



Editorial

La investigación biomédica en la encrucijada y vías de salida

Biomedical research at the crossroads and ways through

Joan M.V. Pons^{a,b,*} y Paula Adam^{a,b}^a Agència de Qualitat i Avaluació Sanitàries de Catalunya, AQUAS, Barcelona, España^b CIBER de Epidemiología y Salud Pública, CIBERESP, España

La investigación biomédica constituye la rama del árbol del conocimiento con el afán más noble: aliviar en los humanos, si no revertir, las enfermedades y lesiones, el sufrimiento y la muerte. Sus triunfos, a partir mediados del siglo XIX hasta la actualidad, han sido extraordinarios. Desde el fin de la 2.^a Guerra Mundial se han multiplicado los proyectos, investigadores, grupos y centros que persiguen aportar, aún, más beneficios a la especie dominante del planeta. Las inversiones globales por I+D en salud, sector público y privado, alcanzaron en 2009 la cifra de 240.000 millones de dólares americanos. De los 214.000 millones (\$) invertidos en países con rentas altas, el 60% provinieron del sector empresarial, el 30% del sector público y alrededor del 10% de otras fuentes, incluidas las organizaciones sin ánimo de lucro. Tan solo el 1% del total de inversión en I+D en salud se destinó a las enfermedades descuidadas, aquellas que más afectan a los países en vías de desarrollo¹.

El país con el mayor gasto (o inversión) en investigación biomédica, tanto pública (los *National Institutes of Health*, la *National Science Foundation* y la *Defense Advanced Research Projects Agency*) y privada (industria farmacéutica y de productos sanitarios) es EE. UU. Es allí donde se hacen más visibles los síntomas de agotamiento. La hipótesis de un crecimiento continuado, rápido, indefinido de la investigación biomédica ha dado lugar a un sistema insostenible, hipercompetitivo por unos recursos menguantes². La aparente expansión sin fin parece acabarse, a pesar de que, como decía Vannevar Bush (1890-1974) en su informe al presidente de los EE. UU., y donde se perfiló el modelo de investigación científica imperante, la ciencia sea una frontera interminable (1945). Las últimas décadas han visto un deterioro de todo el engranaje, surgiendo múltiples cuestiones, algunas de gran relevancia³.

Para empezar se podría hablar de la crisis de reproducibilidad. Trabajos, algunos muy citados, cuando se intentan reproducir, no dan lugar a los mismos resultados, o si se replican el efecto es mucho menor. Sorprendente, pero a la vez inquietante. La investigación científica que, en teoría, debería mantener un foco en la

falsabilidad no avanza en esta dirección. Parece como si, como describió Robert Merton (1910-2003), todo el mundo quiera ser el primero en descubrir y publicar (*first takes all*)⁴. Nadie tiene interés en repetir, para asegurar o confirmar, los trabajos de otros. Sin embargo, la replicación no deja de ser un elemento clave de rigor, de fortaleza. La crisis de reproducibilidad sacude un sistema que se *autorregula* a sí mismo, y aunque se acabe rectificando lo que se muestre falso o erróneo, genera equívocos y acaba derrochando gran cantidad de recursos⁵.

A pesar del fomento de la investigación traslacional⁶ para superar la brecha entre los descubrimientos en el laboratorio (*from bench*) y su aplicación en la atención a los enfermos (*to bedside*), persiste la fragilidad de la investigación biomédica preclínica, aquella realizada antes de iniciar la experimentación en humanos, sin que se incorporen las recomendaciones para su mejora^{7,8}. Mientras que el fracaso y la ignorancia impulsan a la ciencia hacia delante⁹, parece como si se temiera la frustración y antes se torturaran los datos para así obtener algo que, engañosamente, parezca de valor. La hipérbole con la que investigadores y medios describen resultados preliminares no contribuye a un juicio ecuánime. Reconocer la ignorancia, lo que no sabemos, incluso aquello que no sabemos que desconocemos supone humildad, virtud escasa en los escalafones más altos de la comunidad científica. Algo semejante ocurre con los sesgos: se ven los de otros pero no los propios. Y si alguna cosa caracteriza a los sesgos es su sutileza.

La escasa o nula reproducibilidad en la investigación biomédica socava aquello de la producción y acumulación del conocimiento (*a hombros de gigantes*) y contribuye no solo a un retraso en el desarrollo de nuevas intervenciones médicas, sino que añade un dispendio considerable a la investigación. Se estima que la falta de reproducibilidad supera el 50% de los estudios preclínicos evaluados, y esto se traduce en un coste (pensar como coste de oportunidad) de más de 28.000 millones de dólares. Las causas principales se clasifican como sigue: reactivos biológicos y materiales de referencia (36,1%), diseño de los estudios (27,6%), análisis de los datos y publicación (25,5%) y protocolos de laboratorio (10,8%)¹⁰.

Pero no es solo un problema de reproducibilidad o de modelos animales con escasa congruencia con las enfermedades humanas. Hay también un problema de pureza en referencia a los problemas de contaminación, como la descrita en una multiplicidad de

* Autor para correspondencia.

Correo electrónico: joan.ponsr@gencat.cat (J.M.V. Pons).

líneas celulares que circulan y se comercializan y no son lo que dicen ser. Las células HeLa, por el nombre de la paciente de quien se obtuvieron, se han esparcido y se ha contaminado gran número de laboratorios. No ha de extrañar que *Nature* solicite a los autores de artículos que impliquen líneas celulares que hayan sido verificadas previamente¹¹. Cosa parecida ocurre con los anticuerpos monoclonales, ya que algunos tampoco son lo que dicen ser o no son tan específicos como afirman³. Al problema de replicación se añade el de autenticación.

Uno de los factores causales más relevantes es el exceso de dependencia de lo que ha venido en llamarse significación estadística, la famosa «p» de Ronald Fisher (1890-1962) y que convencionalmente se interpreta como significativa cuando el valor es inferior a 0,05, y erróneamente lleva a concluir sobre la certeza de los resultados. Nada más lejos de la verdad. Forma parte de la confusión imperante en medicina entre significación estadística y significación clínica. El valor de p no es ninguna medida sobre la verdad o verosimilitud de la hipótesis evaluada, como tampoco sobre la magnitud del efecto o la importancia de los resultados¹². Lo mismo incumbe a los estudios genéticos a gran escala, incluso cuando se ponga a 5×10^{-8} el umbral de significación¹³. Basar las conclusiones en si el valor de p está por debajo de un umbral determinado y buscar alcanzarlo torturando los datos, haciendo análisis de grupos *post hoc*, persiguiendo la significación a cualquier precio solo conduce a un exceso de resultados estadísticamente significativos, pero de escasa utilidad práctica.

En una comunidad biomédica que no para de aumentar y bajo una presión creciente por unos recursos menguantes no ha de sorprender que para sobrevivir (*publish or perish*) se acabe entrando en conductas inapropiadas. Los incentivos existentes, los elementos que se valoran en la carrera profesional, solo pueden conducir al desastre. La recompensa por publicar ha crecido en exceso y relativamente a otras actividades de importancia, como pueden ser la docencia o la asistencia sanitaria basada en la implementación de los resultados de la investigación. Hoy en día hay más casos de fraude que antes y, aun siendo un porcentaje pequeño sobre el global que se publica, el número de retracciones crece año tras año, sobre todo en revistas con más renombre¹⁴. Autores o editores retiran el artículo publicado por la razón que sea, aunque suele ser por plagio, falsificación o manipulación o invención de datos o imágenes. En dos terceras partes de los casos no es por una equivocación, que cualquiera puede tener o excusar, sino por lo que se considera una mala conducta (*misconduct*) o falta de integridad. Tanto es el afán de publicar que han surgido, como modelo de negocio bien provechoso para incautos, las revistas llamadas depredadoras¹⁵. Hay un exceso de revistas donde se publican demasiados artículos, la mayor parte de escasa utilidad o equivocados. Las recompensas no pueden ser en función de la cantidad que se publica.

El problema afecta a científicos, instituciones, revistas, agencias financiadoras y por supuesto al sector privado. Todos estos agentes deberían reflexionar sobre el papel que están desempeñando en esta deriva tan peligrosa. Para revertir la tendencia habría que aprender de experiencias próximas y, como en todo, realizar una estrategia multifactorial. La bibliometría «a peso» está siendo abandonada y se están introduciendo otras formas de evaluar la ciencia y los investigadores. Importantes posicionamientos internacionales, como el *Leiden Manifesto*¹⁶ o el informe sobre *The Metric Tide*¹⁷, donde se aboga por el uso «responsable» y contextualizado de los indicadores, se espera que fomenten el abandono del mal llamado factor de

impacto. La misma comunidad científica con el *Manifeto for Reproducible Research*¹⁸, así como el uso creciente del movimiento *Open Science* suponen alternativas y nuevos mecanismos para los editores y los procesos de revisión por semejantes. Los esfuerzos de las agencias financiadoras para reducir la futilidad¹⁹ pueden mejorar, como así muestra el *National Institute of Health Research (NIHR)* de Gran Bretaña. En este país, se ha producido un cambio en el sistema de incentivos combinando un mayor énfasis en el impacto o su valoración, en el sentido de verdadera aportación relevante, que han de ser elementos claves más que la «excelencia» abstracta que solo beneficia a los pocos bien relacionados. También las universidades británicas han cambiado el sistema de financiación, progresión y recompensa profesional a través del *Reference Excellence Framework (REF)* de 2014 y la consulta abierta para el desarrollo del REF 2021²⁰.

Pero además de los incentivos es necesario un cambio cultural, un cambio en valores y por tanto también un proceso global de gestión del cambio. Desde la óptica del ecosistema de la ciencia el uso de la evidencia (o evaluación) en las políticas científicas puede aportar luz y soluciones, si bien dependerá del contexto y de la implicación de los actores clave²¹. Sin duda alguna es un problema de corresponsabilización que afecta a todos.

Bibliografía

1. Røttingen JA, Regmi S, Eide M, Young AJ, Viergever RF, Årdal C, et al. Mapping of available health research and development data: What's there, what's missing, and what role is there for a global observatory? *Lancet*. 2013;382:1286–307.
2. Alberts B, Kirschner MW, Tilghman S, Varmus H. Rescuing US biomedical research from its systemic flaws. *PNAS*. 2014;111:5773–7.
3. Harris R. Rigor mortis. How sloppy science creates worthless cures, crushes hope and wastes billions. New York: Basic Books; 2017.
4. Merton RK. Priorities in scientific discovery: A chapter in the sociology of science. *Am Sociol Rev*. 1957;22:635–59.
5. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gülmezoglu AM, et al. How to increase value and reduce waste when research priorities are set. *Lancet*. 2014;383:156–65.
6. Contopoulos-Ioannidis DG, Alexiou GA, Gouvias TC, Ioannidis JPA. Life cycle of translational research for medical interventions. *Science*. 2008;321:1298–9.
7. Horrobin DF. Modern biomedical research: An internally self-consistent universe with little contact with medical reality? *Nat Rev Drug Discov*. 2003;2:151–4.
8. Pons JMV, Pemaner-Miranda G, Camí J, Rodés J. La experimentación animal y el progreso de la medicina. *Med Clin*. 2014;143:448–54.
9. Firestein S. Failure. Why science is so successful. New York, EE. UU.: Oxford University Press; 2016.
10. Freedman LP, Cockburn IM, Simcoe TS. The economics of reproducibility in preclinical research. *PLoS Biol*. 2015;13:e1002165, <http://dx.doi.org/10.1371/journal.pbio.1002165>
11. Time to tackle cells' mistaken identity. Announcement. *Nature*. 2015;520:264.
12. Wasserstein RL, Lazar NA. The ASA's statement on p-values: context, process, and purpose. *Am Stat*. 2016;70:129–33.
13. Sham PC, Purcell SM. Statistical power and significance in large-scale genetic studies. *Nat Rev Genet*. 2014;15:335–45.
14. Fang FC, Grant R, Casadevall A. Misconduct accounts for the majority of retracted scientific publications. *PNAS*. 2012;109:17028–33.
15. Beall J. Predatory publishers are corrupting open access. *Nature*. 2012;489:179.
16. Hicks D, Wouters P, Waltman L, Rijcke S, Rafols I. Bibliometrics: The Leiden manifesto for research metrics. *Nature*. 2015;520:429–31.
17. Wilsdon J, Allen L, Belfiore E, Campbell PH, Curry S, Hill S, et al. The metric tide: Report of the independent review of the role of metrics in research assessment and management. 2015, <http://dx.doi.org/10.13140/RG.2.1.4929.1363>
18. Munafo M, Nosek B, Bishop D, Button K, Chambers CH, Percie du Sert N, et al. A manifesto for reproducible science. *Nat Human Behav*. 2017;1:0021.
19. Nasser M, Clarke M, Chalmers I, Gundro Brurberg K, Nykvist H, Lund H, et al. What are funders doing to minimise waste in research? *Lancet*. 2017;389:1006–7.
20. The second Research Excellence Framework [consultado 7 Nov 2017]. Disponible en: <http://www.hefce.ac.uk/rsrch/ref2021/>
21. Adam P, Ovseiko P, Grant J, Graham K, Boukhris OF, Dowd AM, et al. ISRIA Statement: Ten-point guidelines for an effective process of research impact assessment. *Health Research Policy and Systems*. 2017, <http://dx.doi.org/10.1186/s12961-018-0281-5>