

EPISTEMOLOGÍA DE LA MEDICIÓN

EPISTEMOLOGY OF MEASUREMENT

Santiago Cuéllar Rodríguez

Real Academia Nacional de Farmacia

autor de correspondencia: santiago.cuellar.rodriguez@gmail.com

REVISIÓN

El texto de este artículo se corresponde, in extenso, con la ponencia que fue expuesta bajo el mismo título por el autor el día 17 de febrero de 2022, en la mesa redonda "Medición y ciencia. Aspectos epistémicos y repercusiones metodológicas y regulatorias en investigación clínica", dentro del ciclo "Epistemología, metodología y regulación de la medición científica en el ámbito clínico", celebrado en la Real Academia Nacional de Farmacia los días 17 y 22 de febrero de 2022.

RESUMEN

Medir designa el proceso por el que se interacciona con un sistema con el objetivo de representar alguno de sus aspectos mediante términos abstractos o que tienen un carácter meramente ideal, estableciendo referencias comunes que permitan compararse con un patrón prefijado y acotado por márgenes de incertidumbre predefinidos. Obviamente, la medición es un integrante esencial de la ciencia; uno de sus sellos distintivos, fuente primaria de conocimiento y el imprescindible nexo de unión entre la teoría y la experimentación, permitiendo que se retroalimenten reciprocamente y mejoren continuamente. En el estudio científico de cualquier fenómeno, la intervención del observador y el propio hecho de la observación perturban al fenómeno estudiado. Por ello, siempre es preciso tomar en consideración el grado de perturbación que produce, algo que en investigación clínica puede ser particularmente complejo por la multiplicidad y variabilidad de las manifestaciones humanas. La complejidad del sujeto observado (ser humano), dada su condición de ente pensante, emocional, moral y social, multiplica la incertidumbre de sus respuestas tanto individuales como colectivas, extremando las exigencias metodológicas y éticas del científico; además, la respuesta de cada individuo a una intervención clínica incorpora otros múltiples factores no relacionados con los efectos objetivamente achacables a la intervención, pero que deben ser debidamente ponderados. Por todo ello, para poder medir científicamente el efecto de una intervención en un ensayo clínico, es fundamental definir y justificar previamente la magnitud mínima del cambio que puede dar lugar a un efecto que sea clínicamente relevante para el paciente y para el observador.

ABSTRACT

Measuring designates the process by which a system is interacted with in order to represent some of its aspects by means of abstract terms or those that have a merely ideal character, establishing common references that allow comparison with a predetermined pattern and bounded by predetermined margins of uncertainty. Obviously, measurement is an essential component of science; one of its hallmarks, a primary source of knowledge and the essential link between theory and experimentation, allowing them to feed off each other and continuously improve. In the scientific study of any phenomenon, the intervention of the observer and the very fact of the observation disturb the phenomenon studied. For this reason, it is always necessary to take into consideration the degree of disturbance it produces, something that in clinical research can be particularly complex due to the multiplicity and variability of human manifestations. The complexity of the observed subject (human being), given his condition as a thinking, emotional, moral and social entity, multiplies the uncertainty of his responses, both individual and collective, and extreme methodological and ethical demands of the scientist; In addition, the response of each individual to a clinical intervention incorporates multiple other factors unrelated to the effects objectively attributable to the intervention, but which must be duly weighed. For all these reasons, in order to scientifically measure the effect of an intervention in a clinical trial, it is essential to previously define and justify the minimum magnitude of the change that can give rise to an effect that is clinically relevant for the patient and for the observer.

Palabras Clave:

método científico
filosofía de la ciencia
epistemología; medición
experimentalidad;
reproducibilidad
investigación clínica

Keywords:

scientific method
philosophy of science
epistemology
measurement
experimentality
reproducibility
clinical research



1. INTRODUCCIÓN

La ciencia implica tanto adquisición sistemática de conocimiento como su acumulación organizada y su desarrollo evolutivo. A ello cabe agregar un tercer elemento fundamental: la ciencia es una referencia del conocimiento o, al menos, de un cierto tipo de conocimiento. Por eso, la ciencia es — debería ser — pública; es decir, accesible para cualquier persona suficientemente cualificada en la materia en cuestión, con la única exigencia de conocer el método científico estándar; también debería ser concreta, expresada en términos claramente definidos y no ambiguos; como dice Mario Bunge, “la ciencia constituye una rebelión contra la vaguedad y la superficialidad del sentido común” (1). Asimismo, debería estar sujeta a una arquitectura simbólica específica y compartida (matemática, química, etc.), independiente de la lengua particular de cada científico. Finalmente, la ciencia debe ser objetiva, es decir, ajena a ideas particulares (2).

Existe un amplio consenso referido a que el principal objetivo de la ciencia es crear una imagen o un modelo de la realidad hasta donde la evidencia empírica permita establecer, sin olvidar que todo modelo es infiel a la realidad en muchos más aspectos de los que es fiel, ya que solo busca ilustrar aspectos muy concretos de esa realidad. Además, en ocasiones, la simulación computacional de un modelo muestra más la habilidad del programador que la plausibilidad de una explicación. Como dijo Rodney Brooks, “el problema con las simulaciones es todas que están condenadas al éxito” (3). Es decir, tanto lo bueno como lo malo de los modelos es que son fácilmente manipulables.

El método científico, a pesar de no ser uniforme, siempre se fundamenta en la retroalimentación entre teoría y experimento; de hecho, es este último — el experimento — lo que diferencia al método científico de otras formas rigurosamente racionales y sistemáticas de conocimiento, como la metafísica o la historia. Por tanto, la ciencia no se basa solo en deducciones teóricas ya que, por muy bien construida que esté una formulación matemática, no demuestra por sí misma la realidad del fenómeno modelizado. Fuera de la evidencia experimental, la ciencia entra en la desnuda especulación y ahí está obligada a ceder el testigo a la metafísica, que tiene la obligación — y el derecho — de plantearse preguntas que la ciencia tiene el derecho — y la obligación — de ignorar.

Ni siquiera la unanimidad de los observadores garantizan la conformidad de la interpretación de un fenómeno con la realidad porque, como decía Thomas Kuhn, “todas las observaciones están cargadas de teoría”. Más allá de la mera perspectiva científica del investigador, su producción de ciencia no puede evitar la influencia de su carrera profesional, ni la de los objetivos e intereses de las instituciones o empresas donde desarrolla su labor, ni la de los va-

lores morales, sociales y culturales imperantes en su comunidad. También para Karl Popper toda teoría científica tiene un excedente metafísico, porque de hecho afirma más de lo que puede ser testado (4).

En cualquier caso, es difícil no estar de acuerdo con Richard Feynman, premio Nobel de Física, quien aseguraba que “el experimento es el único juez de la verdad científica”. El problema es que, como decía Sydney Brenner, premio Nobel de medicina, “Nos estamos ahogando en un mar de datos y estamos hambrientos de conocimiento”.

En resumen, la utilidad de la ciencia como pretensión de observación y explicación racional de la realidad, solo alcanza hasta donde llegan sus pretensiones legítimas de conocimiento experimental. Más allá, se abre un océano desconocido donde la ciencia no tiene referencias ni salvavidas a los que agarrarse. Así pues, la esencia del método científico es la retroalimentación continua entre una teoría cada vez más explicativa y un experimento cada vez más esclarecedor; de hecho, lo que caracteriza al método científico de otras formas de conocimiento es la experimentación y, particularmente, la reproducibilidad de los resultados experimentales.

Si algo han aprendido los científicos es que la inmutabilidad de las leyes físicas no existe y su simetría — su aplicabilidad general en todo orden, tiempo y espacio — está plagada de fisuras. Además, algunas formas de simetría científica, como la invarianza de escala, han sido desestimadas por la teoría de la relatividad y la mecánica cuántica. Paradójicamente, estas rupturas de la simetría han permitido obtener conocimientos más amplios, exactos y precisos; por ello, la aparente fragilidad de la ciencia constituye muchas veces su verdadera fortaleza.

La urgencia oportunista por acaparar méritos académicos, profesionales, empresariales o institucionales para la promoción de su posición profesional y económica, lleva a algunos científicos a dejarse en la cuneta aspectos tan relevantes como el rigor metodológico o la ética, formulando a veces hipótesis delirantes con retazos deshinchados de otras bien fundamentadas y documentadas, ocultando o falsificando sus métodos reales de investigación, haciéndola intencionadamente irreproducible u obteniendo conclusiones arbitrarias pero a la medida de sus intereses (5).

2. REDUCCIONISMO Y DETERMINISMO CIENTÍFICO

Existen diferentes planteamientos sobre la relación que existe entre el conocimiento científico y la realidad. Van desde el *realismo ingenuo*, en el que los enunciados de la ciencia corresponden directamente a la realidad, hasta el *instrumentalismo* puro, donde los enunciados son meros artefactos solo útiles para representar a los fenómenos. En cualquier caso, la ciencia trabaja con



modelos que, aunque pueden ser muy útiles, son inevitablemente *reduccionistas*. Son los investigadores translacionales y los tecnológicos los que se encargan de transitar desde el conocimiento científico más general, básico o fundamental, hasta sus aplicaciones más concretas.

Con el término reduccionismo tratamos de expresar la necesidad que tiene la ciencia de proceder a una simplificación de la realidad, empujándola al considerar solo una mínima cantidad de sus elementos, a los que convierte artificialmente en parámetros que, además, solo pueden ser manejados matemáticamente al mismo tiempo en un pequeño número de forma fiable. Más allá de no tomar en cuenta muchos de los elementos de esa realidad, hay una cantidad indeterminada de otros elementos que son desconocidos, con el agravante de no saber cuántos ni cómo son, ni cuál es su peso relativo en el conjunto, ni las interacciones existentes entre ellos y los propios elementos conocidos. Confundir el modelo con la realidad que pretende describir es a lo que alude la famosa frase acuñada por Alfred Korzybski en 1931: "el mapa no es el territorio" (6).

Por otro lado, la ciencia es obligadamente *determinista* porque si no limitase artificialmente el grado de incertidumbre causal de la realidad profunda, así como los efectos de la imprecisión e incompletitud de nuestros métodos de investigación, la ciencia no pasaría de ser un decepcionante relato voluntarista, en el que la estéril vaguedad sustituiría a la duda racional y a la buena predicibilidad de sus asertos.

Porque, por encima de todo, la ciencia busca establecer una explicación de la forma más independiente posible del observador y del emisor de la hipótesis. Con la ciencia se pretende entender y, si es posible, mostrar la existencia de una relación entre fenómenos, entre posibles motivos y previsibles consecuencias. Según el *principio de causalidad*, atribuido a Leibniz, todo suceso está determinado por una causa; es decir, todo tiene — al menos — una razón o causa suficiente o, por decirlo de otra manera, todo lo que ocurre tiene una razón o explicación justificativa para ser de una determinada manera y no de otra, y la causa precede siempre al efecto y lo determina, y éste incorpora todas las características de la causa. Llevado a su máximo extremo, si conociésemos el estado actual del mundo con total precisión, podríamos predecir cualquier evento en el futuro, como aseguraba Laplace.

Sin embargo, la teoría de la relatividad y — muy especialmente — la física cuántica han supuesto importantes retos para la visión determinista de la ciencia. Más allá están las denominadas *ciencias de sistemas*, que estudian las relaciones entre los componentes de cada sistema complejo, en los que la realidad está estratificada en niveles, cada uno de los cuales está limitado — pero no absolutamente determinado — por los niveles inferiores de realidad.

Esto vale tanto para una máquina elemental como para el cosmos global.

Todo el conocimiento humano avanza sobre los hallazgos previamente obtenidos por nuestros antecesores, ampliándolos, corrigiéndolos o desechándolos. Es decir, no existe la obtención de conocimientos *ex nihilo* (a partir de la nada), sino que cualquier sistema racional se fundamenta evolutivamente sobre las bases de los antecedentes, aunque cada cierto tiempo se produzcan verdaderos "terremotos" que reorganizan el marco racional de entendimiento y la búsqueda de nuevos hallazgos que profundizan en la comprensión de la realidad. La ciencia, obviamente, no es una excepción a esta regla; antes bien, representa su ejemplo más paradigmático.

Por ello, la ciencia muestra un nivel de credibilidad muy superior a cualquier otro sistema de conocimiento. La sociedad española, como la del resto de los países avanzados, valoran positivamente a la ciencia. Según el *Observatorio Español de I+D+i*, el 46% los ciudadanos españoles tenían en 2020 una imagen positiva de la ciencia y la tecnología, considerando que los beneficios que aporta superan con creces a los perjuicios que puede ocasionar (7). En ese mismo año, España publicó 105.461 documentos científicos, lo que representa un 3,3% de la producción científica mundial, ocupando la posición 11ª (China era la 1ª, con 750.000, seguida de Estados Unidos, con algo más de 600.000, y Gran Bretaña, con 200.000). El porcentaje de publicaciones españolas de excelencia — aquellas se incluyen en el conjunto del 10% de los artículos más citados de su área — se situó en un 15,1%, en línea con Francia (15,3%), Estados Unidos (15,2%) o Alemania (15,7%).

3. TEORÍA Y EXPERIMENTACIÓN

La experimentación no siempre ocupa el mismo lugar cronológico en la evolución del conocimiento científico. Einstein formuló la teoría de la relatividad sentado en su despacho de la oficina de patentes de Berna (Suiza) y de forma prácticamente exclusiva a partir de conceptos teóricos. No resolvió casi ningún problema experimental; de hecho, tuvieron que pasar bastantes años hasta que llegase la primera confirmación empírica de sus predicciones teóricas. El desarrollo de la mecánica cuántica, en cambio, les vino impuesto a los teóricos por los impactantes resultados de los experimentos; podríamos decir que ningún científico teórico en su sano juicio habría inventado la mecánica cuántica sin verse forzado a ello por los datos experimentales. Lo que no impidió que se convirtiese en la teoría más exitosa de la historia de la ciencia (8).

Uno de los aspectos más sobresalientes de la ciencia durante los siglos XIX, XX y, desde luego, en el XXI es su matematización. La formulación matemática de cualquier hipótesis o ley



científica se hace cada vez más prolija, hasta el punto de que incluso puede llegar a suplantar a la propia experimentación. A este respecto, el premio Nobel de física Roger Penrose — nada sospechoso de estar en contra de las matemáticas en ningún ámbito, dada su condición de físico teórico y matemático — decía que la coherencia matemática está lejos de ser un criterio suficiente para decirnos si es probable que estemos en la pista correcta; de hecho, normalmente son las condiciones físicas las que tienen la mayor importancia (9).

Como el propio Penrose indica, el progreso de la ciencia física se ha conseguido al encontrar el equilibrio correcto entre las restricciones, tentaciones y revelaciones de la teoría matemática, por una parte, y de la observación precisa de las acciones del mundo físico, por otra, normalmente mediante experimentos cuidadosamente controlados. El riesgo ocurre — como afirmaba Guillermina Díaz — cuando la matemática se establece como una cárcel de la realidad, momento en el que la matemática pierde por completo todo su sentido (10).

Lo que le proporciona al agua sus características físico-químicas tan particulares — tanto que, gracias a ellas, existe la vida — no es la estructura aislada de su molécula (dos átomos de hidrógeno unidos a uno de oxígeno: H_2O) sino los efectos de las interacciones que se producen cuando se juntan muchas de ellas. Es decir, hay propiedades que emergen de la multiplicidad, como una manifestación de la complejidad.

Algunos científicos consideran que algo parecido puede suceder con el espacio y con otros elementos que, hasta ahora, considerábamos como constituyentes primarios — irreductibles — de la realidad, como la energía, el tiempo o la materia. Según esta perspectiva, el continuo espacio-tiempo podría considerarse como un fenómeno que emerge a partir de algo más fundamental; es decir, no sería un elemento de partida, sino una consecuencia — un resultado — de la complejidad. En términos matemáticos, una dimensión no es más que es un eje de coordenadas, un concepto simbólico abstracto que no está necesariamente vinculado al mundo material.

Albert Einstein unificó con su teoría de la relatividad especial el espacio y el tiempo, reduciendo las categorías físicas fundamentales a tres: espacio-tiempo, materia y energía; y no se conformó solo con eso, sino que con la relatividad general incluyó a la gravedad en la estructura del propio espacio-tiempo, considerándola como una deformación de éste, es decir, un elemento geométrico. Theodor Kaluza consideró que el electromagnetismo también podría ser explicado como un artefacto de la geometría del espacio-tiempo, comprobando que esa nueva dimensión — la quinta, tras las cuatro espacio-temporales “convencionales” — encajaba perfectamente en la formulación matemática. Para intentar darle un sentido físico a esa dimensión extra, Oskar Klein propuso

que cada punto del espacio-tiempo cuatridimensional estaría rodeado por una minúscula dimensión “compacta”, que él calculó que tendría un radio de 10^{-30} cm, es decir, varias órdenes de magnitud más pequeño que el del átomo del hidrógeno, de 10^{-8} cm. La incorporación de estas dimensiones compactas adicionales — que estarían enrolladas sobre sí mismas — resultaron muy útiles en la formulación matemática para incluir a las fuerzas nucleares débil y fuerte en el modelo atómico; en una primera formulación se añadieron otras cinco a las cinco previas, lo que nos lleva a las 10 dimensiones de las primeras versiones de la “teoría de cuerdas”.

Una “cuerda” es definida como una especie de hilo unidimensional en vibración, que sería el constituyente de las partículas de materia (fermiones) y de interacción (bosones), las cuales solo se diferenciarían según el modo de vibración de la cuerda, y gracias a las dimensiones extra “compactas”, las trayectorias de las partículas no serían líneas sino planos envolventes o tubos. De acuerdo con Edward Witten, las diversas versiones de la teoría de cuerdas formarían parte de una teoría más general, la “teoría M”, la cual requiere una dimensión extra adicional hasta totalizar 11. Versiones posteriores de la teoría de cuerdas elevan el número de dimensiones hasta 26.

Sin embargo, conviene no olvidar que todas ellas no son más que modelos matemáticos, aunque extremadamente sofisticados e incluso intrigantes — algunas evidencian interconexiones inesperadas entre diversas ramas fundamentales de las matemáticas — pero que están muy lejos de ser demostrables empíricamente. Por otro lado, la(s) teoría(s) de cuerdas (o de supercuerdas, si se considera la supersimetría) se enfrenta con el reto de no ser una teoría científica en sentido estricto, ya que presenta una enorme multiplicidad de posibles soluciones, lo que le hace perder su capacidad predictiva. Con todo, incluso aunque no se correspondiese con la verdad profunda de la realidad, la teoría de cuerdas — incluyendo las dimensiones “extra” — ha hecho notables aportaciones a la fundamentación matemática de la ciencia.

4. REPRODUCIBILIDAD Y REPETIBILIDAD EXPERIMENTAL

Como indicábamos anteriormente, la teoría y la experimentación son las dos caras inseparables de la ciencia. Sin embargo, las pruebas o evidencias experimentales proporcionadas por un estudio científico proceden tanto de construcciones artificiales en torno a una hipótesis fundamentada y formalizada matemáticamente, como de determinaciones precisas de parámetros que han de poder ser reproducibles. Con todo, la relación entre la experimentalidad y el marco conceptual científico — la teoría — es siempre compleja, entre otros muchos motivos, porque los criterios empleados en la verificación experimental inevitablemente son diversos, evolucionan



rápida, dependen del marco conceptual científico del que se parte y, en algunas ocasiones, pueden estar sesgados por intereses personales, editoriales, empresariales o institucionales

Antes mencionábamos la reproducibilidad como una característica propia de las pruebas o evidencias científicas. En este sentido, podemos definir la **reproducibilidad experimental** como la cualidad por la que un experimento científico es capaz de producir idénticos resultados — o muy similares, dentro de un estrecho margen de tolerancia — cuando se repite en las mismas condiciones y con los mismos métodos descritos para dicho experimento.

Aunque muchos científicos consideran que la reproducibilidad experimental es un sinónimo de la **repetibilidad** o **replicabilidad**, habitualmente se define a esta última como la propiedad de obtener resultados consistentes al replicar — repetir — un experimento científico pero partiendo de un conjunto distinto de datos de los empleados en el estudio original, aunque coherentes con estos últimos, siguiendo la misma metodología experimental. En cualquier caso, es evidente la estrecha relación que vincula ambos términos.

Parece obvio que la experimentalidad de un estudio científico debe implicar necesariamente que la hipótesis pueda ser comprobada; es decir, que pueda ser refutada — desmentida — o, utilizando la terminología de Karl Popper, *falsable*. La falsabilidad significa que los contraejemplos — datos experimentales contrarios a la hipótesis testada — pueden ser lógicamente posibles y que la viabilidad para ser observados es real. En resumen, si no puede probarse experimentalmente que una hipótesis podría ser falsa, tampoco podría probarse que es científicamente cierta (11).

La reproducibilidad es un aspecto fundamental del método científico, sin la que se corre el riesgo de facilitar la multiplicación de estudios estériles que llegan a producir una hipertrofiada bibliografía, repleta de información falsa o, al menos, de una cuestionable fiabilidad, que por desgracia casi nunca se llega a corregir o depurar. De hecho, uno de los principales motivos de falta de reproducibilidad es la opacidad de algunos artículos científicos publicados en cuanto a referir detalladamente protocolos, métodos, conjuntos y origen de los datos, procesos de cálculo y de validación estadística, entre otros motivos, lo que impide reproducir hallazgos potencialmente (o solo aparentemente) relevantes.

La falta de reproducibilidad científica — total o parcial — no es, por desgracia, algo infrecuente. Un estudio elaborado por John Ioannidis y publicado en 2018, mostraba que dos de cada tres artículos destacados de psicología y uno de cada tres de economía experimental no eran reproducibles experimentalmente (12). Estudios aún más recientes han mostrado que al intentar reproducir 193 experimentos en biología del cáncer (*The Reproducibility Project: Cancer Biology*), que habían sido descritos en 53 artículos de alto impacto, solo se consiguió con 50 de tales experimentos (26%), en 23 de los artículos (43%) (13).

5. EPISTEMOLOGÍA DE LA MEDICIÓN

La última edición (23ª) del *Diccionario de la lengua española*, de la Real Academia Española, define a la epistemología como “teoría de los fundamentos y métodos del conocimiento científico” (14). Esta definición es excesivamente genérica, por lo que podríamos concretar algo más si decimos que la epistemología tiene como objeto el origen, la naturaleza y la validez del conocimiento científico o — en otras palabras — qué valor tienen, cómo se generan y cómo se validan los conocimientos proporcionados por la ciencia.

La **medición** es una parte esencial de la ciencia; uno de sus sellos distintivos, una fuente primaria de conocimiento y, al mismo tiempo, el imprescindible nexo de unión entre la teoría y la experimentación, en un proceso complejo mediante el que ambos se complementan y retroalimentan recíprocamente de forma continua.

¿Pero, qué significa “medir”? Básicamente, designa el proceso por el que se interacciona con un sistema con el objetivo de presentar alguno de sus aspectos mediante términos abstractos (clases, números, vectores, etc.) o que tienen un carácter meramente ideal (paciente promedio, etc.), estableciendo referencias comunes que permitan ser comparadas con un patrón prefijado y acotado por márgenes de incertidumbre predeterminados (15).

Martin Heidegger citaba la frase de Max Planck, “Es real lo que se deja medir”, como quintaesencia del pensar calculador que caracteriza a la ciencia moderna; la mensurabilidad es lo que decide lo que debe valer en la ciencia. Sin embargo, Heidegger advertía que no debemos entender el término “calcular” en el simple sentido de “operar con números”, sino en “contar con algo”, “poner en ello nuestra expectativa”; de este modo, toda objetivación de lo real es calcular, ya sea persiguiendo los efectos de las causas, ya haciéndose imágenes sobre los objetos, ya asegurando conexiones de secuencia y de orden (34).

Es bien conocido que, en el estudio de cualquier fenómeno, la intervención del **observador** y el propio proceso de la **observación** afectan significativamente (es decir, perturban) al fenómeno estudiado. Esta perturbación limita en cierto grado e incluso llega a “congelar” en el tiempo la percepción de la realidad, por lo que la visión del observador es siempre incompleta. En cualquier caso, esta perturbación provocada por la medición afecta a todos los fenómenos reales y a cualquier nivel o escala: desde el nanomundo de la física cuántica hasta las descomunales dimensiones de la astrofísica, pasando — por supuesto — por la investigación farmacológica básica y clínica; de ahí que siempre sea preciso tomar en consideración el grado de perturbación que produce la propia medición científica. O, lo que es lo mismo, la metaciencia, la ciencia investigándose a sí misma.



Nadie puede negar que la ciencia moderna está sujeta a una rápida y continua evolución, ampliando y diversificando sus métodos. Podemos decir, por tanto, que cada vez es más **multidisciplinar**; es decir, exige la participación de áreas cada vez más diversas de conocimiento; pero también es más **interdisciplinar**, ya que combina esas áreas para alcanzar objetivos compartidos y, además, las conecta matemática y metodológicamente; y esto significa que también es cada vez más **transdisciplinar**.

Tal multitud y diversidad de disciplinas científicas interconectadas implica el empleo de técnicas, unidades y terminologías que no siempre son coherentes ni directamente inteligibles entre ellas. Y ello hace que algunas mediciones no se validen de la manera adecuada. A esto cabe agregar que, en determinadas ocasiones, los datos primarios o “en bruto” — *raw data* — son sometidos por intereses espurios a ciertos procedimientos intencionadamente sesgados de refinado estadístico — *cocinado* — que desvirtúan el resultado real.

Nadie puede negar que la teoría química funciona excepcionalmente bien al ser transpuesta al ámbito empírico. Pero, como admiten muchos de sus más reconocidos especialistas, incluyendo varios Premios Nobel, la química está, de hecho, entretrejida de ficciones tan convenientes como cuestionables, como son la electronegatividad, el estado de oxidación, el tautomerismo o la acidez (16). En este sentido, decía el premio Nobel de química Roald Hoffmann que cualquier definición rigurosa del enlace químico está condenada a ser empobrecedora; “vale la pena divertirse con la riqueza difusa de la idea”.

Ciertamente, en ciencia se trabaja muchas veces con aspectos de la realidad cuya naturaleza profunda — *su esencia* — es desconocida, como la energía, la materia, el espacio o el tiempo, a pesar de ser capaces de medir, contar, obtener resultados e incluso predecir con enorme precisión y exactitud muchos fenómenos ligados a ellos, así como algunas de sus interacciones, conexiones e interconversiones.

Pese a todo, hemos sido capaces de desplegar un enorme esfuerzo tecnológico — y económico — en beneficio de la medición científica. Para ilustrar la extraordinaria magnitud que llegan a alcanzar algunas instalaciones, podemos observar en la figura 1 una foto del dispositivo CMS (*Compact Muon Solenoid; Solenoide compacto de muones*) del Gran Colisionador de Hadrones (*Large Hadron Collider; LHC*) del CERN (*Centre Européen pour la Investigation Nucléaire; Organización Europea para la Investigación Nuclear*), en Ginebra. El CMS es uno de los ocho diferentes experimentos que actualmente hay en CERN, junto con el ATLAS, ALICE, TOTEM, LHCb, LHCf, FASER y MoEDAL. El CMS pesa 14.000 toneladas, mide 29 metros y tiene un diámetro de 15 metros, pero no es el mayor del CERN, ya que el ATLAS (*A Toroidal LHC Apparatus;*

Aparato Toroidal del LHC) mide 46 metros, tiene un diámetro de 25 y pesa alrededor de 7.000 toneladas, mientras que ALICE (*A Large Ion Collider Experiment; Experimento Colisionador de Grandes Iones*) mide 26 metros, tiene un diámetro de 16 y pesa alrededor de 10.000 toneladas.

Podríamos citar también otras megainstalaciones científicas como el Super Kamiokande (Japón), el LIGO (EEUU) o el proyecto de anillo de 100 km para sustituir al actual de 27 km del LHC del CERN (en China hay otro con este mismo tamaño, ya en construcción).

Paradójicamente, a pesar de contar con instalaciones de investigación científica tan impresionantes como las mencionadas, seguimos sin saber exactamente qué son la energía, la materia, el espacio y el tiempo. Como ocurre con la química, los físicos también reconocen su frustración al no saber qué es, en realidad, una partícula elemental (17).

Con frecuencia, confundimos **predictibilidad con comprensibilidad** porque, aunque la ciencia nos ofrece una imagen de que el mundo puede ser bastante predecible, eso no significa necesariamente que sea suficientemente comprensible. De hecho, nada impide — en principio — que dos o más hipótesis científicas (modelos) diferentes puedan ser coherentes con un mismo conjunto de datos u observaciones.

El físico David Mermin, de la Universidad de Cornell, afirmaba en 1989 que “si me viera obligado a resumir en una frase lo que me dice la interpretación de Copenhague de la mecánica cuántica, sería *¡Cállate y calcula!*”. Es decir, deberíamos contentarnos con una interpretación instrumentalista — un formalismo matemático puramente simbólico — que funcione; no importa que la teoría no reproduzca la realidad, en tanto que permita encajar adecuadamente los resultados de los experimentos. Cabe indicar que Mermin se desdijo de tal afirmación algunos años después.

La mecánica cuántica funciona excepcionalmente bien en cuanto que tiene una impresionante capacidad para predecir y justificar con enorme exactitud y precisión numerosos resultados experimentales; de hecho, fue desarrollada precisamente para dar cuenta de fenómenos observados que no podía ser explicados mediante la física clásica. ¿Por qué preocuparnos por lo que significa? Esa parece ser la actitud de muchos científicos, tal como se desprende de una encuesta realizada entre físicos por parte de la Universidad de Aarhus de Dinamarca en 2016, que mostró que solo una minoría entendía realmente el significado de la interpretación de Copenhague o incluso conceptos fundamentales de la mecánica cuántica. Y sin embargo, la exploración de cuestiones filosóficas, inútiles solo aparentemente, puede tener profundas consecuencias prácticas (18).

Es habitual la confusión entre el (in)determinismo y la (im)previsibilidad. En realidad, el (in)determinismo o (in)certidumbre es una propiedad ontológica, relativa a la esencia de las enti-

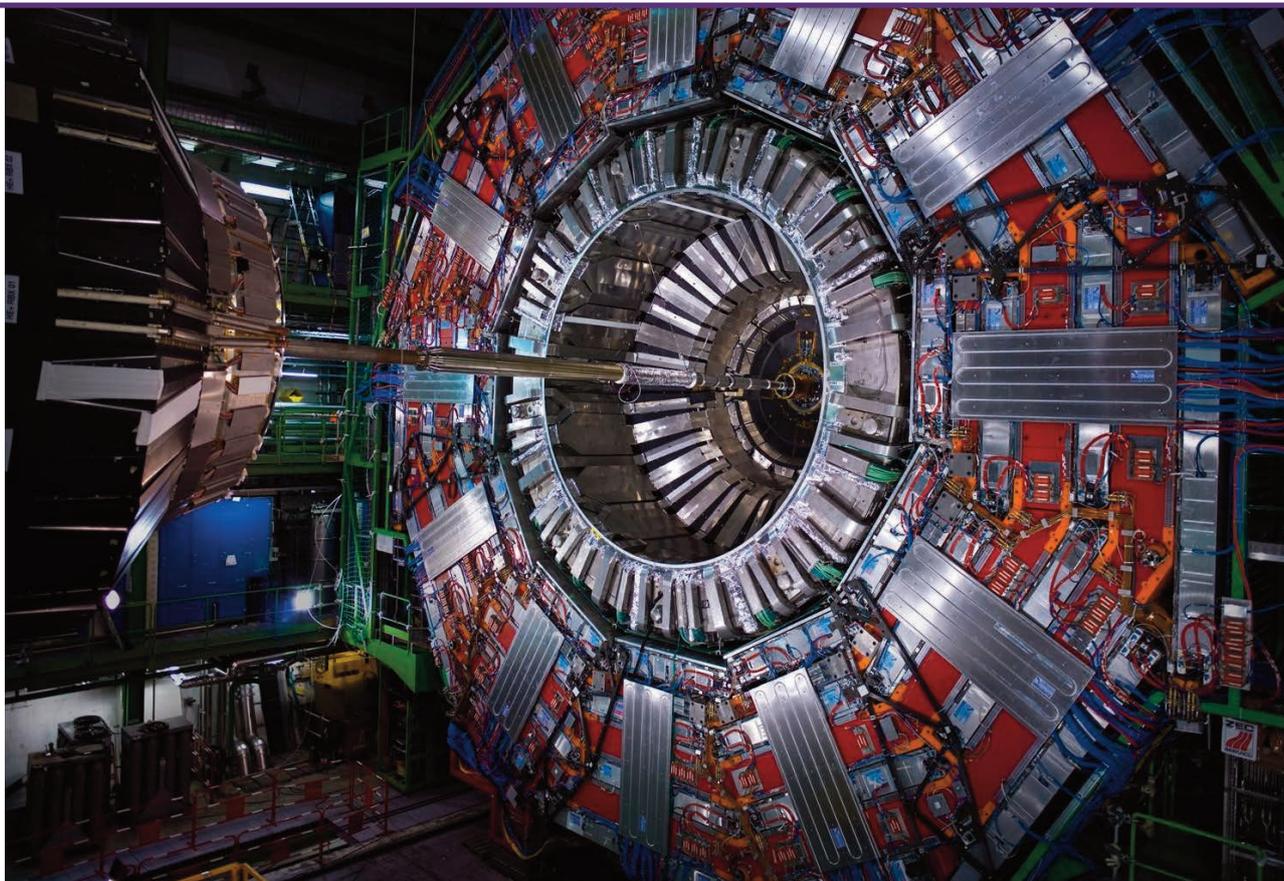


Figura 1. Dispositivo CMS (Compact Muon Solenoid; Solenoide compacto de muones) del Gran Colisionador de Hadrones (Large Hadron Collider; LHC) del CERN.

dades y los fenómenos, mientras la previsibilidad es una propiedad epistémica, es decir, relativa a la capacitación para llegar a entenderlos e interpretarlos predictivamente. Un fenómeno caótico como el clima puede ser muy difícil de prever — o incluso ser imprevisible — en la práctica, por la extrema dificultad metodológica de su estudio a partir de las causas y de los factores implicados, pero eso no significa que sea un proceso intrínsecamente incierto. La incertidumbre define el carácter íntimo de un fenómeno real, no la forma ni la complejidad a través de las que puede llegar a ser conocido por nosotros.

La brecha entre predictibilidad y comprensibilidad amenaza con volar por los aires todo el proyecto científico, como indica David Krakauer, debido a que parece que hemos llegado a un límite en el que la comprensión y la predicción se están desalineando (35). Durante los inicios de la ciencia moderna, la de Bacon y Newton, todo parecía encajar satisfactoriamente en lo que podríamos llamar un “círculo virtuoso”. Teorías rigurosas bien respaldadas por observaciones permitieron avanzar hacia lo que parecía ser la comprensión de la totalidad de la realidad, desde la mecánica cósmica hasta el electromagnetismo y la genética mendeliana.

Sin embargo, el crecimiento exponencial del “big data” está acrecentando la brecha entre predictibilidad y comprensibilidad que mencionábamos anteriormente, hasta el punto de que el vínculo

entre la comprensión y la predicción parece haber dejado de ser válido. Los modelos matemáticos, alimentados por monstruosas cantidades de datos de origen muy diverso, proporcionan predicciones cuasi-milagrosas pero, al mismo tiempo, nos ocultan la explicación hasta el punto de llevarnos a pensar que la razón llega a ser prácticamente innecesaria, siempre que los resultados obtenidos merezcan la pena.

El delicado equilibrio entre la **comprensión del significado de la realidad** (la ontología), es decir, el conocimiento profundo de la naturaleza del mundo, y la **predicción de las manifestaciones de la realidad** (la epistemología), el dominio del proceso mediante el que adquirimos conocimiento, busca optimizar nuestra comprensión actual llevándonos a apreciar nuevas características, lo que nos faculta para formular nuevas leyes fundamentales de la realidad que, a su vez, permiten hacer nuevas predicciones que deberán ser confirmadas o refutadas.

Sin embargo, en la actualidad, tal equilibrio parece estar desplazándose hacia la parte experimental y, especialmente, la computación de ingentes cantidades de datos. Es cierto que, de otra manera, seguiríamos atascados en problemas de muy difícil solución. En este sentido, podemos considerar que la mecánica cuántica vino a primar la justificación de los resultados por encima de todo, lo



que implica desarrollar una teoría nueva con el fin fundamental de justificar y predecir resultados que de otra manera “nos mantendrían atascados en el barro”.

Pero, a cambio de unos resultados extraordinariamente exactos, la mecánica cuántica nos impuso como tributo no poder intuir el mecanismo que los produce. Como reconocían algunos de sus científicos más significados, “la mecánica cuántica es fundamentalmente incomprensible” (Niels Bohr), “si crees que entiendes la mecánica cuántica, no entiendes la mecánica cuántica” (Richard Feynman), hasta llegar a la abnegada aceptación de John von Neumann de que “no entiendes la mecánica cuántica, simplemente te acostumbras”.

Idealmente la confirmación de una hipótesis científica requiere considerar los siguientes aspectos:

- Robustez explicativa del modelo sometido a ensayo; es decir, que éste disponga de sólidos fundamentos teóricos y un andamiaje lógico-matemático apropiado, así como que sea capaz de explicar la mayor cantidad y diversidad posibles de aspectos previamente observados del fenómeno objeto de estudio.
- Diseño y ejecución de experimentos no solo confirmatorios sino también *diferenciadores* que permitan obtener datos o mediciones suficientemente distintos de los obtenidos con otras hipótesis alternativas.
- Datos correctamente registrados y debidamente *contrastados*; es decir, observados, medidos y reproducidos empleando procedimientos y escalas previamente validadas, preferiblemente independientes del propio modelo testado e incluso de los propios investigadores.
- Prevenir, corregir o paliar los efectos de posibles *errores accidentales y sistemáticos (sesgos)*, hasta reducirlos a un mínimo predeterminado que sea aceptable por la comunidad científica.

Una cuestión fundamental a la que ya hemos aludido es que el marco conceptual científico es muy relevante para la medición en ciencia, porque el contexto científico estándar — eso que Thomas Kuhn llamaba *paradigma* — determina el método y, con ello, la precisión y la exactitud de las mediciones realizadas en un estudio.

Por ejemplo, el valor de la edad del universo que aceptamos actualmente como más probable está en torno a 13.770 millones de años. Sin embargo, este valor no procede de una medición directa sino que es una cantidad teórica calculada a partir del *Modelo Cosmológico Estándar*, que ha sido extrapolada hacia atrás en el tiempo hasta el “Big Bang”; es decir, es una estimación que va más allá del intervalo de observación física (19).

Con otros modelos cosmológicos no homogéneos — no copernicanos — el resultado sería muy diferente e incluso, partiendo de paradigmas estacionarios, como el defendido por Fred Hoyle, su origen se remontaría al infinito. Curiosamente, Hoyle fue el res-

ponsable de acuñar el término burlesco de “Big Bang”, con el ánimo de desprestigiar la propuesta — que hoy constituye el estándar cosmológico — de lo que se conoce como “Gran Explosión” u, originalmente, como “Átomo primigenio” (Lemaître), el punto inicial en el que se formó la materia, el espacio y el tiempo del universo, a partir de trabajos de — entre otros — Alexander Friedman (1922), Georges Lemaître (1927), Edwin Hubble (1929) y George Gamow (1948).

Por supuesto, todo lo indicado no implica — en absoluto — minusvalorar el extraordinario valor de los modelos científicos estándares actuales. Antes, al contrario, la *teoría cuántica de campos*, es decir, el formalismo teórico empleado para describir el comportamiento de las partículas elementales (electrones, fotones, quarks, etc.), ha tenido y tiene rotundos éxitos. Baste recordar que, entre sus numerosas y excelentes predicciones teóricas confirmadas experimentalmente, el valor del momento magnético del electrón es la más exacta de la historia de la ciencia, con una desviación de menos de 1 parte en cien mil millones ($< 1:10^{11}$).

Pese a ello, experimentos recientes muestran que el valor del momento magnético del muón — uno de los “hermanos pesados e inestables” del electrón, junto con el tauón — es ligera pero significativamente mayor de lo que predice la teoría (20). El muón pertenece a la segunda generación de leptones (el tauón es el correspondiente miembro de la tercera) que, como el electrón, tiene carga negativa aunque su masa es 207 veces mayor que este último y además es muy inestable: su vida media es de $2,2 \cdot 10^{-6}$ segundos, mientras que la del electrón es de — al menos — 66.000 cuatrillones ($6,6 \cdot 10^{28}$) de años, unos 5 trillones de veces mayor que el valor aceptado para la edad actual del universo.

En definitiva, ningún modelo o teoría científica — ni siquiera la prodigiosa electrodinámica cuántica (QED) — está exenta de ser sometida a una permanente revisión, en un proceso ilimitado de retroalimentación entre teoría y experimentación, conectadas a través de la medición. Esa es, precisamente, una de las mayores fortalezas de la ciencia.

6. ERRORES, SESGOS Y FALSIFICACIÓN (Y RETRACTACIÓN) DE DATOS CIENTÍFICOS

Como se indicó anteriormente, para confirmar o refutar cualquier hipótesis científica también se requiere prevenir, corregir o paliar los efectos de posibles errores accidentales y sistemáticos. Los primeros, los accidentales o aleatorios, afectan fundamentalmente a la precisión de los resultados, aunque también pueden afectar a la exactitud. El grado de *precisión* se refiere a la mayor o menor dispersión del conjunto de valores obtenidos de mediciones repetidas de una magnitud y, por tanto, cuanto menor es la dispersión mayor la precisión. Afortunadamente, hay procedimientos que



permiten evitar los errores aleatorios en muchas ocasiones o, al menos, corregirlos o compensarlos en cierto grado.

Mucho más relevantes para cualquier estudio científico son los errores sistemáticos o *sesgos*, porque reducen o incluso llegan a anular la *exactitud* de la medición, es decir, la veracidad de los resultados, su correspondencia con la realidad. A esto debe añadirse que los sesgos no afectan en muchos casos a la precisión, debido a que los resultados sesgados frecuentemente son homogéneos, lo que puede proporcionar una falsa seguridad para el investigador. En otras palabras, los sesgos representan la diferencia entre lo que creemos que estamos midiendo y lo que estamos midiendo realmente (21).

En el ámbito de la investigación clínica, los sesgos pueden producirse tanto en la fase experimental, como en la elaboración de las conclusiones y en la difusión de éstas. En concreto, podemos producir sesgos durante cada cualquiera de las fases experimentales:

- En la planificación: desde la definición de objetivos y la selección de la metodología de investigación (incluyendo el diseño, la población de muestra y los criterios de inclusión y exclusión, la definición y elección de variables, unidades y escalas, los métodos de validación e inferencia estadística, etc.
- En la ejecución: por incumplimiento sistemático de procedimientos, de calendarios u horarios de intervenciones, etc.
- En la obtención de resultados: diseño o empleo impropio de métodos y unidades de registro de datos, de análisis estadístico, de representación gráfica, etc.

Pero también son muy relevantes en investigación clínica los sesgos relacionados con una incorrecta discusión y elaboración de las conclusiones del estudio. Por ello, es muy importante acotar y discutir sistemáticamente todo aquello que los resultados experimentales de un estudio permiten (o no) afirmar o refutar, con el fin de establecer racional y éticamente los límites del conocimiento generado por el estudio. Este apartado específico de "limitaciones del estudio" es cada vez más valorado en las revistas científicas de mayor impacto y, de hecho, en muchas de ellas constituye un apartado obligatorio para cualquier artículo original. En este sentido, entre los aspectos más relevantes a considerar, podemos destacar los siguientes:

- Calibrar el grado de confianza en los resultados en función de los objetivos y métodos usados.
- Justificar qué (y qué no) se puede esperar sobre el impacto científico del estudio en su área y qué repercusión potencial puede tener sobre la ciencia y la sociedad.
- Señalar si son posibles otras explicaciones racionales alternativas a la hipótesis testada, tomando en consideración los resultados obtenidos.

- Ponderar si los resultados son generalizables o extrapolables a otras circunstancias diferentes a las empleadas.
- Valorar la representatividad de la muestra seleccionada con relación al objetivo del estudio y por qué no se seleccionaron otras posibles alternativas.
- Demostrar la idoneidad de los métodos utilizados y las mediciones realizadas, así como los motivos por los que se desestimaron otras posibles alternativas.
- Confirmar, en su caso, la pertinencia de la programación cronológica del estudio y de la toma de datos, en relación a otras alternativas posibles.
- Justificar las tasas de conformidad o tolerancia empleadas en las mediciones,
- Declarar la existencia y, en su caso, aclarar su pertinencia de intereses económicos o de otro tipo por parte de los autores.
- Y, no menos importante que todo lo anterior, es hacer un higiénico ejercicio de humildad científica y de inteligencia práctica y emocional, porque quien menosprecia su propia ignorancia, la agrava con su estupidez

Lamentablemente, comprobamos que en bastantes artículos originales publicados en revistas científicas el apartado destinado a describir la metodología empleada en el experimento tiene un papel cada vez más marginal, quedando reducido en muchos casos a la mínima expresión y, además, empleando una terminología manifiestamente imprecisa. Además, no es infrecuente el recurso a métodos y herramientas de investigación no validados previamente de forma empírica o el empleo de procedimientos inapropiados de inferencia estadística; en este sentido, las pruebas de significación de la hipótesis nula con informes de valores de p — el famoso $p < 0,05$ — están integradas en infinidad de estudios, a pesar de ser un método inferencial deficiente para la mayoría de ellos (22).

El **sesgo editorial o de publicación** es un tipo de error debido a la confianza infundada del científico en que todo el conocimiento antecedente sobre una determinada materia reside solo los estudios previamente publicados, generalmente en revistas de alto impacto. Es lo podríamos describir con la frase: "lo que no se publica no existe". En realidad, muchos artículos originales que no aparecen en revistas de amplio impacto, acaban haciéndolo en otras de divulgación más marginal, o son archivados en bases documentales de difícil acceso.

La importancia del sesgo de publicación en investigación científica es extraordinaria, ya que puede falsear la fundamentación de los resultados anteriores sobre los que se apoya un estudio, alterando de forma relevante la hipótesis sometida a investigación.

El sesgo editorial puede afectar gravemente a la elaboración de revisiones sistemáticas y metanálisis, valiosas herramien-



tas que permiten combinar y sintetizar los resultados de estudios previos y, en el ámbito biosanitario, evaluar el efecto de una intervención en salud y elaborar o actualizar guías de práctica clínica.

Con la generalización actual de los llamados repositorios o servidores de preimpresión (preprint servers) en línea, previsiblemente el impacto del sesgo editorial tiende a disminuir. Además, hay metodologías que permiten cuantificar su impacto a través de modelos de selección, gráficos en embudo (funnel plot), asimetría de las desviaciones estandarizadas, etc.

Por su parte, el **sesgo de edición** o, en palabras llanas, el sectarismo científico, consiste en alterar, retardar o negar la publicación de un artículo original por motivos extracientíficos; es decir, por prejuicios sobre autores, sus ideas originales, país o continente de origen, centro académico, empresa o institución, publicación previa o no en la revista, costes de edición (pago por publicar), pertenencia o no a determinados grupos cooperativos, idioma, etc. Cabe decir que es difícil encontrar a un solo científico que, al menos en algún momento de su carrera, no haya sido víctima de algún sesgo de edición.

Como indicábamos anteriormente, una herramienta que ha ayudado a paliar, al menos hasta cierto punto, los sesgos editoriales y de edición son las bases de datos o repositorios en línea de acceso libre y gratuito, que permiten hacer públicos artículos científicos en internet para adelantar o asegurar su difusión, antes de ser sometidos a una revisión por pares (peer review) y, en su caso, ser aceptados o no por publicaciones soportadas por empresas editoriales privadas, casi siempre previo pago de onerosas tarifas e importantes retrasos en la publicación. Por tanto, las ventajas obvias de este tipo de repositorios gratuitos y en línea son la gran rapidez de publicación y la ausencia de costes asociados; pero su desventaja principal es, por supuesto, la ausencia de una revisión científica.

El repositorio gratuito en línea más popular es **arXiv**, creado en 1989 por Paul Ginsparg como una simple lista de correo electrónico para unas pocas docenas de físicos teóricos, ha crecido hasta superar el día 3 de enero de 2022 los dos millones de artículos originales de matemáticas, física, ciencias de la computación, biología cuantitativa y estadística, a un ritmo actual de 1.200 originales al día, pero creciendo (23). Actualmente hay cientos de análogos (.Xiv) en múltiples áreas científicas y tecnológicas, soportados y patrocinados por diversas instituciones, mayoritariamente académicas previa rigurosa.

Una de las patologías o corrupciones más graves que afectan a la investigación científica es la **falsificación**, entendida como modificación o invención voluntaria de datos o mediciones. Su incidencia real es muy difícil de conocer, pero desde hace ya algunas décadas se viene denunciando de forma creciente. En este sentido, resultan especialmente reveladores los resultados de un estudio re-

alizado sobre 7.000 científicos holandeses, donde se informa que el 8% de los investigadores — y, específicamente, el 10% en el ámbito de la biomedicina — que respondieron a la encuesta — más de la mitad de los solicitados — reconocieron haber falsificado o fabricado datos y cerca del 50% confesaron haber realizado “alguna” práctica de investigación cuestionable (24).

Todo lo cual no quiere decir necesariamente que el nivel de corrupción científica sea ahora mayor que en el pasado, sino que actualmente tenemos muchos y mejores métodos para detectarla y denunciarla.

Por otro lado, se produce una **retractación** de un artículo científico cuando alguno o todos los autores de un estudio publicado, o los propios editores de la publicación, desautorizan expresamente en el mismo medio lo previamente publicado. Los motivos de retractación son muy diversos y van desde la inclusión inadvertida en el texto de graves errores científicos (diseño, muestreo, procedimientos de ejecución, análisis de datos, omisión de información, limitaciones conceptuales del estudio, etc.) hasta la comisión de un fraude u otro tipo de violación ética, tales como plagio de ideas, métodos, resultados o conclusiones de otros autores sin consignar la debida atribución; también se incluyen como motivos de retractación la falsificación de datos o fabricación “ad hoc”, la redundancia o duplicado de datos previamente publicados en la misma o en otra publicación, la proposición de conclusiones inadecuadamente soportadas por los datos obtenidos, etc.

Se ha constatado una tendencia creciente de **retractaciones** de artículos científicos en los últimos años, debido a errores, plagio o autoplagio y falsificación o fabricación “ad hoc” de datos. Gran parte de las retractaciones se producen antes de los dos primeros años desde la publicación y suelen ser solicitadas por alguno o todos los autores del original; sin embargo, a pesar de que las retractaciones se producen en los mismos medios donde fueron publicados los originales, gran parte de ellos fueron citados incorrectamente en otros originales (5).

En este sentido, la práctica de validar los estudios previamente publicados no es tan común en biomedicina como lo es en las ciencias naturales, lo que ha ido mermando la credibilidad de la investigación clínica, como ha sido puesto de manifiesto por algunos especialistas en edición científica (25).

Retraction Watch es una base de datos y sitio web, ambos organizados y administrados por profesionales de la información científica, sin ánimo de lucro y actualmente financiado por sus propios lectores, dedicado a detectar, recopilar y publicar los casos de retractaciones. La misión del *Center for Scientific Integrity*, la organización matriz de *Retraction Watch*, es promover la transparencia y la integridad en la ciencia y las publicaciones científicas, y difundir las mejores prácticas y aumentar la eficiencia en la ciencia. Para

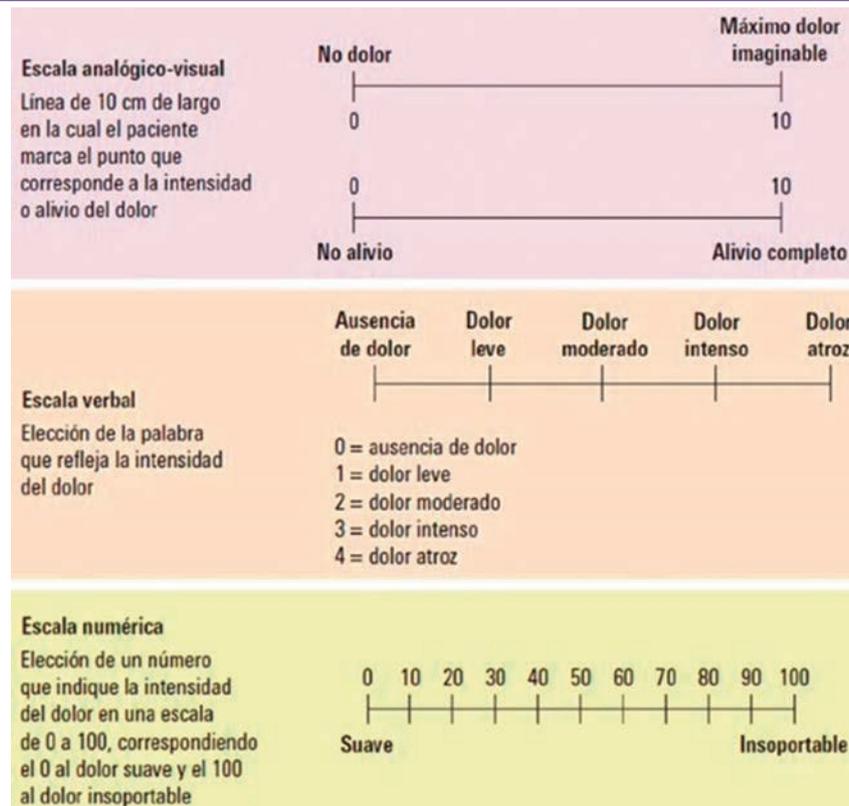


Figura 2. Escalas sobre autovaloración del dolor.

esta organización, retirar un original o retractarse significa lo mismo, porque considera que las revistas que lo han retirado generalmente lo hacen para no justificar por qué se retracta un artículo y, a veces, para hacer que un artículo desaparezca sin dejar rastro. Su base de datos está disponible gratuitamente para científicos, académicos y cualquier otra persona interesada en analizar la información (26).

7. PARTICULARIDADES DE LA MEDICIÓN EN LA INVESTIGACIÓN CLÍNICA: ¿OBJETIVAR LO SUBJETIVO?

El gran reto de la investigación clínica — aquello que la hace particularmente problemática desde una perspectiva científica — es tratar de objetivar fenómenos que, en muchas ocasiones, son esencialmente subjetivos — como el dolor o la depresión — además de otros que, estos sí, pueden ser más fácilmente parametrizados.

El motivo de esta dificultad adicional es obvio: **el sujeto de la investigación es el ser humano**, considerado tanto individual como colectivamente, lo que implica la condición de persona pensante (que es consciente de su intervención en un estudio), emocional (que responde a partir de estados mentales o de actitudes), moral (que se rige por principios que supeditan su comportamiento personal) y social (que es consciente y partícipe de su relación con otras personas y colectivos). Todo ello, entre otros factores, multiplica

la incertidumbre de las respuestas y obliga a extremar las exigencias metodológicas y éticas del científico responsable del estudio clínico. Si bien se puede lograr una medición precisa en el laboratorio de parámetros bioquímicos, previamente estos han debido ser correlacionados de forma biunívoca y preciso con algún elemento de la salud humana.

¿Qué es bienestar? ¿Según qué estándares podemos llamar a una población saludable?, se pregunta Sean Valles, mostrando la preocupación por la inconstancia y la contingencia de la medición, que deja mucho espacio para que los valores y las motivaciones de los profesionales den forma — para bien o para mal — a los resultados del proceso de medición (27).

Por otro lado, es evidente que desconocemos muchas de las variables que podrían afectar a los resultados de un experimento clínico y, esperamos — uno poco “mágicamente” — que la aleatorización de la muestra distribuya de forma equitativa los posibles “factores de confusión”, es decir, los numerosos aspectos no considerados en el diseño del estudio clínico, como la existencia de hábitos alimentarios personales desconocidos — entre otras muchas circunstancias posibles — que podrían afectar significativamente al metabolismo de un fármaco.

Es cierto que se han desarrollado multitud de escalas sobre autovaloración del dolor (como los que se muestran en la figura 2) y de cuestionarios que integran y ponderan diferentes as-



pectos afectivos, sensoriales, etc., transformándolos en cifras abstractas e intervalos numéricos discretos. ¿Pero cómo interpretar en cada paciente si, por ejemplo, una unidad es o no equiparable a lo largo de toda la escala considerada? ¿Representa la variación de 10 a 20 el mismo grado de magnitud del cambio fisiopatológico que el descrito con la variación de 70 a 80? ¿Esa valoración personal es siempre la misma, con independencia de la intervención terapéutica?

¿Es objetivable el nivel de comprensión de cada paciente sobre lo que le pregunta el observador clínico y la capacidad del paciente para expresar lo él que experimenta en cada momento? ¿Cómo interpretar resultados que proceden de combinar datos subjetivos de múltiples pacientes? ¿Es la aleatorización de grandes series de pacientes una respuesta científicamente rigurosa? ¿Tiene sentido hablar de "paciente promedio"? En resumen: ¿Hasta qué punto — si es que lo hay — podemos objetivar (es decir, cuantificar sistemáticamente) sensaciones e interpretaciones subjetivas, como el dolor, la pérdida de motivación o una alucinación visual?

Por otro lado y por si fuera poco, la respuesta individual de cada persona sometido a una intervención clínica está mediada por múltiples factores no relacionados con los efectos objetivos propios de la intervención; factores que, por supuesto, deben ser debidamente ponderados, entre otros (28):

- Los omnipresentes y poderosos efectos placebo y nocebo.
- El empleo incontrolado de otros tratamientos, adyuvantes o no.
- Los múltiples factores que conforman "el estilo de vida" y "el entorno" de cada paciente, pero que no son considerados — o declarados — en la población seleccionada — o excluida — para el estudio.
- La constitución genética completa y no solo la presencia, ausencia o mutación de determinados genes considerados "ad hoc" por el investigador.
- La actitud psicológica del paciente en relación con la intervención y con el observador clínico, entendida como predisposición, resistencia, desconfianza, esperanza, etc.
- El nivel de comprensión de cada paciente sobre lo que le pregunta el observador clínico y la capacidad del paciente para expresar lo él que experimenta en cada momento
- La existencia real de procesos espontáneos e inesperados de remisión de la enfermedad.
- Diversos fenómenos sociológicos y sesgos estadísticos asociados al propio ensayo, como se indica a continuación.

Desde hace varias décadas, un aspecto fundamental de la investigación de nuevos medicamentos es mostrar que el medicamento experimental es más efectivo que el placebo, para demostrar objetivamente su eficacia. En este sentido, el primer ensayo clínico

controlado con placebo y publicado se realizó en el año 1931, en el que se estudiaba el aurotiosulfato (Sanocrysin®) como tratamiento de la tuberculosis, utilizándose agua destilada como placebo (29).

Desde entonces se han desarrollado estrategias especiales para predecir el comportamiento de los pacientes en relación al placebo, con el fin de excluirlos antes de que sean admitidos a un ensayo clínico aleatorizado y controlado con placebo o para eliminar o reducir los efectos más relevantes de este último, especialmente con el fin de evitar los que se conoce como "paradoja de la eficacia" (*efficacy paradox*). Esta paradoja implica que los grupos (brazos) de tratamiento con placebo del estudio podrían llegar a experimentar un efecto mayor que el tratamiento supuestamente activo o basado en la evidencia. Por otro lado, es importante considerar que tasas elevadas de respuesta al placebo obstaculizan los esfuerzos para detectar la eficacia real de los nuevos fármacos, contribuyen a que los resultados de los ensayos resulten fallidos o ambiguos y, en última instancia, retrasan la incorporación de nuevos tratamientos a la clínica (30).

Un fenómeno particularmente interesante es el efecto positivo que se observa en la conducta de las personas participantes en un ensayo clínico, aparentemente debido al simple hecho de saberse observadas. Este efecto es bien conocido e incluso utilizado en psicología industrial.

Un ejemplo paradigmático de ello es el experimento realizado en los años 20 del siglo pasado, que describió lo que acabó denominándose como *efecto Hawthorne*. El nombre hace referencia a un pueblo próximo a Chicago, en Estados Unidos, en el que se ubicaba la *Western Electric Company*, donde se realizó un conjunto de experimentos psicológicos relacionados con los condicionantes de la productividad laboral de los más de 20.000 empleados de la compañía. En concreto, se observaron mejoras en el rendimiento laboral tras la implementación de cambios en el sistema de producción, pero con la peculiaridad de que esta mejora podía observarse tanto si los cambios eran objetivamente positivos para el proceso de producción, como si no lo eran. Tras varios años de observación sistemática, la conclusión fue que una parte relevante del beneficio observado se debía al hecho de que los trabajadores se sabían observados y no al efecto propio relacionado con la — supuesta — mejora implementada en el proceso productivo (31).

Otro interesante fenómeno es el denominado *regresión a la media*, que refleja el hecho de que, en la medición de una variable clínica de naturaleza continua en un ensayo clínico, los individuos que se encuentran más alejados de los valores medios de la variable tienen más probabilidades de obtener, en mediciones posteriores, un resultado que tiende a converger con el valor medio de la población estudiada. El fenómeno de regresión a la media es mucho más común en estudios con enfermos de patologías de ca-



rácter fluctuante, como la artritis reumatoide o la colitis ulcerosa, en las que el paciente solicita atención médica — o acepta participar en un ensayo clínico — cuando los síntomas son más graves, como ocurre en un brote agudo, por lo que en observaciones posteriores, cuando el paciente se siente observado y los síntomas han remitido o disminuido, la estimación de la gravedad del cuadro patológico es claramente inferior (32).

El *efecto Rosenthal*, también conocido como *efecto Pigmalión* o *efecto esperado*, se manifiesta en contextos terapéuticos, donde tanto los clínicos como los evaluadores, con el paso del tiempo, pueden dar menor importancia a los síntomas comunicados porque esperan que los pacientes mejoren con el tiempo, lo cual podría conducir a una falsa impresión de mejoría. Este efecto también influye en la interacción entre clínico y paciente, dando como resultado un mayor efecto placebo o psicoterapéutico y, en cualquier caso, una verdadera mejoría.

Por su parte, el *efecto halo* implica que la mejora en un síntoma da lugar a expresiones de optimismo y bienestar por el paciente, que reducen el impacto negativo de otros síntomas, incluso si estos últimos realmente no han mejorado.

Debido a las notables y abundantes dificultades para los parametrizar fenómenos subjetivos de interés clínico, es fundamental, definir y justificar previamente la magnitud mínima del cambio que puede producir un efecto clínicamente relevante para el paciente y para el observador (*minimum clinically important difference*, MCID), para poder medir correctamente el efecto de una intervención en cada estudio. Como indicaba Kathy Liu en relación con los ensayos clínicos para testar fármacos en pacientes con enfermedad de Alzheimer, “sin MCID, los patrocinadores podrían encontrar motivos para potenciar los ensayos para detectar significación estadística solo para efectos pequeños y potencialmente intrascendentes en los resultados clínicos. Las MCID benefician a pacientes, familiares, cuidadores y a los sistemas de atención sanitaria, por lo que deben incorporarse en los ensayos clínicos y en la guía de desarrollo de fármacos para la enfermedad de Alzheimer” (33).

8. CONCLUSIONES

- La medición es un integrante esencial de la ciencia; uno de sus sellos distintivos, fuente primaria de conocimiento y el imprescindible nexo de unión entre la teoría y la realidad, permitiendo que las hipótesis y los experimentos se retroalimenten y mejoren continuamente.

- En el estudio científico de cualquier fenómeno, la intervención del observador y el propio hecho de la observación perturba al fenómeno estudiado. Por ello, siempre es preciso tomar en consideración el grado de perturbación que produce, algo que

en investigación clínica puede ser particularmente complejo por la multiplicidad y variabilidad de las manifestaciones humanas.

- El ámbito científico es cada vez es más multidisciplinar, interdisciplinar y transdisciplinar, lo que implica el empleo de métodos, términos y unidades no siempre son coherentes ni directamente inteligibles entre ellos.

- La predictibilidad de un fenómeno no presupone necesariamente su comprensibilidad. Nada impide — en principio — que dos o más hipótesis científicas distintas puedan ser coherentes con un mismo conjunto de datos u observaciones, por lo que es prioritario diseñar experimentos con mediciones para obtener datos que puedan ser suficientemente diferentes con cada modelo.

- Los sesgos reducen la veracidad y perturban la divulgación de los resultados de un experimento; y en muchos casos lo hacen sin afectar a la precisión de los datos, lo que puede atribuirles una falsa fiabilidad.

- Uno de los grandes retos de la investigación clínica es objetivar fenómenos que son esencialmente subjetivos, debido a la complejidad del sujeto observado (ser humano), por su condición de ente pensante, emocional, moral y social, lo que multiplica la incertidumbre de sus respuestas tanto individuales como colectivas, y extrema las exigencias metodológicas y éticas del científico.

- La respuesta de cada individuo a una intervención clínica incorpora múltiples factores no relacionados con los efectos objetivamente achacables a la intervención, pero que deben ser debidamente ponderados en cualquier caso.

- Para poder medir científicamente el efecto de una intervención en un ensayo clínico, es fundamental definir y justificar previamente la magnitud mínima del cambio que puede dar lugar a un efecto que sea clínicamente relevante para el paciente y para el observador.

9. REFERENCIAS

1. Bunge M. La ciencia. Su método y su filosofía. Editorial Laetoli, 2013. ISBN 9788492422593
2. Cuéllar Rodríguez S. Some thoughts on science. An Real Acad Farm. 2015; 81(3): 230-238
3. Brooks RA, Mataric MJ. Real Robots, Real Learning Problems. En “Robot Learning”; Jonathan H. Connell, Sridhar Mahadevan, eds. Springer, Boston (USA); pp.193-213. DOI: 10.1007/978-1-4615-3184-5 (1993)
4. Popper KR. Conjectures and Refutations: The Growth of Scientific Knowledge (Routledge Classics). ISBN: 9780415285940.
5. Cuéllar Rodríguez S. Retracción y normalización de la investigación científica de la COVID-19. Panorama Actual Med. 2021; 45(444): 595-601.
6. Korzybski A. A Non-Aristotelian System and its Necessity for Rigour



- in Mathematics and Physics. Artículo presentado ante la American Mathematical Society, en New Orleans, Louisiana. American Association for the Advancement of Science, December 28, 1931.
7. ICONO: Observatorio Español de I+D+i. Percepción Social de la Ciencia y la Tecnología en España. Fundación Española para la ciencia y la tecnología (FECYT). Ministerio de Ciencia e Innovación. Disponible en: <https://icono.fecyt.es/principales-indicadores/produccion-cientifica> (consultado el 23 de diciembre de 2021).
 8. Moyer M. ¿Es digital el espacio? *Investigación y Ciencia*. 2012; abril, pp. 16-23.
 9. Penrose R. El camino de la realidad. Una guía completa de las leyes del universo. Editorial Random House Mondadori, 2006. ISBN: 9788483066812
 10. Díaz G. Esbozo de una filosofía zubiriana de la matemática. Actas del II Congreso de la Sociedad de Lógica, Metodología y Filosofía de la Ciencia en España. Bellaterra, 6-8 de febrero de 1997, pp. 141-5.
 11. Cuéllar Rodríguez S. Productividad científica y promiscuidad editorial. *Panorama Actual Med*. 2020; 44(438): 1225-1233
 12. Ioannidis JPA. Rethink Funding. *Sci Am*. 2018; 319(4): 52-5. Doi: 10.1038/scientificamerican1018-52
 13. Errington TM, Denis A, Perfito N, Iorns E, Nosek BA. Reproducibility in Cancer Biology: Challenges for assessing replicability in preclinical cancer biology. *eLife* 2021;10:e67995 DOI: 10.7554/eLife.67995
 14. Real Academia Española (RAE). Diccionario de la lengua española. 23ª edición, octubre 2013. Versión en línea: <https://dle.rae.es/>
 15. Cuéllar Rodríguez S. Medición y ciencia. *Panorama Actual Med*. 2021; 45(447): 1034-1039
 16. Ball P. Beyond the bond. *Nature*. 2011; 469: 26-8. DOI: 10.1038/469026a
 17. Wolchover N. What Is a Particle? *Quantamagazine*. Noviembre 12, 2020. <https://www.quantamagazine.org/what-is-a-particle-20201112/>
 18. Baggott J. Calculate but don't shut up. <https://aeon.co/essays/shut-up-and-calculate-does-a-disservice-to-quantum-mechanics?> (diciembre 2021)
 19. Uzan JP. ¿Hay que revisar el modelo cosmológico? *Investigación y Ciencia*. Noviembre 2021
 20. Fermilab. Muon g-2. April 27, 2021; <https://muon-g-2.fnal.gov/>
 21. Cuéllar Rodríguez R. Sesgos de publicación y edición científica: la metaciencia al rescate. *Panorama Actual Med*. 2020; 44(430): 37-47.
 22. Ioannidis JPA. Massive citations to misleading methods and research tools: Matthew effect, quotation error and citation copying. *Eur J Epidemiol*. 2018; 33(11): 1021-3. DOI: 10.1007/s10654-018-0449-x.
 23. Garisto D. ArXiv.org reaches a milestone and a reckoning. *Scientific American*. January 10, 2022; <https://www.scientificamerican.com/article/arxiv-org-reaches-a-milestone-and-a-reckoning/>
 24. Chawla DS. 8% of researchers in Dutch survey have falsified or fabricated data. *Nature*. 2021 Jul 22. doi: 10.1038/d41586-021-02035-2
 25. González-Rayas JM, Rayas-Gómez AL, González-Yáñez JM, García-González JJ, Hernández-Hernández JA, López-Sánchez RC. La nueva normalidad en la investigación médica: replicabilidad y reproducibilidad. *Rev Biomed*. 2020; 31(3): 108-10
 26. Retraction Watch. Disponible en: <https://retractionwatch.com/>
 27. Valles S. Philosophy of Biomedicine. *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/sum2020/entries/biomedicine/>
 28. Cuéllar Rodríguez S. Metascience: reflections on the reproducibility of research in pharmacology. *An Real Acad Farm*. 2019; 85(1): 4-11
 29. Amberson JB, McMahon BT, Pinner M. A clinical trial of Sanocrysin in pulmonary tuberculosis. *Am Rev Tuberc*. 1931; 24: 401-35.
 30. McQueen D, Cohen S, St John-Smith P, Rampes H. Rethinking placebo in psychiatry: How and why placebo effects occur. *Advances in Psychiatric Treatment*. 2013; 19(3): 171-180. doi:10.1192/apt.bp.112.010405
 31. Sedgwick P, Greenwood N. Understanding the Hawthorne effect. *BMJ*. 2015; 351: h4672. doi: 10.1136/bmj.h4672
 32. Whitney CW, Von Korff M. Regression to the mean in treated versus untreated chronic pain. *Pain*. 1992; 50(3): 281-5.
 33. Liu KY, Schneider LS, Howard R. The need to show minimum clinically important differences in Alzheimer's disease trials. *Lancet Psychiatry*. 2021 Nov; 8(11): 1013-1016. doi: 10.1016/S2215-0366(21)00197-8.
 34. Heidegger, Martin. *Filosofía, ciencia y técnica*. Santiago de Chile, Editorial Universitaria (1997). ISBN 9561113112
 35. Krakauer DC. At the limits of thought. *Aeon Magazine*; 4 feb 2022. <https://aeon.co/essays/will-brains-or-algorithms-rule-the-kingdom-of-science?>

Si desea citar nuestro artículo:
Epistemología de la medición

Santiago Cuéllar Rodríguez
An Real Acad Farm [Internet].

An. Real Acad. Farm. Vol. 88. nº 1 (2022) · pp 31-44
DOI: <http://dx.doi.org/10.53519/analesranf.2022.88.01.02>