuencia, un punto de corte imperfecto.

Para ilustrar las limitaciones de los puntos de corte veamos el siguiente ejemplo. El cociente volumen de aire expulsado en el primer segundo de una espiración forzada/capacidad vital forzada (FEV<sub>1</sub>/FVC) es el parámetro espirométrico para definir obstrucción, así que casi cualquier neumólogo aceptaría que tener un FEV<sub>1</sub>/FVC < LIN significa obstrucción. En consecuencia, un resultado > LIN sería considerado como no obstrucción. Sin embargo, cuando se utilizan definiciones de obstrucción menos estrictas (FEV<sub>1</sub>/FVC < 0.7) ya se identifican anormalidades funcionales concomitantes al compararse con los sujetos catalogados como normales por el criterio más estricto (< LIN). Dichas anormalidades funcionales incluyen menor difusión pulmonar de monóxido de carbono, menor eficiencia ventilatoria en el ejercicio, así como menor consumo de oxígeno<sup>7</sup>.

En el área bajo la curva considerada como normal también existen diferencias. No es lo mismo, en términos de mortalidad, ser normal cuando el FEV<sub>1</sub> se encuentra a -1.5 de puntaje Z, que ser normal con FEV<sub>1</sub> de +0.5 de puntaje Z. Ambos serían normales, pero el primero tiene mayor mortalidad que el segundo. Ese gradiente de mortalidad en sanos se ha demostrado recientemente en un estudio llevado a cabo en los EE.UU. que incluyó a dos cohortes con seguimiento de 20 años<sup>8</sup>. Por consiguiente, está claro que la naturaleza no entiende de puntos de corte. La función pulmonar es una variable continua, por lo que resulta un tanto infructuoso agrupar sujetos por intervalos de función respiratoria. Un estudio británico reciente que incluyó a más de 300 mil sujetos demostró (tal como lo hizo John Hutchinson en 1846) que la función pulmonar en sus unidades originales (litros de FVC y de FEV<sub>1</sub>), sin ajuste por covariables, es predictor de mortalidad a 12 años de seguimiento<sup>9,10</sup>.

Ante las muchas limitaciones de la definición de normalidad que utilizamos en la actualidad, algunas de ellas expuestas arriba, viene la interrogante acerca de qué deberíamos hacer para tener una lectura más eficiente de los resultados de las PFR y darles un mejor sentido. Concluir que de tal a tal valor de FEV<sub>1</sub> o de FVC es normal o anormal, y que de tal a tal valor es leve, moderado o grave es, claramente, simplificar los hallazgos funcionales.

Una alternativa ante la problemática señalada es realizar mediciones repetidas de la función respiratoria, lo que permite describir el curso temporal de los parámetros funcionales, también llamados trayectorias funcionales. Existe diversos factores pre- y posnatales que en la edad pediátrica afectan el desarrollo pulmonar y que, en consecuencia, afectan la función pulmonar que alcanzará el sujeto en la vida adulta<sup>11</sup>. La función pulmonar que alcanza el sujeto a los 20-25 años depende, en resumen, de factores genéticos y ambientales. El reto clínico y epidemiológico es identificar e incidir sobre los factores modificables con el propósito de cambiar la trayectoria de la función respiratoria para que el sujeto alcance una función pulmonar óptima. De lo contrario, el sujeto llegará a la vida adulta con una función pulmonar reducida y, a partir de ese punto, iniciará la disminución funcional progresiva asociada, tanto al envejecimiento, como a factores exposicionales o a enfermedades crónicas. Así, conviene dejar un poco de lado el análisis de la función pulmonar de forma transversal para hacerlo de forma longitudinal, en la cual el sujeto es su mismo comparador. Existen diversos métodos para evaluar los cambios longitudinales, tanto en niños como en adultos<sup>1</sup>. El método de FEV1Q es uno de los métodos más sólidos para dicho propósito en población adulta<sup>12</sup>. FEV1Q se obtiene al dividir el valor de FEV<sub>1</sub> (l) entre 0.5 en hombres y entre 0.4 en mujeres. Dichos divisores son el percentil 1 del FEV<sub>1</sub> en población sana para cada sexo. Mientras el cociente (FEV1Q) se acerque más a 1, más deteriorada se encuentra la función respiratoria. Un descenso de una unidad en dicho cociente toma, en condiciones normales, 18 años. Si la reducción de una unidad se observa, digamos, en 10 años; el individuo está presentando una disminución acelerada de la función respiratoria. Para otros parámetros funcionales, diferentes de los espirométricos, la información acerca del comportamiento longitudinal es limitada y es una importante área de investigación. La gran mayoría de las trayectorias funcionales se han establecido a partir de estudios transversales; sin embargo, se requiere un mayor número de estudios de cohorte a largo plazo para disminuir las fuentes de variación.

Otra alternativa en el proceso de interpretación de las PFR que va más allá del análisis transversal o longitudinal es la medicina de sistemas y la inteligencia artificial (IA)<sup>13</sup>. La investigación biomédica ha evolucionado de manera sorprendente. Hasta hace algunos años la medicina era reactiva y enfocada en enfrentar los problemas de salud que emergían en la población. Se basaba en descripciones clínicas, series de casos o ensayos clínicos simples. Posteriormente, vinieron estudios que proponían identificar fenotipos y que se basan especialmente en ómica. La multiómica, que estudia la interacción entre las ómicas, ha permitido entender mejor la biología de sistemas, en donde lo más importante no son los elementos individuales sino las interacciones entre ellos, y ha sido un paso adelante en la generación e implementación de lo que conocemos como la medicina de precisión<sup>14</sup>. En este contexto las PFR se suman, mediante un conglomerado funcional, a la red de información de las gigantescas bases de datos, lo que permite crear modelos biológicos complejos que sirven para anticipar, con mínimo error, la respuesta terapéutica y el pronóstico de un sujeto individual. En esas circunstancias le estaremos dando un mejor sentido a la valoración funcional respiratoria. Mientras tanto, la interpretación de las PFR será siempre (como dice el Dr. Alberto Neder) un estudio con  $n = 1$ <sup>15</sup>.

A manera de corolario, me gustaría enunciar lo siguiente. Primero, los resultados transversales de PFR generan errores de clasificación inherentes al punto de corte que se utilice. Segundo, la normalidad estadística no equivale necesariamente a salud clínica. Tercero, el análisis individual de la trayectoria funcional respiratoria parece ser una estrategia más eficiente en la interpretación de las PFR. Cuarto, la incorporación de la función respiratoria en modelos complejos basados en IA podría ser útil para identificar mejor las prioridades en salud.

## Financiamiento

La realización del presente trabajo no ha recibido ayudas específicas provenientes de agencias del sector público, sector comercial o entidades sin ánimo de lucro.

## Conflicto de intereses

El autor declara no tener conflicto de intereses.

## Consideraciones éticas

**Protección de personas y animales.** Los autores declaran que para esta investigación no se han realizado experimentos en seres humanos ni en animales.

**Confidencialidad, consentimiento informado y aprobación ética.** El estudio no involucra datos personales de pacientes ni requiere aprobación ética. No se aplican las guías SAGER.

**Declaración sobre el uso de inteligencia artificial.** Los autores declaran que no utilizaron ningún tipo de inteligencia artificial generativa para la redacción de este manuscrito.

## Referencias

1. Stanojevic S, Kaminsky DA, Miller MR, Thompson B, Aliverti A, Barjaktarevic I, et al. ERS/ATS technical standard on interpretive strategies for routine lung function tests. *Eur Respir J.* 2022;60(1):2101499. <https://doi.org/10.1183/13993003.01499-2021>
2. Eisen EA, Wegman DH, Louis TA, Smith TJ, Peters JM. Healthy worker effect in a longitudinal study of one-second forced expiratory volume (FEV1) and chronic exposure to granite dust. *Int J Epidemiol.* 1995;24(6):1154-61. <https://doi.org/10.1093/ije/24.6.1154>
3. Bhakta NR, Bime C, Kaminsky DA, McCormack MC, Thakur N, Stanojevic S, et al. Race and ethnicity in pulmonary function test interpretation: an Official American Thoracic Society Statement. *Am J Respir Crit Care Med.* 2023;207(8):978-95. <https://doi.org/10.1164/rccm.202302-0310st>
4. Moffett AT, Bowerman C, Stanojevic S, Eneanya ND, Halpern SD, Weissman GE, et al. Global, race-neutral reference equations and pulmonary function test interpretation. *JAMA Netw Open.* 2023;6(6):e2316174. <https://doi.org/10.1001/jamanetworkopen.2023.16174>
5. The Standard Normal Distribution [Internet]. Scibbr [consultado 26 mayo 2025]. Disponible en: <https://www.scribbr.com/statistics/standard-normal-distribution/>
6. Standardizing a normal distribution [Internet]. Scibbr [consultado 26 mayo 2025]. Disponible en: <https://www.scribbr.com/statistics/standard-normal-distribution/>
7. Neder JA, Milne KM, Berton DC, de-Torres JP, Jensen D, Tan WC, et al.; CRRN (Canadian Respiratory Research Network) and the CanCOLD (Canadian Cohort of Obstructive Lung Disease) Collaborative Research Group. Exercise tolerance according to the definition of airflow obstruction in smokers [letter]. *Am J Respir Crit Care Med.* 2020;202(5):760-2. <https://doi.org/10.1164/rccm.202002-0298le>
8. Cannon MF, Goldfarb DG, Zeig-Owens RA, Hall ChB, Choi J, Cohen HW, et al. Normal lung function and mortality in World Trade Center Responders and National Health and Nutrition Examination Survey III Participants. *Am J Respir Crit Care Med.* 2024;209(10):1229-37. <https://doi.org/10.1164/rccm.202309-1654oc>
9. Zhou L, Yang H, Zhang Y, Wang Y, Zhou X, Liu T, et al. Predictive value of lung function measures for cardiovascular risk: a large prospective cohort study. *Thorax.* 2024;79(3):250-8. <https://doi.org/10.1136/thorax-2023-220703>
10. Kouri A, Dandurand RJ, Usmani OS, Chow CW. Exploring the 175-year history of spirometry and the vital lessons it can teach us today. *Eur Respir Rev.* 2021;30(162):210081. <https://doi.org/10.1183/16000617.0081-2021>
11. Agusti A, Faner R. Lung function trajectories in health and disease. *Lancet Respir Med.* 2019;7(4):358-64. [https://doi.org/10.1016/s2213-2600\(18\)30529-0](https://doi.org/10.1016/s2213-2600(18)30529-0)
12. Balasubramanian A, Wise RA, Stanojevic S, Miller MR, McCormack MC. FEV<sub>Q</sub>: a race-neutral approach to assessing lung function. *Eur Respir J.* 2024;63(4):2301622. <https://doi.org/10.1183/13993003.01622-2023>
13. Torre-Bouscoulet L. Medicina respiratoria de sistemas. *Neumol Cir Torax.* 2023;82(1):36-7. <https://dx.doi.org/10.35366/114227>
14. Torre-Bouscoulet L. Los retos de la medicina personalizada. *Neumol Cir Torax.* 2015;74(4):238-9.
15. Neder JA. The new ERS/ATS standards on lung function test interpretation: some extant limitations. *Eur Respir J.* 2022;60(2):2200252. <https://doi.org/10.1183/13993003.00252-2022>