



Big Data, reproducibilidad y ciencia abierta: hacia una nueva cultura científica

Big Data, reproducibility, and Open Science: towards a new scientific culture

Big Data, reprodutibilidade e Ciência Aberta: rumo a uma nova cultura científica

Lázaro Pablo Linares-Cánovas¹ 

¹ Universidad de Ciencias Médicas de Pinar del Río. Policlínico Docente Universitario "Luis Augusto Turcios Lima". Pinar del Río, Cuba.

Citación: Linares Cánovas LP. Big Data, reproducibilidad y ciencia abierta: hacia una nueva cultura científica. Rev Ciencias Médicas [Internet]. 2026 [citado: fecha de acceso]; 30(2026): e7272. Disponible en: <http://revcmpinar.sld.cu/index.php/publicaciones/article/view/7272>

Recibido: 02 de octubre de 2026
Aceptado: 02 de octubre de 2026
Publicado: 03 de octubre de 2026

Estimados lectores:

La ciencia biomédica vive hoy un cambio profundo, impulsado por la convergencia entre los datos masivos, la evidencia del mundo real y los principios de la ciencia abierta. Es innegable que cada día se generan volúmenes enormes de información; lo que, si bien abre oportunidades inéditas para comprender mejor diferentes procesos y enfermedades complejas, y tomar decisiones sanitarias más precisas, obliga a mirar con seriedad la calidad de los datos, su gobernanza y su uso responsable.⁽¹⁾ Más que una simple innovación tecnológica, se está configurando una nueva cultura científica, en la que integrar plataformas de datos e inteligencia artificial se ha vuelto una prioridad internacional.⁽²⁾

La evidencia del mundo real (RWE, por sus siglas en inglés), se ha convertido en uno de los pilares de esta transformación. A diferencia de los ensayos clínicos tradicionales, que suelen desarrollarse en condiciones muy controladas, la RWE permite observar cómo funcionan las intervenciones en la práctica cotidiana, en pacientes diversos y en contextos asistenciales reales. Por eso está ganando peso en la regulación de medicamentos, la evaluación de tecnologías sanitarias y la vigilancia poscomercialización.⁽³⁾ Sobre todo es sistemas de salud cada vez más exigidos, donde disponer de evidencia útil para la vida real ya no es un complemento, sino es una necesidad estratégica.

Paralelamente, las bases de datos masivas han ampliado de manera extraordinaria las posibilidades analíticas de la medicina moderna. Las historias clínicas electrónicas, los registros poblacionales, así como los biobancos y plataformas ómicas, entre otros; permiten la generación de flujos continuos de información, garantizando el estudio a escalas antes poco imaginadas de las enfermedades, permitiendo que el avance científico en esta área se dé a una velocidad poco mensurable. Si eso se combina con el empleo de métodos avanzados de análisis, surge nuevas perspectivas para la identificación de factores de riesgo, predicción de desenlaces clínicos, y el consiguiente diseño de intervenciones más efectivas.⁽⁴⁾ No obstante, el entusiasmo por el volumen de datos no debe eclipsar un principio fundamental, la calidad metodológica sigue siendo el principal determinante de la validez científica. Ante ello, los grandes conjuntos de datos no sustituyen el rigor; por el contrario, exigen mayores estándares de transparencia, trazabilidad y control de sesgos.

Durante la última década, una de las grandes preocupaciones, con la amplia expansión de los datos masivos es la crisis de reproducibilidad científica. Y es que múltiples disciplinas han documentado dificultades para replicar resultados previamente publicados, fenómeno que

ha generado debates sobre la confiabilidad de la literatura científica. Frente a esta realidad, las revistas de alto impacto y las agencias financiadoras promueven cada vez más el preregistro de protocolos, la disponibilidad de códigos analíticos, la publicación de datos subyacentes y la adopción de estándares de reporte más estrictos.⁽⁵⁾ La reproducibilidad no debe entenderse únicamente como un requisito técnico, sino como una expresión de integridad científica y un compromiso ético con la sociedad que financia y utiliza los resultados de la investigación. El fortalecimiento de los estudios de replicación constituye además una prioridad emergente en la agenda biomédica internacional.

En este nuevo escenario, compartir datos deja de ser una opción secundaria y pasa a ser una condición clave para acelerar el conocimiento, por ello, la ciencia abierta propone que resultados, bases de datos, y códigos puedan encontrarse y reutilizarse de manera responsable. Dicho enfoque facilita la colaboración y mejora la eficiencia de la inversión científica, algo especialmente valioso frente a desafíos sanitarios globales. De la misma forma obliga a desarrollar marcos de gobernanza que protejan la privacidad y promuevan estándares comunes, haciendo posible un intercambio confiable de información.

Desde la perspectiva cubana, estos desafíos adquieren matices particulares. Cuba posee fortalezas reconocidas en materia de vigilancia epidemiológica, Atención Primaria de Salud e integración de redes asistenciales, elementos que generan una valiosa fuente de datos para la investigación en Salud Pública. Sin embargo, la limitada disponibilidad de infraestructuras tecnológicas avanzadas, las restricciones económicas (propias del embargo económico al que está sometido la isla), y la necesidad de fortalecer ecosistemas nacionales de gestión de datos científicos continúan representando retos relevantes. Ante esta situación, resulta imprescindible promover políticas institucionales que favorezcan la digitalización progresiva, la interoperabilidad de sistemas, la alfabetización en ciencia de datos y la participación activa en iniciativas regionales e internacionales de ciencia abierta. Estas acciones permitirían potenciar la visibilidad y el impacto global de la producción científica nacional.

A la luz de estas reflexiones, la Revista de Ciencias Médicas de Pinar del Río, en su condición de espacio para el diálogo científico y la difusión del conocimiento, reconoce en la transición hacia una cultura sustentada en datos abiertos, investigación reproducible y evidencia del mundo real una oportunidad estratégica para fortalecer el impacto social de la ciencia y su contribución a la toma de decisiones en salud. La excelencia editorial ya no depende únicamente de publicar resultados novedosos, sino también de garantizar transparencia, accesibilidad y capacidad de verificación.

En consecuencia, promovemos una investigación comprometida con los principios de la ciencia abierta, la ética de los datos y el rigor metodológico, entendiendo que estos elementos son esenciales para construir confianza pública y avanzar hacia sistemas sanitarios más equitativos y sostenibles. El futuro de la investigación médica no estará definido por quién acumule más información, sino por quién sea capaz de transformarla en conocimiento útil, reproducible y compartido para el beneficio colectivo.

REFERENCIAS BIBLIOGRÁFICAS

1. Cremin CJ, Dash S, Huang X. Big Data: Historic Advances and Emerging Trends in Biomedical Research. Curr Res Biotechnol [Internet]. 2022 [citado 28/06/2026]; 4:138-151. Disponible en: <https://doi.org/10.1016/j.crbiot.2022.02.004>

2. Raghupathi W, Raghupathi V. Big data in healthcare and medicine revisited: design and managerial challenges in the age of artificial intelligence. Health Inf Sci Syst [Internet]. 2026 [citado 28/06/2026]; 14:12. Disponible en: <https://doi.org/10.1007/s13755-026-00433-2>
3. Burns L, Le Roux N, Kalesnik-Orszulak R, Christian J, Dudinak J, Rockhold F, et al. Real-world evidence for regulatory decision-making: updated guidance from around the world. Front Med (Lausanne) [Internet]. 2023 [citado 28/06/2026]; 10:1236462. Disponible en: <https://doi.org/10.3389/fmed.2023.1236462>
4. Beesley LJ, Salvatore M, Fritsche LG, Pandit A, Rao A, Brummett C, Willer CJ, Lisabeth LD, Mukherjee B. The emerging landscape of health research based on biobanks linked to electronic health records: Existing resources, statistical challenges, and potential opportunities. Stat Med [Internet]. 2020 [citado 28/06/2026]; 39(6):773-800. Disponible en: <https://doi.org/10.1002/sim.8445>
5. Ellis RJ. Questionable Research Practices, Low Statistical Power, and Other Obstacles to Replicability: Why Preclinical Neuroscience Research Would Benefit from Registered Reports. eNeuro [Internet]. 2022 [citado 28/06/2026]; 9(4):ENEURO.0017-22.2022. Disponible en: <https://doi.org/10.1523/eneuro.0017-22.2022>