

Epistemología

filosofía de la mente y bioética

La dureza de la ciencia dura: ¿mito o realidad? Elementos para una discusión

Luis Felipe Orozco-Cabal¹

El pensamiento no es letra muerta. Ningún argumento ni verdad poseen valor si no convencen al sujeto que obra; si la voluntad de ese último, esclarecida por la razón, no lo conduce a vivir iluminado por la verdad; por ello la voluntad de saber que moviliza todo el proceso del pensar termina por la aceptación voluntaria de la verdad. (1)

Resumen

Objetivo: Discutir aspectos metodológicos relacionados con la calidad de las investigaciones preclínicas y la aplicabilidad de principios de investigación clínica y medicina basada en la evidencia para evaluar la validez y calidad de los estudios preclínicos. *Método:* Análisis crítico de publicaciones selectas. *Resultados:* El ejercicio de la investigación (preclínica y clínica) exige indicadores rigurosos que permitan determinar la validez y generalización de sus hallazgos. Es deseable el desarrollo de metodologías para determinar su peso y hacer posible su revisión sistemática. *Conclusiones:* Se exige un replanteamiento del quehacer científico, a la luz del reconocimiento de sus limitaciones, la urgencia de sus aportes y la responsabilidad que esta tiene frente al bienestar de la humanidad.

Palabras clave: metodología, medicina basada en evidencia, investigación.

Title: The Hardness of Hard Science: Truth or Myth?

Abstract

Objective: To discuss methodological issues concerning the quality and applicability of clinical research and evidence-based medicine principles to pre-clinical research. *Methods:* Critical analysis of selected writings. *Results:* Rigorous standards to determine the quality and generalization of pre-clinical findings are needed; methods to systematically review the evidence and its weight have not been developed for pre-clinical research. *Conclusions:* Never before has a debate about the limitations of pre-clinical research been more relevant, given the momentousness of its contributions and its obligation towards the well-being of humankind.

Key words: Methodology, evidence-based medicine, research.

¹ MD, PhD. Doctor en Neurociencias, Residente de Psiquiatría General. Departamento de Psiquiatría y Salud Mental, Facultad de Medicina, Pontificia Universidad Javeriana, Bogotá, Colombia.

La verdad en las ciencias

La ciencia, en su concepción moderna, es tanto un proceso por el cual se adquieren conocimientos, como, también, el cuerpo de conocimientos resultantes de este proceso. Tiene como fin último la búsqueda de la verdad en cuanto realidad accesible al ser humano, y en el caso de la ciencia con enfoque positivo, nace luego del reconocimiento de la posibilidad de descubrir lo que las cosas son (verdad) en sí mismas, partiendo de la observación y por medio del estudio metódico y sistemático de las cosas mismas (2). Se trató de un esfuerzo contrario al enfoque metafísico clásico heredado de los griegos. En la ciencia moderna, el proceso de adquisición sistemática de conocimiento de un fenómeno determinado se denomina el *método científico*, y el sistema que busca “explicar” es generalmente la *naturaleza*, entendida como el sistema de leyes que la rigen.

Dado que el método científico se basa fundamentalmente en la observación sistemática, la verdad derivada de este método está circunscrita a las condiciones de observación, descripción y verificación de los fenómenos en estudio. Luego, más que verdades absolutas, la ciencia permite extraer principios generales con un carácter universal relativo acerca del mundo natural, con el propósito de realizar predicciones con grados de certeza proporcionalmente relativos acerca de los fenómenos estudiados (3).

Por definición, estos principios generales que constituyen teorías y luego “paradigmas”, según los cuales se explica el mundo natural, están sujetos al escrutinio permanente de las disciplinas y pueden o no ser confirmados, a medida que la capacidad de observación se agudiza (4). En este orden de ideas, la verdad es siempre relativa a las condiciones de observación y a la evolución dinámica de los paradigmas que rigen su articulación teórica; está sujeta a verificación y se fundamenta en la reproducibilidad de las observaciones bajo múltiples condiciones de registro o descripción (3).

Pero si bien la búsqueda de la verdad sigue siendo el motivo más importante del descubrimiento científico (3), su ejercicio, como producto humano, es vulnerable a la presión de la incertidumbre y la variabilidad y el error. El reconocimiento de esta vulnerabilidad en el método científico y en el corpus científico y la necesidad de establecer la validez de las observaciones científicas fueron los principales impulsores de la estadística durante el siglo XVIII (5). Hoy parece irrefutable que en toda observación, por más “científica” que sea, existe cierto grado de incertidumbre y posibilidad de error, lo que hace necesaria la utilización de metodologías estandarizadas y herramientas, como la estadística, para cuantificar ambos fenómenos, en aras a restringir el alcance de las interpretaciones que se derivan de

nuestro quehacer científico, acercándonos a la verdad.

La medicina basada en la evidencia: un cambio de paradigma

La práctica médica ha sufrido una profunda transformación en sus paradigmas con el surgimiento de la *medicina basada en la evidencia* (MBE) durante el último siglo. Este nuevo paradigma promulga el escrutinio de la evidencia derivada de la investigación clínica bajo estándares rigurosos, realiza una jerarquización de la evidencia y alerta acerca de los errores derivados de una práctica médica basada en la autoridad individual o experticia. Adicionalmente, defiende la interpretación de los hallazgos a la luz de la relación riesgo/beneficio, con base en las características de los pacientes (6,7). Luego, no es sorprendente que la MBE haya tenido un profundo impacto en la práctica médica, en el diseño de los estudios clínicos y en la interpretación de las observaciones.

Los parangones de la práctica médica se han modificado sustancialmente y muchos de los mitos que guiaban el ejercicio de la medicina han sido desacreditados. Una de las disciplinas que más se ha beneficiado de este cambio de paradigma es la psiquiatría y, más específicamente, la terapéutica de los trastornos psiquiátricos. Basta con leer los reportes recientes relacionados con los conflictos de

interés y el sesgo de publicación como fuentes de confusión en la evaluación de la efectividad de algunos grupos de medicamentos ampliamente utilizados en nuestra práctica, como los antidepresivos y moduladores del afecto (8-11).

Lo anterior ha producido una desconfianza profunda en la literatura clínica y ha llevado al escrutinio cuidadoso de la investigación clínica. Esto, sin demeritar el conocimiento derivado de las descripciones de casos individuales o la investigación cualitativa que hacen parte fundamental del corpus teórico de la psiquiatría, que, por su carácter particular, obliga a recurrir a este tipo de aproximaciones, aunque parezcan resistir un análisis de tipo *categorial* y “científico” (12).

La necesidad de realizar un escrutinio de los resultados de las investigaciones clínicas se ha visto fortalecida por factores adicionales, como el bajo poder de los estudios, la dificultad para encontrar asociaciones replicables a lo largo del tiempo y las inconsistencias entre los avances de las investigaciones preclínicas y la prueba de hipótesis basadas en esos avances en el campo clínico. Autores como Ioannidis (13) han aludido a errores estadísticos, entre otros, como razones para estos fenómenos. Otros, como Willett (14), han llamado la atención hacia la calidad de la evidencia preclínica como fuente de la heterogeneidad e inconsistencia entre los hallazgos derivados de

los experimentos clínicos. Ésta es una sugerencia interesante, ya que debemos recordar cómo la psiquiatría, así como otras especialidades médicas, fundamenta gran parte de su quehacer en los resultados de las investigaciones preclínicas en neurociencias y disciplinas afines.

Si las investigaciones preclínicas son de baja calidad, las hipótesis derivadas de éstas para ser probadas en el ámbito clínico también lo son. Esta posibilidad obliga a examinar si el proceso de investigación en el ámbito preclínico está sujeto a una estrategia de aseguramiento de la calidad, similar a la propuesta por la MBE en el campo clínico, que disminuya la “mala ciencia”, lo que permitiría un incremento en la capacidad de transferir sus resultados. Esta pregunta es especialmente relevante si consideramos la razón entre el número de publicaciones preclínicas y las clínicas, tanto en psiquiatría y salud mental como en otras disciplinas, ampliamente en favor de la primera.

Para ilustrar la desigualdad de esta proporción tomaré como ejemplo el número de estudios publicados en el 2008 acerca del estrés, un fenómeno relacionado con la génesis de la enfermedad mental. En el 2008 ingresaron en MedLine 49.612 registros de artículos de investigación primarios relacionados con estrés realizados en animales o sistemas celulares, comparado con tan sólo 5.887 registros de investigaciones primarias de corte clínico o en hu-

manos (cálculos propios). Luego, las cifras hablan por sí mismas; el número de reportes primarios de literatura preclínica supera en una razón de diez o más el número de reportes clínicos. La desigualdad de esta proporción sugeriría que la investigación preclínica es una fuente inagotable de hipótesis para probar en el campo clínico. Sin embargo, esto parece ser un supuesto inválido, especialmente en algunos campos de interés esencial, como las enfermedades neurodegenerativas, las adicciones, los trastornos del afecto y los trastornos del desarrollo.

Numerosos factores pueden ser responsables del limitado potencial traslacional del conocimiento preclínico. La complejidad del sistema nervioso y de los mecanismos que subyacen a su funcionamiento hace difícil el establecimiento de relaciones causales que disminuyan la distancia entre neurofisiología y comportamiento. Los avances en la comprensión de la fisiología cerebral van de la mano con el desarrollo de tecnología que permita el estudio de los fenómenos neuronales; el último siglo ha visto grandes avances al respecto (i. e. patch-clamp, imagenología de calcio, expresión de receptores acoplados a opsinas, entre otros).

La especificidad regional, celular y el estado-dependiente de muchos de los sistemas de regulación de la función cerebral limitan nuestra capacidad para describir mecanismos canónicos acerca de éste, lo cual favorece la especialización del conoci-

miento y la particularización de éste. Los fenómenos mentales humanos son difíciles de replicar en el ámbito preclínico, lo cual limita la validez de constructo y predecibilidad de los modelos animales. Aunque similares, la conectividad del cerebro humano impone ámbitos de procesamiento de información que no existen en animales experimentales, como roedores o insectos; luego hay un nivel de complejidad que simplemente no es reproducible en el laboratorio. Por eso, la tendencia a la búsqueda de sistemas neurales simples, conocidos y fáciles de manipular y validar en el ámbito clínico. Una historia exitosa de este tipo de aproximaciones es la carrera de Erick Kandel y otros grupos de investigadores interesados en los mecanismos biológicos de la memoria (15).

Pero, más allá de las limitaciones más fundamentales impuestas por la naturaleza misma del cerebro humano, se encuentran aquéllas más asequibles a nuestro entendimiento, como las relacionadas con la calidad y la interpretación crítica de resultados de la investigación preclínica.

La calidad de la evidencia preclínica y del quehacer científico

En general, la validez de los hallazgos preclínicos se fundamenta en las condiciones controladas del ambiente experimental, el rigor del diseño experimental, la metodolo-

gía de medición o exploración, el análisis de datos y su cuidadosa interpretación. Su validación, especialmente en Estados Unidos y el Reino Unido, se realiza por medio de un sistema de revisión por pares, como casi todas las actividades académicas, y debe estar a disposición de ésta (16). Específicamente, el proceso de revisión por pares tiene como funciones principales: la diseminación responsable de los resultados de las investigaciones, crear y preservar el conocimiento canónico de las disciplinas, asegurar la calidad de las investigaciones, y la asignación y aseguramiento de los derechos de autor (17).

Estos mecanismos de aseguramiento de la calidad de la evidencia y de escrutinio por parte de la comunidad son tan antiguos como las comunidades académicas mismas y han sostenido múltiples modificaciones a lo largo del tiempo, pero están lejos de ser perfectos. La pobre *predictibilidad* de la evidencia preclínica acerca de asociaciones epidemiológicas relevantes sugiere que ni la tecnología, la rigurosidad de método o el proceso de revisión por pares son suficientes para asegurar la calidad de la investigación y la relevancia para la clínica de los resultados de la investigación preclínica. Tal es el caso de los estudios sobre el potencial carcinogénico de la sacarina, o la pobre validez del constructo de los modelos animales de depresión o la disparidad entre el número de estudios relacionados

con la enfermedad cerebrovascular que contrasta con el reducido número de alternativas terapéuticas en uso (18-21).

En el campo del estrés y patologías mentales asociadas con el estrés también existe una importante discrepancia entre los avances de estudio animales sobre las acciones del factor liberador de corticotropina como modulador de la transmisión sináptica en el sistema nervioso y la virtual ausencia de estudios de fase III con análogos de esta molécula (22).

Metodológicamente existen grandes disparidades en el aseguramiento de la calidad de las investigaciones preclínicas que realizamos. Pound y colaboradores (23), respecto a los estudios en animales relacionados con la identificación de efectos terapéuticos de nuevas sustancias, han señalado algunas deficiencias metodológicas y conceptuales de estos estudios: la marcada heterogeneidad intra- e interespecies, las diferencias en el metabolismo de los animales que producen variación en la eficacia y toxicidad de los compuestos; heterogeneidad entre los modelos de enfermedad y su validez de constructo; variaciones en la dosificación de los compuestos experimentales y regímenes de administración, y su relevancia con los esquemas de dosificación en humanos; la heterogeneidad de los métodos de aleatorización y brazos de los estudios; muestras insuficientes y estudios

de poder cuestionable, la falta de control para factores de confusión y la ausencia de la intención por tratar en los análisis; el empleo de pruebas estadísticas inadecuadas para el diseño experimental utilizado; diferencias o incapacidad de establecer ciegos en el desarrollo de los experimentos; selección de variables de desenlace inadecuadas y su relevancia con los constructos humanos que intentan cuantificar; tiempos de seguimiento variables e inadecuados, y su pobre correspondencia con el curso natural de los fenómenos en humanos. Lo anterior desacredita la afirmación de “*before testing a new treatment in man, test it first in animals if possible*” (24).

Adicionalmente, es frecuente observar la publicación de revisiones no sistemáticas de la literatura preclínica en todas las disciplinas. Con frecuencia se indica en la sección de métodos: “Se realizó una revisión de la literatura”. No existe una metodología rigurosa para la realización de revisiones de este tipo de hallazgos, lo que favorece la manipulación de la información, el establecimiento de falsas asociaciones cuya magnitud no se cuantifica y la libre interpretación de la evidencia. En un estudio por Roberts y colaboradores (25) se demostró que sólo 1 de 10.000 artículos de investigación en animales calificaba como meta-análisis frente a 1 en 1.000 para la literatura clínica.

Lo discutido hasta aquí sugiere que es hora de parar, pensar y ac-

tuar respecto a nuestro quehacer científico. Especialmente en un país donde las ciencias biomédicas están en sus albores (26). ¿Por qué no tomar las lecciones aprendidas de la epidemiología y la MBE y ver su aplicabilidad al servicio del escrutinio de la evidencia preclínica?

Consideraciones acerca de la aplicabilidad de los principios para el análisis de los datos epidemiológicos a las investigaciones preclínicas

En el 2001, Abramson y Abramson (27) proponen que el análisis de la validez de los datos epidemiológicos implica la evaluación sistemática de los errores sistemáticos (sesgos), seguida de la evaluación del error aleatorio (azar) y de la causalidad en las asociaciones.

La primera fase de análisis, según explica Ghaemi (5), busca determinar la presencia de posibles sesgos y variables de confusión en los estudios. Una de las principales consideraciones en los estudios clínicos para controlar sesgos de selección y el control de factores de confusión es que no existan diferencias significativas entre los sujetos participantes en el estudio. En general, esto se logra con la determinación adecuada de la población de estudio, los controles y la aleatorización.

En el caso de los estudios en animales, la homogeneidad en la muestra de estudio se trata de asegurar

por medio de múltiples procesos. Por ejemplo, las camadas de roedores o de animales experimentales se compran a un mismo proveedor, quien asegura la identidad genética de las poblaciones. Estos sujetos se exponen a las mismas condiciones de transporte, proceso de aclimatación, de hotelería, ciclos luz/oscuridad, alimentación, control de ruido, temperatura, humedad y manipulación. No obstante, investigaciones recientes sugieren que aun con la igualdad de las condiciones antes descritas, se observan diferencias características entre los miembros de una misma camada, que pueden representar un factor de confusión en los experimentos. Uno de estos ejemplos proviene de la literatura de adicciones, donde se han descrito, dentro de una misma camada de roedores, unos que tienen naturalmente una mayor predisposición a tener un estilo de comportamiento de corte impulsivo (28).

Estas diferencias características sólo son evidentes luego de un sencillo proceso de tamización, al utilizar una prueba comportamental relacionada con la medida de desenlace principal del estudio. De no cuantificar este rasgo, los niveles de impulsividad podrían convertirse en un factor de confusión relevante en un experimento sobre comportamiento apetitivo, especialmente con tamaños de muestra limitados. Lo anterior obligaría a realizar tamización comportamental de los sujetos del estudio y a ampliar la tamización

genética de las camadas por parte de los proveedores; ambas medidas incrementarían considerablemente los costos del estudio, tanto para los investigadores como para el proveedor.

Tal y como mencioné antes, la aleatorización de los sujetos experimentales es otro de los métodos para el control del sesgo de selección. Ésta es una condición exigida en los estudios preclínicos y se utiliza con frecuencia. Sin embargo, existen condiciones experimentales en los estudios preclínicos que pueden limitar el efecto de control que ejerce la aleatorización. Uno de estos factores es el limitado tamaño de la muestra empleado en los estudios preclínicos, aunque la aleatorización puede disminuir la probabilidad del sesgo de selección y las variables de confusión, las muestras suelen ser insuficientes para que las diferencias individuales puedan diluirse entre la población. Esto es todavía más inquietante en estudios no comportamentales, como los estudios celulares, moleculares o electrofisiológicos.

En los estudios electrofisiológicos los números de células disponibles para los experimentos suelen ser escasas y es natural que el electrofisiólogo seleccione activamente las células de mayor vitalidad para los experimentos, al imprimir un sesgo de selección. Adicionalmente, los experimentos suelen ser tan extensos y complicados que la muestra está limitada al tiempo de

los estudios, los recursos del proyecto o las limitaciones impuestas por los comités de ética respecto al número de animales disponibles para experimentación.

La aplicación de ciegos en los estudios clínicos es otro de los métodos para controlar los sesgos de medición. Es este un método aplicable a los estudios en animales, tanto de índole comportamental como celular o molecular. Si bien es posible crear los mecanismos para establecer algún tipo de ciego en los procedimientos de la investigación preclínica, desde el punto de vista logístico parece bastante difícil. Las dificultades nacen de diversos factores. Generalmente, quien realiza la medición es la misma persona que aleatoriza y manipula los animales. De no ser así se impone un estrés adicional al animal que confunde los resultados. La elaboración de un protocolo de un experimento requiere un grado de conocimiento elevado acerca de los reactivos por emplear, el animal y lo que se busca probar.

En el caso de experimentos moleculares, aplicar el ciego implica tener laboratorios grandes, donde exista un gran número de técnicos y estudiantes, de manera que por lo menos quien analiza los datos está ciego a las condiciones experimentales. En países como el nuestro, donde los grupos de investigación están limitados por los recursos económicos, es bastante difícil sostener este tipo de logística. No obstante, los investigadores preclínicos debe-

rían estar conscientes del posible efecto del sesgo de medición en sus investigaciones y cuantificarlo.

El tamaño de la muestra en los estudios preclínicos raramente se calcula tomando en cuenta un análisis de sensibilidad. Este tipo de práctica está limitada al número de animales admitidos para experimentación por parte de los comités de ética y a los recursos disponibles para la investigación. Adicionalmente, con el fin de superar estos inconvenientes, se ha planteado que todos los investigadores deben reportar el tamaño del efecto en sus observaciones. Sin embargo, tan sólo unos pocos grupos editoriales e investigadores han planteado o exigen estos datos a la fecha (29).

Respecto al error aleatorio, en la investigación preclínica suele seguirse el parámetro de admitir hasta un 5% de error en los análisis. Sin embargo, con frecuencia se encuentran diseños experimentales de medidas repetidas, en los que los investigadores no restringen el valor del error aleatorio (corrección de Bonferroni) permitido para sus pruebas estadísticas. Lo anterior lleva a la inflación del error tipo I en los análisis y las comparaciones entre grupos suelen ser erróneas. Es esencial que se haga un escrutinio de los métodos estadísticos empleados para el análisis de los datos y que se tomen las medidas para corregir este tipo de imprecisiones. También es frecuente el uso de pruebas estadísticas inadecua-

das o el uso de estadística no paramétrica, sólo con el fin de probar asociaciones significativas.

En tercer lugar, es necesario examinar las relaciones de causalidad. La complejidad de los sistemas biológicos limita enormemente esta posibilidad y suelen pasar años hasta que las intrincadas relaciones entre sistemas regulatorios son descubiertas. Tomemos como ejemplo las relaciones causales entre diferentes segundos mensajeros en fenómenos como la potenciación a largo plazo. Tomó por lo menos cincuenta años de investigaciones de múltiples grupos de investigadores y campos para descubrir estas relaciones (15).

Hoy en día existe una tendencia por parte de los comités editoriales de solicitar a los autores “historias completas” para la publicación de sus hallazgos. Por la frase historias completas hacen referencia a que en el artículo se reporte o se elabore el mecanismo que subyace al fenómeno reportado. Esto ha llevado a que los investigadores escriban, en sus discusiones, complicadas explicaciones para sustentar los hallazgos de su investigación. Éstas, muchas veces, se salen de contexto, y así como pueden resultar en un *insight* valioso, muchas otras veces resultan sólo basura. Probar la causalidad en este tipo de sistemas exige cantidades variables de horas-hombre en experimentos y recursos para pobrar relaciones causales biológicas para un fenómeno.

Una de las consideraciones que hace Ghaemi en su texto (5) es que la causalidad es un problema filosófico más que metodológico, y probablemente esté en lo cierto. Los criterios para determinar la causalidad son diversos y parecen propios a cada disciplina. Por ejemplo, en el campo de la neurofarmacología existe el concepto de que la causalidad entre un receptor específico y un efecto se prueba cuando se cumple una serie de condiciones experimentales: la estimulación natural del receptor produce un efecto que es replicable por la aplicación de un agonista, la aplicación de un antagonista para el receptor debe bloquear el efecto logrado por el agonista endógeno o el exógeno, la eliminación condicional o natural del receptor en un ratón mutante condicional o *KO* precluye el efecto, la sobreexpresión del receptor debe potenciar el efecto. Estos hallazgos deben ser reproducibles bajo múltiples condiciones.

Este tipo de aproximación para comprobar la causalidad es reminiscente de los criterios de Hill para relaciones causales en epidemiología (5). Pero, ¿esto es posible o, por lo menos, cierto? Pocos grupos de investigación logran los fondos necesarios para mantener líneas de investigación maduras por tiempos suficientes que permitan probar este tipo de relaciones causales. Es necesario unir los esfuerzos de varios grupos de investigación para este tipo de análisis. Adicionalmente, desde el punto de vista metodológico,

Barratt y colaboradores (30) han planteado el uso de medidas experimentales multidimensionales para el estudio de contratos complejos que brinden mayor soporte a las relaciones causales. Esto en el ámbito preclínico es equivalente a refutar las asociaciones utilizando múltiples tipos de aproximaciones experimentales.

En el caso del receptor arriba mencionado, no sólo utilizar criterios funcionales, sino también imagenológicos, moleculares, para blindar la asociación. La anterior reflexión sugiere que las relaciones causales reportadas en estudios individuales deben tomarse con cautela y deben construirse matrices de resultados a escala multidimensional, para evaluar el peso de las relaciones causales en la literatura preclínica.

Conclusiones

Tal y como se indicó en la introducción, este texto pretende realizar una reflexión preliminar acerca de la calidad de las investigaciones preclínicas y la aplicabilidad de principios de investigación clínica para evaluar la validez de los datos conseguidos en este tipo de investigaciones. No pretende ser una revisión comprehensiva del tema, dado que la extensión y complejidad de este tipo de labor excede los alcances de este trabajo y del conocimiento del autor. No obstante, la reflexión aquí introducida sugiere que la literatura preclínica y el ejer-

cicio mismo de la investigación pre-clínica exigen una reevaluación de nuestro quehacer como científicos. Es necesario realizar un escrutinio del error sistemático, aleatorio y de la causalidad en los estudios preclínicos con base en parámetros estrictos, como los publicados por la Asociación Estadounidense de Psicología (APA, por sus siglas en inglés *American Psychological Association*) recientemente (29).

Finalmente, las reflexiones aquí plasmadas sugieren que la complejidad de los sistemas orgánicos obliga a las ciencias biomédicas a ser más rigurosas, más integrales, y a los investigadores a ser más humildes. Pero, también, exige que los investigadores participen en los procesos económicos y políticos que afectan el ejercicio científico y la producción de conocimiento con alta calidad. Bien lo mostró K. Popper en su *Lógica de la investigación científica* (31), al señalar que la ciencia, al hablar de verdad, alude a una conjetura cuyo valor reside en poder ser falsada por otro. Todas las verdades de la ciencia son conjeturas que se basan en el mejor argumento. Y cuando esto no es posible, como lo dice el mismo Popper, el juego de la ciencia ha terminado.

Referencias

1. Orozco LE. La formación Integral. Bogotá: Uniandes; 1999.
2. Bacon R. *Novum Organum*. Buenos Aires: Porrua; 1556.
3. Lecourt D. El orden y los juegos. Buenos Aires: La Flor; 1985.
4. Kuhn TS. The structure of scientific revolutions. Chicago: University of Chicago; 1962.
5. Ghaemi SN. A clinician's guide to statistics and epidemiology in mental health : Measuring truth and uncertainty. Cambridge: Cambridge University; 2009.
6. Haynes R, Sackett R, Gray J, Cook DJ, Guyatt GH. Transferring evidence from research into practice: 1. the role of clinical care research evidence in clinical decisions. *ACP J Club*. 1996;125(3):A14-6.
7. Guyatt G, Meade MO, Rennie D, Cook DJ. User's guides to the medical literature: Essentials of evidence-based clinical practice. New York: McGraw-Hill; 2008.
8. Turner E, Mathews A, Linardatos E, Tell R, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med*. 2008;358(3):252-60.
9. Gómez-Restrepo C. Acerca del sesgo de publicación en literatura sobre antidepresivos. *Rev Colomb Psiq*. 2008;37(1):1-2.
10. Caballero A, Gómez-Restrepo C, Rodríguez M. Conflictos de interés y sesgos de publicación: ¿un problema en la literatura psiquiátrica? *Rev Colomb Psiq*. 2008;37(4):627-38.
11. Nassir Ghaemi S, Shirzadi A, Filkowski M. Publication bias and the pharmaceutical industry: the case of lamotrigine in bipolar disorders. *Medscape J Med*. 2008;10(9):211.
12. Williams DD, Garner J. The case against "the evidence": a different perspective on evidence-based medicine. *Br J Psych*. 2002;180:8-12.
13. Ioannidis JP. Why most discovered true associations are inflated. *Epidemiology*. 2008;19(5):640-8.
14. Willett WC. The search for truth must go beyond statistics. *Epidemiology*. 2008;19(5):655-6.
15. Kandel ER. In search of memory: the emergence of a new science of mind. New York: WW Norton & Company; 2006.
16. King DA. The scientific impact of nations. *Nature*. 2004;430(6997):311-6.
17. Rowland F. The peer-review process [Internet]. Acceso: 27 nov. 2009. Disponible

- en: http://www.jisc.ac.uk/uploaded_documents/rowland.pdf; 2002.
18. Bracken MB. Why are so many epidemiology associations inflated or wrong? Does poorly conducted animal research suggest implausible hypotheses? *Ann Epidemiol.* 2009;19(3):220-4.
 19. Cryan JF, Valentino RJ, Lucki I. Assessing substrates underlying the behavioral effects of antidepressants using the modified rat forced swimming test. *Neurosci Biobehav Rev.* 2005; 29(4-5): 547-69.
 20. Rupniak NM. Animal models of depression: challenges from a drug development perspective. *Behav Pharmacol.* 2003;14(5-6):385-90.
 21. Horn J, de Haan RJ, Vermeulen M, Luiten PGM, Limburg M. Nimodipine in animal model experiments of focal cerebral ischemia: a systematic review. *Stroke.* 2001;32(10):2433-8.
 22. Orozco-Cabal L, Pollandt SW, Liu J, Shinnick-Gallagher P, Gallagher JP. Regulation of synaptic transmission by CRF receptors. *Rev Neurosci.* 2006;17(3):279-307.
 23. Pound P, Ebrahim S, Sandercock P, Bracken MB, Roberts I; Reviewing Animal Trials Systematically (RATS) Group. Where is the evidence that animal research benefits humans? *BMJ.* 2004;328(7438):514-7.
 24. Saderock P, Roberts I. Systematic reviews of animal experiments. *Lancet* 2002 360:586.
 25. Roberts I, Kwan I, Evans P, Haig S. Does animal experimentation inform human healthcare? Observations from a systematic review of International animal experiments on fluid resuscitation. *BMJ* 2002, 324:474-76.
 26. Aldana E. El sistema universitario y los de ciencia y tecnología y de innovación de Colombia en el periodo 2000-2008. Santiago de Chile: CINDA, 2009.
 27. Abramson JH, Abramson ZH. Making sense of data: a self-instruction manual on the interpretation of epidemiological data. New York: Oxford University Press 2001.
 28. Dalley JW, Fryer TD, Brichard L, Robinson ES, Theobald DE, Laane K, Pena Y, Murphy ER, Shah Y, Probst K, Abakumova I, Aigbirhio FI, Richards HK, Hong Y, Baron JC, Everitt BJ, Robbins TW. Nucleus accumbens D2/3 receptors predict trait impulsivity and cocaine reinforcement. *Science* 2007, 315, 1267-1270.
 29. American Psychological Association. Publication Manual of the American Psychological Association. 6th Ed. New York: APA 2009.
 30. Orozco-Cabal LF, Barratt ES, Buccello RR. Implicaciones para el estudio de la neurobiología de la experiencia consciente. *Rev Latinoam de Psicol* 2007 39(1):109-126.
 31. Popper K. Lógica de la investigación científica. Ed. Paidós. Buenos Aires, 1934.

Conflicto de interés: el autor manifiesta que no tienen ningún conflicto de interés en este artículo.

Recibido para evaluación: 29 de noviembre del 2009

Aceptado para publicación: 10 de febrero del 2010

Correspondencia
Luis Felipe Orozco-Cabal
Departamento de Psiquiatría y Salud Mental
Facultad de Medicina. Pontificia Universidad Javeriana
Carrera 7ª No. 40-62
Bogotá, Colombia
L.orozco@javeriana.edu.co