

## Big Data y medicina de precisión: una visión general hacia el futuro

Walter Masson 

*Jefe de Prevención Cardiovascular, Hospital Italiano.  
Ciudad Autónoma de Buenos Aires, Argentina.*

*Acta Gastroenterol Latinoam 2023;53(3):203-206*

*Recibido: 21/08/2023 / Aceptado: 21/09/2023 / Publicado online: 30/09/2023 / <https://doi.org/10.52787/agl.v53i3.345>*

A diferencia del enfoque "un talle único para todos", la medicina de precisión utiliza la información médica (incluyendo la genética), conductual y ambiental de una persona para individualizar aún más su atención. Esto podría conducir a la mejor predicción de la enfermedad de una persona en riesgo y al diagnóstico y tratamiento más eficaz cuando la afección estuviera presente. Los grandes datos permiten mucha más precisión y adaptación de lo que nunca antes fue posible mediante la vinculación de diversos conjuntos de datos que revelan correlaciones y vías causales hasta ahora desconocidas. Pero también se plantean cuestiones éticas relacionadas con el

equilibrio de intereses, la viabilidad del anonimato, las implicancias familiares y de grupo, así como la discriminación genética.

Factores de comportamiento y estilo de vida (como la dieta, las actividades diarias, incluso el uso de las redes sociales) interactúan con factores genéticos para causar enfermedades. Esto significa que los datos relacionados con el estilo de vida también contienen información sobre los determinantes clave del riesgo para muchas de las enfermedades crónicas comunes que aquejan a las sociedades modernas. Estos factores de comportamiento pueden interactuar con factores biológicos y causar enfermedades. El acceso a los datos relacionados con esos factores conductuales no sólo permite una mejor comprensión de los efectos biológicos, sino también identificar cambios de comportamiento que pueden mitigar los efectos de las variantes biológicas en la susceptibilidad a la enfermedad.

El término genérico "macrodatos" o "Big Data" se refiere al conjunto de datos estructurados y no estructurados que, por ser extremadamente grande y complejo, no permite la recopilación, procesamiento, almacenamiento y análisis según los métodos tradicionales.<sup>1</sup> En otras palabras, esta gran cantidad de información sólo se puede procesar y analizar mediante el uso de nuevas herramientas tecnológicas y digitales.

Algunas de las principales características del concepto "Big Data" son el volumen (enorme cantidad de datos), la velocidad (crecimiento continuo y exponencial de la

---

**Correspondencia:** Walter Masson  
Correo electrónico: [walter.masson@hospitalitaliano.org.ar](mailto:walter.masson@hospitalitaliano.org.ar)

información), la variedad (datos de diferentes fuentes de información, alta complejidad), la veracidad (fuentes de información confiables) y el valor agregado (generación de nuevas oportunidades de desarrollo).<sup>2-3</sup> El avance de la tecnología y la informática ha permitido procesar estos macrodatos, buscando tendencias o patrones en la información analizada, con el objetivo de responder preguntas originadas en diferentes disciplinas.

Las tecnologías avanzadas desarrolladas para Big Data han impulsado su aplicación en muchas áreas, como la lucha contra el crimen y cuestiones relacionadas con la inseguridad, la administración empresarial, las finanzas, el Sistema de Posicionamiento Global (GPS), el comercio, el turismo, la meteorología, la biología o el medio ambiente, entre otras.<sup>4</sup> Por otro lado, las aplicaciones de Big Data en las ciencias de la salud en general, y en la medicina en particular son múltiples, abarcando prácticamente todos los ámbitos: genómica, epidemiología, ensayos clínicos, algoritmos diagnósticos y pronósticos, telemedicina, gestión administrativa, etc. Es decir, la aplicabilidad es infinita e incluye cuestiones relacionadas con la investigación, la docencia y las tareas asistenciales. La fuente de datos médicos sobre los que podríamos aplicar técnicas analíticas de Big Data son muy variados, desde datos personales y socioeconómicos, características clínicas, información administrativa hasta datos relacionados con los estudios complementarios de laboratorio o imágenes. La información que surge luego de analizar los datos con técnicas de Big Data permiten optimizar la práctica médica, haciéndola más “personalizada” o “centrada en el paciente”.<sup>5</sup> En ese sentido, el análisis de los macrodatos permite “modelar” al paciente y ofrecerle a cada individuo lo que mejor se adapte a sus características personales. Adicionalmente, las aplicaciones de Big Data podrían hacer que la medicina sea más participativa.<sup>6</sup> Con la información generada, el paciente puede participar activamente en las decisiones relacionadas con su salud, facilitando, entre otras cosas, una mejor adherencia a los tratamientos.

Otros dos escenarios en los que se aplican las técnicas analíticas de Big Data son la medicina preventiva y la medicina predictiva, mucho más tangibles para el médico general. En el primer caso, podemos citar como ejemplos la posibilidad de realizar más efectivamente la vigilancia de brotes epidémicos, situaciones de emergencia sanitaria, farmacovigilancia<sup>7-8</sup> o el desarrollo de mejores campañas de salud pública, desde la vacunación a la prevención de la obesidad o las conductas suicidas.<sup>9-11</sup> En el segundo caso, partiendo del análisis basado en las técnicas de Big Data, pueden crearse modelos predictivos

más precisos, como por ejemplo para estimar el riesgo de presentar una emergencia cardiovascular, hospitalizarse, o tener una buena o mala evolución luego de un diagnóstico oncológico.<sup>12-14</sup>

Finalmente, deberían tenerse en cuenta algunas preocupaciones vinculadas a las técnicas de Big Data. En este contexto es importante considerar ciertas cuestiones metodológicas al momento de interpretar la información, como los sesgos de registro y asociación. Asimismo, deben contemplarse algunas cuestiones éticas, como la privacidad de los pacientes.<sup>15</sup>

En el ejemplo que mostramos a continuación evaluaremos la evidencia a partir de la información provista por Big Data.

### Ejemplo

Una cuestión muy importante de la práctica clínica habitual es la relación existente entre la enfermedad por reflujo gastroesofágico (ERGE) y los denominados síntomas extraesofágicos. Mucho se ha dicho de la relación entre la ERGE y las exacerbaciones de la enfermedad pulmonar obstructiva crónica (EPOC), aunque aún es controvertida. Estudios previos han demostrado que los pacientes con ERGE y EPOC coexistentes tienen peor calidad de vida y dificultad para respirar más severa que aquellos con EPOC, pero sin ERGE. Además, se ha sugerido que la ERGE sería un factor de riesgo de exacerbación aguda. Aún es más controvertida la utilización de inhibidores de la bomba de protones (IBP) para evitar las exacerbaciones agudas en pacientes con EPOC.

En el siguiente ejemplo se analizan los efectos del tratamiento de la ERGE con IBP en el riesgo de las exacerbaciones agudas y neumonía en pacientes con EPOC sobre la base del siguiente estudio científico: Respir Res. 2023 Mar 11;24(1):75. Se trata de un estudio en el que se utiliza una gran base de datos poblacional que cuenta con información referente a los diagnósticos médicos y los tratamientos instaurados.

El objetivo del estudio fue evaluar los riesgos tanto de la exacerbación como de la neumonía después del tratamiento con IBP para la ERGE en pacientes con EPOC.

Este estudio utilizó una base de datos de reembolso de la República de Corea. Se incluyeron pacientes  $\geq 40$  años con EPOC como diagnóstico principal y que recibieron tratamiento con IBP para ERGE al menos durante 14 días consecutivos entre enero de 2013 y diciembre de 2018.

Se realizó un análisis de serie de casos auto controlados para calcular el riesgo de exacerbación moderada y grave

y neumonía, en el que cada sujeto de estudio actuó como su propio control, para minimizar la influencia de los factores de riesgo individuales de exacerbación o neumonía.

## Resultados

Un total de 104.439 pacientes con EPOC recibieron tratamiento con IBP para la ERGE. El riesgo de exacerbación moderada fue significativamente menor durante el tratamiento con IBP que al inicio del estudio. El riesgo de exacerbación grave aumentó durante el tratamiento con IBP, pero disminuyó significativamente en el período posterior al tratamiento. El riesgo de neumonía no aumentó significativamente durante el tratamiento con IBP.

## Conclusiones

El riesgo de exacerbación se redujo significativamente tras el tratamiento con IBP en comparación con el período en el que no fueron tratados. La exacerbación grave puede aumentar debido a la ERGE no controlada, pero luego disminuir después del tratamiento con IBP. No hubo evidencia de un mayor riesgo de neumonía.

## Fortalezas y Limitaciones del estudio

La fortaleza de este trabajo es que utiliza una serie de pacientes con un tamaño muestral de gran envergadura, que permite describir la población y dimensionar el problema (representatividad).

Las grandes bases de datos poblacionales permiten construir hipótesis a partir de sus hallazgos y el impacto de varios factores en forma simultánea. La construcción de modelos explicativos o predictivos a partir de estos datos es otra fortaleza de estas bases de gran tamaño.

A su vez, es una serie de casos auto controlados, lo que favorece la minimización del problema de los confundidores.

Las debilidades la constituyen los sesgos en relación a la población elegida, la calidad de los datos y la validez externa, ya que la muestra se toma de una población de Corea y probablemente no sea trasladable a otros grupos poblacionales. También deben considerarse algunas cuestiones en relación a Big Data, como son la heterogeneidad, el tratamiento de los datos, las técnicas de datos asociados que son diferentes a los tradicionales, el alto costo de expertos, el costo y el manejo de la manipulación de la privacidad de los datos.

En resumen, las técnicas analíticas de Big Data llegaron para quedarse. Su aplicación correcta conducirá a una medicina de precisión, enfocada en la prevención de las enfermedades y su tratamiento, aunque considerando la

variabilidad genética individual, el ambiente y el estilo de vida de cada persona.

**Propiedad intelectual.** El autor declara que los datos presentes en el manuscrito son originales y se realizaron en su institución perteneciente.

**Financiamiento.** El autor declara que no hubo fuentes de financiación externas.

**Conflicto de interés.** El autor declara no tener conflictos de interés en relación con este artículo.

## Aviso de derechos de autor



© 2023 Acta Gastroenterológica Latinoamericana. Este es un artículo de acceso abierto publicado bajo los términos de la Licencia Creative Commons Attribution (CC BY-NC-SA 4.0), la cual permite el uso, la distribución y la reproducción de forma no comercial, siempre que se cite al autor y la fuente original.

**Cite este artículo como:** Masson W. Big Data y medicina de precisión: una visión general hacia el futuro. Acta Gastroenterol Latinoam. 2023;53(3):203-206. <https://doi.org/10.52787/agl.v53i3.345>

## Referencias

1. Gomes MAS, Kovalski JL, Pagani RN, da Silva VL, Pasquini TCS. Transforming healthcare with big data analytics: technologies, techniques and prospects. J Med Eng Technol. 2023 Jan;47(1):1-11. <https://doi.org/10.1080/03091902.2022.2096133>
2. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2014;2(1). <https://doi.org/10.1186/2047-2501-2-3>
3. Pramanik PK, Pal S, Mukhopadhyay M. Healthcare big data: a comprehensive overview. In research anthology on big data analytics, architectures, and applications. IGI Glob 2022;19-47. <https://doi.org/10.4018/978-1-6684-3662-2.ch006>
4. Hong L, Luo M, Wang R, Lu P, Lu W, Lu L. Big Data in Health Care: Applications and Challenges. Data and Information Management. 2018;2(3):175-197. <https://doi.org/10.2478/dim-2018-0014>
5. Schulte T, Bohnet-Joschko S. How can Big Data Analytics Support People-Centred and Integrated Health Services: A Scoping Review. Int J Integr Care. 2022 Jun 16;22(2):23. DOI: 10.5334/ijic.5543

6. Konstantinidis M, Lalla EA. Clinical anisotropy: A case for shared decision making in the age of too much data and patient dis-integration. *J Eval Clin Pract.* 2020 Apr;26(2):604-609. DOI: [10.1111/jep.13312](https://doi.org/10.1111/jep.13312)
7. Bouzillé G, Poirier C, Campillo-Gimenez B, Aubert M-L, Chabot M, Chazard E. Leveraging hospital big data to monitor flu epidemics. *Comput Methods Programs Biomed.* 2018 Feb;154:153-160. DOI: [10.1016/j.cmpb.2017.11.012](https://doi.org/10.1016/j.cmpb.2017.11.012)
8. Trifiro G, Sultana J, Bate A. From big data to smart data for pharmacovigilance: the role of healthcare databases and other emerging sources. *Drug Saf.* 2018;41:143-149.
9. Mills G. Big data drive efficient rabies vaccination. *Vet Rec.* 2021 Feb;188(3):88-89. DOI: [10.1002/vetr.150](https://doi.org/10.1002/vetr.150)
10. Detecting Suicide and Self-Harm Discussions Among Opioid Substance Users on Instagram Using Machine Learning. *Front Psychiatry.* 2021 May 31;12:551296. DOI: [10.3389/fpsy.2021.551296](https://doi.org/10.3389/fpsy.2021.551296)
11. Tu B, Patel R, Pitalua M, Khan H, Gittner LS. Building effective intervention models utilizing big data to prevent the obesity epidemic. *Obes Res Clin Pract.* 2023 Mar-Apr;17(2):108-115. DOI: [10.1016/j.orcp.2023.02.005](https://doi.org/10.1016/j.orcp.2023.02.005)
12. Scali ST, Stone DH. The role of big data, risk prediction, simulation, and centralization for emergency vascular problems: Lessons learned and future directions. *Semin Vasc Surg.* 2023 Jun;36(2):380-391. DOI: [10.1053/j.semvascsurg.2023.03.003](https://doi.org/10.1053/j.semvascsurg.2023.03.003)
13. Schulte T, Wurz T, Groene O, Bohnet-Joschko S. Big Data Analytics to Reduce Preventable Hospitalizations-Using Real-World Data to Predict Ambulatory Care-Sensitive Conditions. *Int J Environ Res Public Health.* 2023 Mar 7;20(6):4693. DOI: [10.3390/ijerph20064693](https://doi.org/10.3390/ijerph20064693)
14. Choi JW, Kang S, Lee J, Choi Y, Kim HC, Chung JW. Prognostication and risk factor stratification for survival of patients with hepatocellular carcinoma: a nationwide big data analysis. *Sci Rep.* 2023 Jun 27;13(1):10388. DOI: [10.1038/s41598-023-37277-9](https://doi.org/10.1038/s41598-023-37277-9)
15. Kayaalp M. Patient Privacy in the Era of Big Data. *Balkan Med J.* 2018 Jan 20;35(1):8-17. DOI: [10.4274/balkan-medj.2017.0966](https://doi.org/10.4274/balkan-medj.2017.0966)