

La necesidad de objetivar en medicina y sus tribulaciones

J. Sanchis Aldás

Departamento de Medicina. Universitat Autònoma de Barcelona. Departamento de Neumología. Hospital de la Santa Creu i Sant Pau. Barcelona.

«Una diferencia, para constituer un diferencia, debe ser una diferencia.»

Introducción

Objetivar es describir algo en términos precisos, distinguirlo del resto. Para ello, se requiere observar cuidadosamente, comparar, medir y, finalmente, describir lo identificado. Los párrafos siguientes son unas pinceladas sobre estos procesos, imprescindibles para el progreso de la medicina y necesarios en la actividad clínica. En modo alguno pretenden ser exhaustivos. Podría escribirse varios artículos más sobre otros aspectos omitidos en éste, y también relevantes en la objetivación, pero el tiempo y el espacio son dimensiones que constantemente imponen sus límites en toda nuestra actividad.

Hasta bien entrado el siglo XVII, ya descrita la circulación sanguínea por Harvey y Servet, la misión atribuida al pulmón en el organismo era la de enfriar el fuego producido por el corazón en su actividad vital y procurarle acolchamiento para mitigar sus golpes. En este contexto, resulta difícil imaginar necesidad alguna de medir nada. El panorama cambió gradualmente en la segunda mitad del siglo XVII y, a un ritmo más vivo, en el siglo XVIII, con la explosión de la investigación científica, en particular en la física, la química y la biología. De forma paralela, se impulsaron y perfeccionaron los sistemas de medidas.

A la vista del tremendo desbarajuste y las dificultades de equivalencia entre unidades que recibían el mismo nombre en países o incluso ciudades distintas, con subdivisiones diferentes, se hizo necesario el intento de un sistema universal que pudiera aplicarse a las tres dimensiones a las que habitualmente reducimos nuestro mundo físico: longitud, masa y tiempo. Antoine-Laurent Lavoisier (1743-1794) destaca, entre otros méritos, por propugnar el uso del sistema decimal, junto a Laplace y Talleyrand-Perigord. Este último, como miembro influyente de la Asamblea Nacional, propuso en 1790 a la Academia de Ciencias el nuevo sistema de pesos y medidas, basado en una constante física apropiada:

la longitud de un péndulo cuyo período es un segundo. En este mismo año, la Academia decidió que la unidad de longitud para el nuevo sistema universal de medidas sería la diez millonésima parte del cuadrante de un meridiano terrestre, posteriormente denominada metro. Delambre y Méchain, dos astrónomos geodésicos, fueron encargados de verificar la medición del cuadrante, y para ello escogieron un sector de aproximadamente la novena parte de un cuadrante, desde Dunquerque, en el canal de la Mancha, a Montjuïc, en Barcelona; una distancia aproximada de 1.000 km en línea recta. Para la medición invirtieron casi siete años¹. Hoy día, el sistema internacional de unidades (SI) se basa en “siete unidades bien definidas que, por convención, se consideran independientes en sus dimensiones: metro, kilogramo, segundo, amperio, kelvin, mol y candela”. Los intentos de decimalizar el tiempo (horas y días) fracasaron ya en el mismo inicio del sistema, y se suspendieron en 1795 “por tiempo indefinido”. El calendario republicano de años con 12 meses de 30 días y 5 días festivos añadidos subsistió hasta 1806 en que fue abolido por Napoleón.

Las necesidades y dificultades de objetivar con mediciones son, en medicina, las mismas que las de las ciencias en que intenta basarse. A ellas se añaden las propias de los procesos biológicos en los que se reúnen fenómenos paramétricos y categóricos. Para estos últimos deben arbitrarse escalas cuando se los quiere manejar en términos cuantitativos. Los párrafos siguientes se refieren someramente a aspectos que ilustran la historia reciente de estas necesidades y dificultades, al tiempo que revisan nociones básicas relacionadas con el concepto de normalidad, la cuantificación, el proceso de medición y sus instrumentos. Por último, se dedican unas breves palabras a la expresión de la probabilidad en el lenguaje con que el clínico se comunica diariamente.

Normalidad y enfermedad

En un breve y extraordinariamente lúcido relato de la evolución de las ideas y técnicas de medición de la enfermedad, Cochrane² describe tres fases en el desarrollo de sus estudios epidemiológicos:

1. La primera se centró en el desarrollo de técnicas para la recopilación de datos de salud en comunidades, que requerían mediciones precisas y reproducibles.

Correspondencia: Dr. J. Sanchis Aldás.
Departament de Pneumologia. Hospital de la Santa Creu i Sant Pau.
Avda. Sant Antoni Maria Claret, 167. 08025 Barcelona.
Correo electrónico: jsanchis@hsp.santpau.es

Recibido: 19-9-2001; aceptado para su publicación: 9-10-2001.

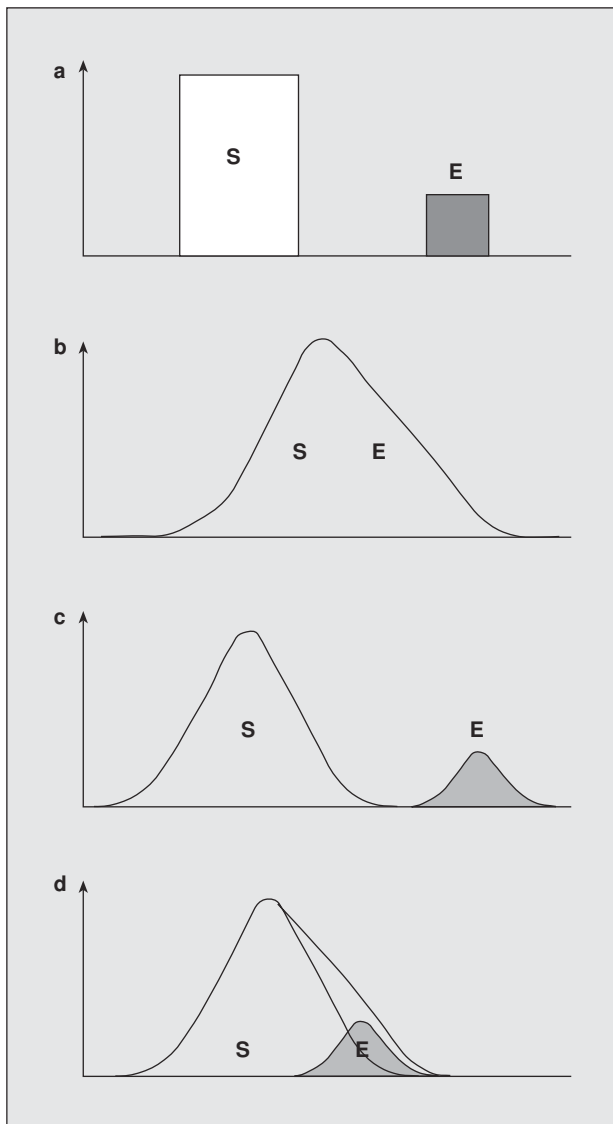


Fig. 1. Distribución de datos de individuos sanos (S) y enfermos (E) en una población general: *a*) como la desea el clínico; *b*) la observada por el epidemiólogo; *c*) la que se obtendría con muestras significativamente amplias, según la creencia generalizada, y *d*) la realidad más común en biología, con dos posibles variantes y un mismo abordaje².

2. La segunda siguió a la de la recolección de los datos. Ésta le enfrentó con el problema de encontrar el significado de las “anomalías” observadas y superar el importante error de asimilar los conceptos de “anomalía” y “enfermedad”; error extraordinariamente común y prevalente, aún hoy día. Cochrane y sus colaboradores abordaron este problema mediante la correlación transversal de las distintas variables registradas, en un intento de interpretar los hallazgos. Así, observaron que algunos de ellos se asociaban claramente con enfermedad, mientras que otros no pasaban de simples anomalías sin repercusión aparente.

3. Cochrane describe la tercera fase muy gráficamente como la “búsqueda de la dicotomía”, que consistió en intentar la detección de la bimodalidad en la distribu-

ción de los datos de las variables, para conseguir “valores de normalidad”, que les permitieran distinguir a los individuos sanos de los enfermos (fig. 1a). El resultado fue exactamente el opuesto al esperado: la mayoría de las veces la distribución fue continua, ligeramente sesgada hacia un lado y no dicotómica (fig. 1b). En un primer momento, la reacción fue la de considerar que las muestras no habían sido suficientemente grandes como para permitir separar claramente las poblaciones de sanos y enfermos (fig. 1c). La realidad de las variables médicas y de la mayoría de las variables biológicas es que se comportan como continuas (fig. 1d). En la medicina clínica persiste la convicción sólida de la existencia general de la distribución dicotómica, que permitiría clasificar sin problemas un dato como normal o patológico, pero la terca realidad es que la dicotomía es extraordinariamente infrecuente.

En este punto debe recordarse, de acuerdo con Murphy³, que un problema particular de la cuantificación radica en la elección de la dimensión que se va a medir. Por ejemplo, en el supuesto de que se quisiera averiguar en un grupo de pacientes si hay dos tipos de hematíes que se diferencian por su tamaño, se podría medir el diámetro, el área o el volumen de los hematíes. La interpretación de los datos dependerá, en parte, de la dimensión escogida, pues cuando la distribución de los datos presenta una moda única, si los datos se elevan a una potencia, incluso a una fracción de potencia, la distribución puede transformarse en bimodal. Esto puede influir decididamente sobre la conclusión a la que se puede llegar⁴.

La búsqueda de valores de normalidad y su definición plantea problemas de parecida dificultad. La idea de que todo lo que se necesita es definir el intervalo comprendido por la media de la distribución ± 2 desviaciones típicas y considerar “anormales” a aquellos individuos o grupos que queden fuera de este intervalo, carece simplemente de pruebas y base racional. Aceptar, por ejemplo, que la desviación típica de una distribución muy sesgada sirva para definir el intervalo buscado o que las desviaciones por encima y por debajo de la media tengan el mismo significado patológico desafía el sentido común más elemental. Ello llevó a Cochrane a preguntar: ¿quiénes de los que tengan su colesterol sérico por debajo de la media o su capacidad vital por encima quisieran tener un valor más próximo al de la media?

El tratamiento de la distribución de datos para determinar intervalos de referencia u otros usos tiene varias posibles soluciones⁵, que dependen en parte de las características del fenómeno que se analice y que exceden el ámbito de este artículo. Para las enfermedades tratables, Cochrane propuso la búsqueda del punto o puntos en la curva de distribución en la que existen pruebas de que el tratamiento produce más beneficio que perjuicio².

Cuantificación

Este aspecto lo ilustra Murphy³ y se refiere a la necesidad de abordar las observaciones en términos cuantita-

tivos, para evitar los juicios de valor, los sesgos y el azar. Los sistemas de medición diferentes del cegesimal se diseñan generalmente *ad hoc*, ante la exigencia impuesta por la interpretación de datos ya existentes. De acuerdo con Murphy, un ejemplo clásico es el denominado cociente de inteligencia (IQ), construido con un conjunto arbitrario de datos de tal manera que proporcionan una distribución Gaussiana (“normal”), lo que requirió atribuir arbitrariamente un peso diferente a sus diversos componentes. Éste es un procedimiento frecuente en los intentos actuales de cuantificación en la medicina clínica. La misma manera en que tratamos de pensar sobre cualquier rasgo clínico tiende a producir una distribución acampanada. Así, tendemos a pensar lógicamente que la mayoría de los individuos poseen una inteligencia mediana (mediocre, precisa Murphy), ordinaria o regular (fig. 2a). A uno de los lados del grupo medio se sitúa un pequeño grupo de personas notables, con talento y, en el lado opuesto, otro grupo pequeño de deficientes mentales. Un número mucho menor de individuos de inteligencia extraordinaria, los que podrían denominarse genios, se sitúan junto a los talentosos. En el otro extremo quedan los intelectualmente incapacitados. No puede, por definición, haber un gran número de individuos excepcionales, por lo que necesariamente los grupos de los extremos son pequeños. La fuerza principal que determina el patrón de la curva es, por decirlo así, la solidaridad de los individuos medianos. Sin embargo, dos fuerzas actúan en contra de la concentración central.

La primera es que los individuos medianos, siendo mayoría, necesitan competir entre sí, para lo que tienden a realzar las diferencias que pueda haber entre ellos, y que para los más inteligentes y los deficientes serían escasamente perceptibles. Puesto que a menudo es el mediano quien decide sobre los estándares, el peso de los datos en la zona central tiende a extenderla, en interés de una mayor discriminación, lo que no ocurriría en una escala de datos objetivos como, por ejemplo, el número de hematíes o la talla. Como el área de la zona central (número de individuos) es la misma cualquiera que sea su forma, la concentración disminuye en el centro si la extensión aumenta hacia los lados (fig. 2b).

La segunda fuerza es la polarización y realiza la falsa agrupación de datos. Por ejemplo, en un momento dado los psicólogos distinguieron los tipos introvertido y extravertido. Más tarde se les atribuyó un riesgo coronario diferente a cada uno, hasta que un análisis cuantitativo más cuidadoso determinó que los individuos *borderline* de dichas agrupaciones superaban con mucho a los dos tipos, haciendo insostenible la existencia de auténticas categorías. Una de las formas más eficaces de abordar el problema del subjetivismo en las escalas es contrarrestar las tendencias mediante la clasificación por percentiles (fig. 2c)⁶.

Resolución, precisión y exactitud

En el intento de cuantificar un fenómeno intervienen tres elementos importantes para la representación de los resultados: resolución, precisión y exactitud (fig. 3). Los tres tienen una definición y significado particulares

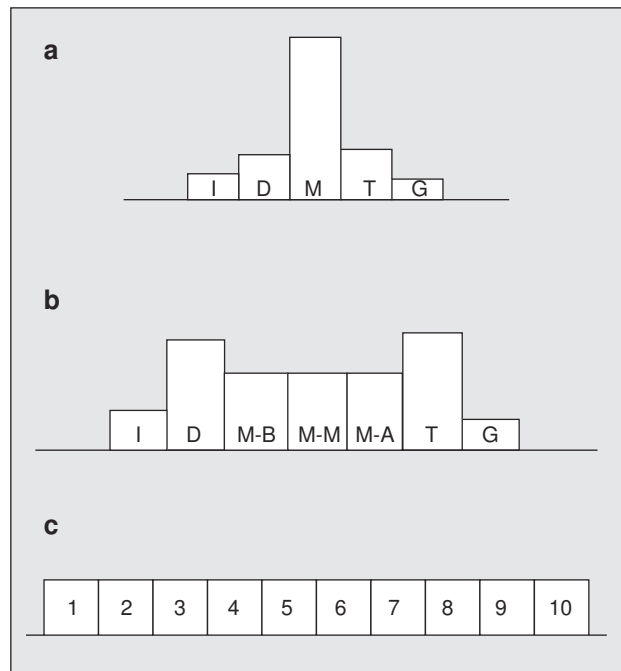


Fig. 2. Escalas categóricas. Vicisitudes de la construcción de una escala cuantitativa para medir el cociente de inteligencia: a) como se imaginó *a priori*; b) como hubo que ajustarla; c) como debería haberse construido. I: incapacitados; D: deficientes; M: medianos; T: talentosos; G: genios; M-B: medianos bajos; M-M: medianos medios; M-A: medianos altos³.

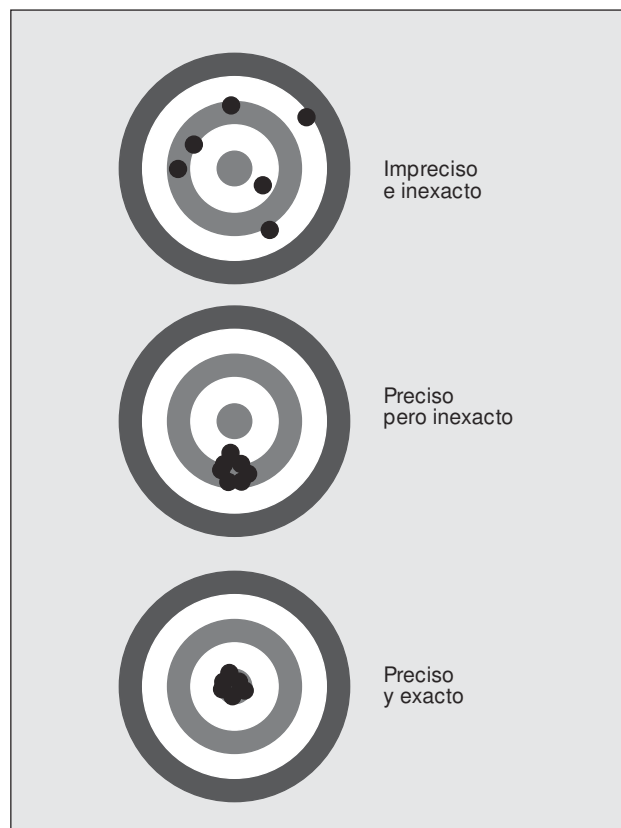


Fig. 3. Precisión y exactitud de los disparos de tres rifles diferentes sobre la diana.

y, a menudo, son confundidos. Aunque se refieren primariamente a datos numéricos, por analogía pueden aplicarse también a las ideas. Los párrafos siguientes son un breve extracto de las reflexiones de Murphy sobre los tres elementos mencionados⁷.

Resolución o poder de resolución de un instrumento es el margen de incerteza con que es capaz de medir (su escala de medida). Un importante factor limitante de la medición es el refinamiento del aparato con que se mide. Al medir una aguja con una escala calibrada en centímetros puede encontrarse que si se coloca un extremo de la aguja en el 0, el otro extremo se sitúe entre los 7 y los 8 cm. Si no se permite al observador interpolar o guiarse por su impresión subjetiva, la medición no puede expresarse con mayor certeza: el poder de resolución es 1 cm. Con dicho instrumento no se puede distinguir entre 7,2 y 7,7 cm. Si se requiere una mayor resolución se necesita calibrar el instrumento en divisiones menores; por ejemplo, milímetros. Pero la mejora de la exactitud no es infinita y, además, depende también de lo que se mide. Por ejemplo, un fisiólogo puede querer medir el tiempo que un músculo debe estar sometido a ejercicio hasta que aparezca dolor. Pero la apreciación del dolor es muy subjetiva, y muchos pacientes tendrían dificultades para decidir con exactitud el momento en que aparecen molestias, y éstas se transforman en dolor. Así que si el tiempo transcurrido se midiera con un cronómetro calibrado en décimas de segundo, la resolución no mejoraría apreciablemente al sustituirlo por un reloj atómico, que permitiera mediciones exactas de 1/200 de segundo.

Los problemas del poder de resolución de un instrumento tienen similitudes con los de las ideas. Un ejemplo sencillo es la medición de la altura de una persona. Si se observa una ficha clínica, la talla aparece como un concepto claro, diáfano. Sin embargo, al enfrentarse con la medición de la talla debe decidirse previamente si el paciente al que se mide debe o no llevar los zapatos puestos. El extremo superior de la persona puede tomarse como el punto más alto de su cabello, o bien determinarse colocando horizontalmente un bloque rígido sobre su cabeza y fijando la posición de la barbilla. Habrá más cuestiones que determinar, como la presión que debe ejercer el bloque rígido de medir sobre el cráneo, si se le pedirá al paciente que se estire tanto como pueda, etc. De esta forma la reproducibilidad de mediciones puede ser muy satisfactoria. La cuestión previa es qué se entiende exactamente por la altura de la persona; cuestión que exige una definición cuidadosa *a priori*.

Precisión, en el contexto científico, se refiere a la reproducibilidad de una estimación. No es una característica de las medidas, es una afirmación sobre un valor al que se ha llegado mediante inferencia indirecta o probabilística. Se expresa por el error típico de la estimación. Para medir la longitud de unas piezas de acero mediante una regla calibrada, se alinea un extremo con el 0 y se determina el punto de la escala alcanzado por el otro extremo. En este proceso, la incerteza se genera por el poder de resolución de la escala calibrada. Pero hasta ahora no se ha dicho nada sobre posibles problemas a la hora de alinear la pieza medida con el 0 de la escala o

en su otro extremo. Puede que exista alguna pequeña variación al colocar el objeto sobre la escala, que ocasione incerteza en las mediciones consecutivas. La precisión, como término técnico, se refiere a la variación obtenida con la repetición "ciega" de determinaciones de la misma medida. La variación puede reflejar errores técnicos, cambios espontáneos del individuo y diferencias entre individuos. Es necesario que las determinaciones sean "ciegas", pues la persona que conoce los resultados de las anteriores tiende siempre a hacer mediciones que concuerdan con las previas.

Una gran precisión indica una buena reproducibilidad de los resultados, pero no garantiza que éstos sean correctos. Es posible ser preciso y también ser precisamente erróneo.

Exactitud es el grado con que la muestra estimada de un valor representa el valor real, verdadero. Si el grado de precisión expresa la variación de la estimación, el de exactitud significa el sesgo de la estimación. Mientras que la precisión puede explorarse y valorarse de forma empírica, la exactitud tropieza con la dificultad, con frecuencia insalvable, de llegar a conocer el valor real. Por convención, el gasto cardíaco se estima indirectamente por la aplicación del principio de Fick, como indicador del trabajo realizado por el corazón. Esta técnica es el estándar para los otros métodos indirectos de medir el gasto. En su favor, podría argumentarse que si sus estimaciones resultan sesgadas es razonable suponer que un corazón cuyo gasto sea 12 l/min está trabajando el doble que otro cuyo gasto sea 6 l/min. Pero el sesgo no tiene que ser necesariamente proporcionado. Además, la premisa exigida por el método de Fick, de que la sangre en la que se va a medir el contenido de O₂ está completamente mezclada, puede que se cumpla en el corazón normal y no en el enfermo. En un mundo ideal debería disponerse de una medición absoluta, auténtica, del gasto que permitiera comparar los métodos indirectos y determinar sus sesgos. Se conseguirían así métodos corregidos satisfactorios. En el mundo real, es preciso a menudo conformarse con las mejores mediciones indirectas y tomar éstas como patrón, aunque no se puede dejar de ser consciente del salto dado.

Las características de la precisión y la exactitud pueden trasladarse al campo de las ideas. Una idea precisa es aquella que comunica la misma noción a todos los individuos competentes. Naturalmente, el grado de coincidencia o acuerdo entre ellos dependerá también del nivel de resolución que se intenta. Así, la pregunta "¿qué trastorno pulmonar produce de forma característica una resonancia normal o aumentada a la percusión, junto con ruidos respiratorios disminuidos o ausentes?" es una pregunta muy precisa. Los clínicos competentes identificarían en ella los signos de neumotórax con colapso pulmonar. En el otro extremo, una idea inexacta puede también generar una respuesta muy reproducible pero falsa.

El ejemplo del tratamiento clínico de la trombosis ilustra el problema de la inexactitud. Un trombo es una agregación de elementos sólidos de la sangre en el interior de un vaso. En la práctica clínica, la presencia de un trombo se reconoce indirectamente por las interfe-

rencias que produce en la circulación del territorio afectado. Como su reconocimiento es tardío, se han dedicado esfuerzos extraordinarios para detectarlo en sus fases más tempranas. En su momento, los investigadores, al observar que la sangre extravasada forma al coagularse una masa sólida, de las mismas características, concluyeron que los dos procesos, la trombosis y la coagulación, estaban íntimamente relacionados. Fue muy decepcionante comprobar que las pruebas diseñadas para valorar la coagulación tenían un poder de predicción muy pobre sobre la aparición de trombos, pese a que dichas pruebas de coagulación eran muy reproducibles. La equivalencia entre coagulación y formación de trombos resultó falsa, sólo con un papel menor en las fases más avanzadas del proceso trombótico, cuando el daño ya está hecho. La lesión de una pared vascular considerada después el mecanismo iniciador de la trombosis era demasiado rápida para atribuirse a la coagulación, y requería fármacos y medidas terapéuticas diferentes.

Las ideas aparentemente claras son muy seductoras, en particular si se basan en mediciones precisas, que pueden ocultar el hecho de que sus resultados no satisfagan las premisas.

Expresiones cuantitativas

En el lenguaje médico, y también en el común, es un hábito y una necesidad usar expresiones cualitativas de probabilidad, como *frecuente*, *común*, *probable* o *raro*, tantas veces empleadas en este mismo artículo. El empleo de expresiones cualitativas de probabilidad por los médicos ha sido analizado por Kong et al⁸, entre otros investigadores. El problema estriba en la ausencia de un significado cuantitativo, expreso, aceptado universalmente, lo que lleva a errores de interpretación en el diálogo entre profesionales y entre médico y paciente. Si *probable* significa 80% de las veces para un médico y 50% para otro, su actuación puede ser diferente ante las mismas circunstancias clínicas.

El problema semántico se complica por el hecho de que el contexto tiene un efecto real sobre el significado de las expresiones cualitativas. Así, por ejemplo, cuando se habla de efectos adversos de fármacos, el término *raro* conlleva una probabilidad mucho menor si se refiere a efectos indeseables de un bloqueador beta que cuando indica los de un antihistamínico. En esta comparación, el 60% de un grupo de médicos entendía < 1% en el caso del bloqueador beta y sólo el 21% de médicos entendió < 1% cuando se refirió al antihistamínico⁹. También el significado cuantitativo cambia según la gravedad de la enfermedad o situación de la que se trata.

En ocasiones, la utilización de una expresión cualitativa en la comunicación científica se debe al desconocimiento de la probabilidad real o a la deficiente definición del aspecto que se valora. Por otro lado, buena parte de la variación interindividual podría reducirse de existir un acuerdo general de correspondencia entre expresiones y números que se aproximara al uso general. No obstante, incluso así, en algunos momentos se siente mayor comodidad en el uso de una expresión cualitativa que de un valor exacto. La solución propuesta por Kong et al es trabajar hacia la consecución de unas equivalencias numéricas de probabilidad de las expresiones, que progresivamente se generalizarían, de la misma forma que el significado de las palabras se difunde y mantiene uniforme gracias a su definición codificada en los diccionarios.

Como conclusión general, cabe reiterar que sin medir no se puede objetivar. La medición es el soporte fundamental y característico de la actividad científica. La comunicación de una observación requiere idéntico rigor al empleado en el proceso de objetivarla. El ejercicio de la medicina científica está sometido a los mismos requisitos y tribulaciones. La aceptación de la disciplina necesaria para dicho ejercicio es a la vez exigente y gratificante.

«La inexactitud es el derroche de lo inapreciable, de lo que no puede recuperarse, el tiempo.»

J.S. Bach

BIBLIOGRAFÍA

1. Sant' Ambrogio G, Dejourns P. On the origin of the Metric System. NIPS 1995;10:46-9.
2. Cochrane AL. The history of the Measurement of Ill health. Int J Epid 1972;1:89-92.
3. Murphy EA. Measures. En: Murphy EA, editor. A companion to medical statistics. Baltimore: The Johns Hopkins University Press, 1985; p. 192-216.
4. Baker GA. Transformations of bimodal distribution. Ann Math Stat 1930;1:334-44.
5. Chin S. Ranges, confidence intervals, and related quantities: what they are and when to use them. Thorax 1991;46:391-3.
6. Herrera L. The precision of percentiles in establishing normal limits in medicine. J Lab Clin Med 1958;52:34-42.
7. Murphy EA. Resolution, precision, and accuracy: practical aspects. En: Murphy EA, editor. The logic of medicine. Baltimore: The Johns Hopkins University Press, 1976; p. 71-80.
8. Kong A, Barnett GO, Mosteller F, Youtz C. How medical professionals evaluate expressions of probability. N Engl J Med 1986; 315:740-4.
9. Mapes REA. Verbal and numerical estimates of probability in therapeutic contexts. Soc Sci Med 1979;13A:277-8.