

La vigencia de la hipótesis en la era de las ómicas

Validity of the hypothesis in the Omics Era

Validade da hipótese na era das ómicas

Amado Alfredo Quintar¹.

Palabras clave: Investigación basada en hipótesis; Investigación basada en datos; ómicas; teoría.

Keywords: hypothesis-driven research; data-driven research; omics; theory.

Palavras-chave: pesquisa orientada por hipóteses; pesquisa orientada por dados; ómicas; teoria.

“*Hypotheses non fingo*” (yo no hago hipótesis) fue la respuesta de Sir Isaac Newton a aquellos que lo desafiaban a dar explicaciones acerca de las causas de la gravedad más allá de sus principios matemáticos. Paradójicamente, él mismo inventó numerosas hipótesis, cristalizando y forjando un estándar para la generación de conocimiento en los próximos siglos. Dicho estándar, que podemos seguir denominando Método Científico, se expandió hacia todas las disciplinas y hegemonizó una forma de construcción cognitiva que permite analizar, interpretar y explicar el mundo en el que vivimos y cómo funciona. Actualmente, el método que practicamos como investigadores, y que enseñamos en nuestros espacios curriculares, siempre revisita lo establecido epistemológicamente por Francis Bacon en el siglo XVII y lo repensado en el debate Whewell-Mill del siglo XIX, con la introducción de la hipótesis como el eje que guía todo proceso de investigación científica. En esta simple historización de la hipótesis en el Método Científico, se deben sumar los postulados de Popper, quien terminó de dar forma al método hipotético-deductivo como la alternativa válida para aplicar en las ciencias biomédicas. La idea de este método es bien conocida: la ciencia procede formulando hipótesis para resolver problemas, las contrasta empíricamente (mediante observación y/o experimentación) y las modifica en función de los resultados obtenidos. Así, la práctica del investigador, en forma generalizada, radica básicamente en la formulación de la hipótesis y en su comprobación o refutación posterior. En el ámbito de la investigación biomédica, esta forma de pensar los proyectos y de “hacer ciencia” se conoce como **investigación guiada por hipótesis** (“*hypothesis-driven research*” en inglés) y centra el proceso en la teorización y la creación de esquemas conceptuales. La carga en la teorización, tanto en el planteo de la hipótesis como en la interpretación de los datos y en el modelado de la realidad para explicar fenómenos biológicos es lo que define a este tipo de investigaciones. De esta forma, han quedado entre nuestros conocimientos la teoría celular o el modelo de mosaico fluido para las membranas celulares, como ejemplos. Los modelos son indispensables en este abordaje, ya que reducen las variables y delimitan el problema de investigación a preguntas que pueden ser respondidas (anticipadamente) mediante una hipótesis; discusión aparte merece la validez de los mismos y las limitaciones de los modelos *in vivo* e *in vitro* de procesos tan complejos y dinámicos como son los referidos a sistemas biológicos o al comportamiento humano⁽¹⁾. Queda bastante claro en este paradigma que la teorización es esencial para explicar la causa de los fenómenos y para revelar los mecanismos que los gobiernan; el mismo Bacon planteaba: “la verdad surge más fácilmente del error que de la confusión”, frase que denota la necesidad de tener una hipótesis o respuesta anticipada para confrontar con los datos del mundo natural. La investigación biomédica de las últimas décadas ha visto el nacimiento y cierta preponderancia de los enfoques ómicos en las publicaciones con mayor factor de impacto. Las metodologías ómicas -ciencias o disciplinas ómicas para algunos- comprenden un grupo de tecnologías de alto rendimiento que posibilitan analizar la totalidad de moléculas (proteínas, ácidos nucleicos, lípidos, metabolitos, etc.) presentes en una única muestra, generando gran cantidad de datos e información en muy poco tiempo. Quizás el término más instalado en nuestro léxico local sea el de genómica (análisis masivo de los genes), a partir del hito que significó la publicación del genoma humano a principios de este milenio⁽²⁾. Para algunos autores, la característica masiva, exploratoria y descriptiva de las ómicas supone un cambio importante del Método Científico centrado en la hipótesis (primero se formulaba una hipótesis y luego se buscaban los datos que la apoyasen o refutasen), al permitir obtener la información del mundo natural al inicio, sobre la cual se generan (o no) teorías. Este enfoque, generalizado como **investigación basada en datos** (*data-driven research* en inglés) ha permitido trasladar los conocimientos científicos, entre otras cosas y con limitaciones, al desarrollo de la Medicina Personalizada de Precisión^(3,4).

1- Universidad Nacional de Córdoba. Facultad de Ciencias Médicas, Centro de Microscopía Electrónica. Córdoba, Argentina. Consejo Nacional de Investigaciones Científicas y Técnicas. Instituto de Investigaciones en Ciencias de la Salud. Córdoba, Argentina. E-mail de contacto: aquintar@unc.edu.ar

Recibido: 2022-12-07 Aceptado: 2022-12-07

DOI: <http://dx.doi.org/10.31053/1853.0605.v79.n3.39641>



<https://creativecommons.org/licenses/by-nc/4.0/>

©Universidad Nacional de Córdoba

La proliferación de infraestructuras digitales, datos masivos (*big data*), vigilancia tecnológica y de organización algorítmica ha generado importantes retos en la práctica del investigador científico. En este sentido, la investigación basada en datos, junto al esfuerzo por entender, reproducir y validar los hallazgos de las ciencias ómicas, implicó necesariamente la búsqueda multidisciplinaria de nuevas formas de mirar, conocer y resignificar los clásicos dogmas del Método Científico y de una nueva variedad de métodos de investigación. En esta búsqueda, en 2008 y casi una década después del secuenciamiento de genoma humano, el editor de la revista *Wired*, Chris Anderson, argumentó que la era de los datos masivos traía consigo el fin de la teoría, haciendo irrelevante cualquier forma de explicación o teorización que modelara los efectos biológicos y las acciones humanas⁽⁵⁾. Según este posicionamiento radical, la enorme cantidad de datos que generan las ómicas posibilita la aplicación de herramientas matemáticas, bioinformáticas y estadísticas para entender los fenómenos sin la ayuda de teorías previas. Así, y con suficientes datos, se materializaría la recurrente frase de “los números hablan por sí mismos”, induciendo a los investigadores académicos a aprender de los potentes algoritmos de Google o Instagram a buscar patrones donde la ciencia no puede. De esta forma, la correlación sería suficiente y reemplazaría a la causación en el eje epistemológico; el progreso de la ciencia podría darse aún sin modelos coherentes, teorías unificadoras o realmente sin ninguna explicación mecanística; no más teorías o hipótesis, no más discusiones acerca si los resultados experimentales refutan o confirman las hipótesis originales⁽⁵⁾. En esta nueva era, lo que importaría son los sofisticados algoritmos y las herramientas estadísticas para filtrar la gran cantidad de datos para encontrar información que pueda transformarse en conocimiento.

El planteo de Anderson despertó una inmediata y enérgica respuesta por parte de científicos de múltiples disciplinas⁽⁶⁻⁹⁾. En una carta publicada por la revista de la Organización Europea de Biología Molecular *EMBO reports* en 2009, Pigliucci argumentó que la falta de teorías, conceptos, modelos o hipótesis transforma a los datos de las ómicas en un mero ruido, y que la ciencia sólo avanza si se ofrecen explicaciones⁽⁷⁾. Este autor va más allá y propone que la investigación basada en datos, sin hipótesis ni modelos, es una forma de colección de los petabytes que ofrecen las ómicas, pero no sería ciencia. Asimismo, se generó un debate epistemológico acerca del tipo de inferencia inductiva, característica de las ómicas. En este sentido, y retomando el valor de la hipótesis en la construcción del conocimiento, los abordajes computacionales ómicos pueden ser tomados como un paso inicial del método científico, lo que los ubicaría en un lugar de generadores de hipótesis, en contraste al carácter comprobatorio que pretendería la ciencia clásica. Los defensores del *big data* apuntan al uso de algoritmos inductivistas como el núcleo de tales abordajes, ya que este pensamiento produciría un estado del conocimiento nunca acabado; los resultados de las inferencias inductivas tienen el potencial de alterar las inferencias previamente realizadas y, así, continuar el razonamiento en forma indefinida. Este aprendizaje permanente, nunca completado, produce un conocimiento imperfecto, pero muy útil en términos prácticos⁽⁹⁾.

El auge de las metodologías ómicas, que generan enormes cantidades de datos masivos, no dictamina que los fenómenos biomédicos sean abordados inevitablemente como *data-driven research*. En un necesario rescate de la hipótesis, la biología de sistemas, por ejemplo, está intentando usar los datos de las ómicas para construir modelos biológicos cuantitativos. Asimismo, las hipótesis son indispensables previamente a la recolección de los datos, ya que los mismos siempre provienen de experimentos diseñados o “modelados” para un fin determinado, bajo ciertas condiciones y con variables controladas. La vigencia de la hipótesis en el universo científico parece estar también promovida por las políticas de asignación de los subsidios de las grandes agencias de financiamiento, como el *National Institutes of Health* (NIH) y la *National Science Foundation* (NSF) en EE.UU. y el *Biological and Biomedical Sciences Research Council* (BBSRC) en Reino Unido⁽¹⁰⁾. Algo similar sucede en nuestra región: las guías para aplicar a subsidios del FONDECYT de Chile favorecen explícitamente la *hypothesis-driven research* como modelo de ciencia orientada por la teoría, considerando a la experimentación como la actividad subsidiaria para testear hipótesis preexistentes⁽¹¹⁾.

Como reflexión final, pensar actualmente en la hipótesis como elemento articulador de todo el proceso de investigación científica lleva a una doble lectura: por un lado, nos interpela acerca del modelo que usamos en nuestra práctica como investigadores y cuántos de nuestros proyectos tienen hipótesis claras. Por otro, la enseñanza en el grado del Método Científico se ve desafiada con la inclusión de nuevos conceptos, a veces contradictorios y en el umbral permanente de quiebre del paradigma clásico. Sin embargo, y después de todo, son debates que la comunidad científica merece y debe darse en el ámbito universitario.

BIBLIOGRAFÍA

1. Enderling H, Wolkenhauer O. Are all models wrong? *Comput Syst Oncol*. 2020 Dec;1(1):e1008. doi: 10.1002/cso2.1008.

2. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigó R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Deslattes Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X. The sequence of the human genome. *Science*. 2001 Feb 16;291(5507):1304-51. doi: 10.1126/science.1058040. Erratum in: *Science* 2001 Jun 5;292(5523):1838.

3. Gout AM, Arunachalam S, Finkelstein DB, Zhang J. Data-driven approaches to advance research and clinical care for pediatric cancer. *Biochim Biophys Acta Rev Cancer*. 2021 Aug;1876(1):188571. doi: 10.1016/j.bbcan.2021.188571.

4. Emmert-Streib F. Severe testing with high-dimensional omics data for enhancing biomedical scientific discovery. *NPJ Syst Biol Appl*. 2022 Oct 21;8(1):40. doi: 10.1038/s41540-022-00251-8.

5. Anderson, C. The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired*. 2008. Disponible en: <https://www.wired.com/2008/06/pb-theory/>

6. Defining the scientific method. Editorial. *Nat Methods*. 2009 Apr;6(4):237. doi: 10.1038/nmeth0409-237.

7. Pigliucci M. The end of theory in science? *EMBO Rep*. 2009 Jun;10(6):534. doi: 10.1038/embo.2009.111.

8. Kraus, W. L. (2015). "Editorial: Would You Like A Hypothesis With Those Data? Omics and the Age of Discovery Science." *Mol Endocrinol* 29(11): 1531-1534.

9. Mazzocchi F. Could Big Data be the end of theory in science? A few remarks on the epistemology of data-driven science. *EMBO Rep*. 2015 Oct;16(10):1250-5. doi: 10.15252/embr.201541001.

10. O'Malley MA, Elliott KC, Haufe C, Burian RM. Philosophies of funding. *Cell*. 2009 Aug 21;138(4):611-5. doi: 10.1016/j.cell.2009.08.008

11. Garrido Wainer JM, Hirmas-Montecinos N, Trujillo Osorio N. The policy of testing hypotheses in Chilean science. The role of a hypothesis-driven research funding programme in the installation of a hypothesis-driven experimental system in visual neuroscience. *Stud Hist Philos Sci*. 2022 Dec;96:68-76. doi: 10.1016/j.shpsa.2022.09.006.

Limitaciones de responsabilidad:

La responsabilidad del trabajo es exclusivamente de quienes colaboraron en la elaboración del mismo.

Conflicto de interés:

Ninguno.

Fuentes de apoyo:

La presente investigación no contó con fuentes de financiación

Originalidad:

Este artículo es original y no ha sido enviado para su publicación a otro medio de difusión científica en forma completa ni parcialmente.

Cesión de derechos:

Quienes participaron en la elaboración de este artículo, ceden los derechos de autor a la Universidad Nacional de Córdoba para publicar en la Revista de la Facultad de Ciencias Médicas y realizar las traducciones necesarias al idioma inglés.

Contribución de los autores:

Quienes participaron en la elaboración de este artículo, han trabajado en la concepción del diseño, recolección de la información y elaboración del manuscrito, haciéndose públicamente responsables de su contenido y aprobando su versión final.