

Carta al editor

“Big data” en ciencias de la salud: Aspectos importantes

“Big data” in health sciences: Important issues

JUAN SANTIAGO SERNA TREJOS¹, STEFANYA GERALDINE BERMUDEZ MOYANO²,
HOOVER LEON GIRALDO³, JUAN PABLO LENIS GONZÁLEZ⁴

Recibido para publicación: 06-07-2023. Versión corregida: 28-12-2023. Aprobado para publicación: 28-05-2023.

Modelo de citación:

Serna Trejos J.S., Bermudez Moyano S.G., Giraldo H.L., Lenis González J.P. “Big data” en ciencias de la salud: Aspectos importantes, Colombia. Arch Med (Manizales). 2023;23(2).
<https://doi.org/10.30554/archmed.23.2.4789.2023>

Keywords: *Big Data ; Medical Informatics Applications; Medical Informatics Computing; Health Information Technology.*

Sr. Editor:

El advenimiento de la era digital ha significado un avance importante en las tecnologías de la comunicación y por ende, todo lo concerniente a ellas, desde la creación de nuevos conocimientos, pasando por su almacenamiento, hasta la generación de modelos y sistemas de gestión y/o procesamiento de datos “*Big Data*” (BD). Sin embargo no solo basta con la generación de información y /o el volumen de información generada a nivel global a través de los medios informáticos, el cual crece a una velocidad exponencial y en cantidades sumamente alarmantes en lo que respecta a el manejo de las mismas, sino también se debe de tomar en

1 Departamento de Epidemiología, Universidad Libre – Seccional Cali. Cali, Colombia; Departamento de Docencia Universitaria, Universidad Piloto - Bogotá, Colombia; Instituto Nacional de Salud. Bogotá, Colombia; Grupo interdisciplinario de investigación en epidemiología y salud pública, Cali, Colombia, Hospital Universitario del Valle, Unidad de Cuidado Intensivo, Cali- Colombia. <https://orcid.org/0000-0002-3140-8995>

2 Hospital Universitario del Valle, Unidad de Cuidado Intensivo, Cali- Colombia.
<https://orcid.org/0000-0002-2259-6517>

3 Grupo de especialistas en manejo integral de enfermedades crónicas (GESENCRO), Departamento de estadística y epidemiología, Palmira- Colombia. <https://orcid.org/0000-0002-8854-105X>

4 Hospital Carlos Holmes Trujillo, Unidad de consulta externa, Cali - Colombia.
<https://orcid.org/0000-0001-9004-475X>.

Correspondencia a: Juan Santiago Serna - Trejos

Correo electrónico: juansantiagosernatrejos@gmail.com

cuenta variables en la BD , como la heterogeneidad de los datos (diversidad); velocidad (rapidez de procesamiento) , veracidad (que tan precisa y cierta es la información) y su valor (qué utilidad supone la información procesada). A un mayor volumen, se precisan de estrategias de manejo de datos óptimas y flexibles para encontrar las variables y/o información que se requiera(1).

En ciencias de la salud, dicho fenómeno no es ajeno al desarrollo de la misma disciplina, puesto que se requiere siempre de toma de decisiones basadas en evidencia concreta y de alta calidad, sin embargo, el número alarmante de información disponible requiere de la generación de modelos de decisión que involucra algoritmos y una serie de conectores y demás criterios de selección explícitos e implícitos en la información disponible para construir la mejor búsqueda posible de la información(1).

Existen múltiples aplicaciones de BD en ciencias de la salud, por ejemplo: en el área de la salud pública supone una herramienta de gran valor en el estudio de las enfermedades crónicas, un estudio realizado por *Guo et al*, pretende examinar la heterogeneidad de los ingresos en diferentes hogares y su impacto en la salud de los mismos, determinando que en hogares con bajos ingresos o bajo capital con adultos mayores como cabezas de hogar, suponen el grupo poblacional con mayor impacto en su salud cuando dicha brecha monetaria se hace cada vez más amplia, generando así recomendaciones para las diferentes entidades gubernamentales en la redistribución de los fondos destinados para la atención en salud como priorización en el acceso de salud de tal población vulnerable(2). Otra aplicación de BD en la salud pública se puede observar en el estudio realizado por *Lin et al*, mediante en el cual implementó un sistema de aprendizaje automático en la predicción de deterioro cognitivo leve en pacientes con enfermedad de Alzheimer, empleando biomarcadores genéticos. Mediante la ejecución de algoritmo logró mostrar mediante el análisis de 29 genes, cuales de estos se asociaron mediante un modelo de decisión, que pacientes estables con la presencia de dichos genes, progresaron a presentar deterioro cognitivo leve, constituyendo una gran herramienta en la medicina de precisión(3).

En otros campos de la salud como la epidemiología, la BD supone gran tamaño de información, que deben de analizarse y, por consiguiente, al generar una muestra grande, conducirá a un poder estadístico mayor y por consiguiente mayor precisión en la misma, generando valores de P pequeños, que se traducen en hallazgos que no se relacionan con el azar, reforzando o descartando con más severidad la generación de diversas hipótesis(4).

Otras disciplinas como la oftalmología, guardan estrecha relación con la BD, gran parte de su aplicación se traslada a manejo de BD, dado que muchos de sus valores de aplicación viene desde valores numéricos (presión intraocular), imágenes bidimensionales (Fotografía de fondo de ojo) imágenes tridimensionales (tomografía de coherencia óptica), generación de lentes convergentes y divergentes basados en dioptrías por análisis de múltiples curvas poblacionales a determinadas

distancias. Existen bases de datos como “ *Smart Eye Database*” (*SMEYEDAT*) la cual almacena gran información en relación con datos oftalmológicos, la cual en tiempo real pueden ilustrar imágenes para precisar en ambientes académicos la identificación rápida y fácil de pacientes con condiciones específicas(5).

Otro caso de aplicación de BD se ha dado en la nefrología, para la construcción de investigación nefrología de base poblacional, es decir, mediante el seguimiento de la función renal y su impacto en el desarrollo de la enfermedad renal crónica (ERC) , se generan diferentes escalas de desarrollo y ciertos valores de referencia de acuerdo a las tasas de filtración glomerular en determinado grupo poblacional, como las variables tomadas en cuenta para la generación de la tasa de filtración glomerular, mientras que unas escalas como *Cockcroft-Gault* toman en cuenta variables como : creatinina, variable , peso, sexo ; otras escalas como *MRDR* toman en cuenta sexo, edad, creatinina y raza (6). En la cardiología clínica, la BD ha incidido en gran medida en la toma de decisiones, por ejemplo: mediante la creación de la base de datos *TriNetX*, la cual se constituye como una red de investigación global d alta importancia, aporta gran información con relación a diagnósticos clínicos, hallazgos de laboratorio, tratamientos recibidos y procedimientos realizados en más de 250 millones de pacientes a nivel global. Dicha información, se enfatiza en pacientes con diagnóstico de hipertensión arterial, diabetes mellitus tipo 2 o enfermedad renal crónica. El blanco de información del portal *TriNetX* está encaminada en revelar el perfil de seguridad y los diferentes resultados cardiovasculares de los fármacos cardiovasculares, la eficacia de los protocolos de rehabilitación y las implicaciones cardiovasculares de los medicamentos que constituyen ensayos clínicos en curso(7)which are generated and released by different sources and channels, like epidemiological surveys, national registries, electronic clinical records, claims-based databases (epidemiological Big Data).

Las aplicaciones de la BD son diversas en el campo del área de ciencias de salud, la generación de información relacionada en este campo dentro de las diferentes disciplinas es cada vez es más cuantiosa, como su conexión transversal entre las mismas áreas, hace necesario análisis más exhaustivos de la misma y la generación de alternativas eficaces de la diferente BD.

Conflictos de interés

Los autores no declaran conflictos de interés

Financiamiento

Autofinanciado.

Contribución de autoría

Todos los autores han contribuido en la concepción, redacción de borrador- redacción del manuscrito final, revisión y aprobación del manuscrito.

Referencias

1. Wang L, Alexander CA. Big data analytics in medical engineering and healthcare: methods, advances and challenges. J Med Eng Technol [Internet]. 2020;44(6):267–83. Available from: doi:10.1080/03091902.2020.1769758
2. Guo H, Yang Y, Pan C, Xu S, Yan N, Lei Q. Study on the Impact of Income Gap on Health Level of Rural Residents in China. Int J Environ Res Public Health [Internet]. 2022;19(13). Available from: doi:10.3390/ijerph19137590
3. Lin RH, Wang CC, Tung CW. A Machine Learning Classifier for Predicting Stable MCI Patients Using Gene Biomarkers. Int J Environ Res Public Health [Internet]. 2022;19(8). Available from: doi:10.3390/ijerph19084839
4. Chan CL, Chang CC. Big Data, Decision Models, and Public Health. Int J Environ Res Public Health [Internet]. 2022;19(14). Available from: doi:10.3390/ijerph19148543
5. Cheng CY, Soh Z Da, Majithia S, Thakur S, Rim TH, Tham YC, et al. Big data in ophthalmology. Asia-Pacific J Ophthalmol [Internet]. 2020;9(4):291–8. Available from: doi:10.1097/APO.0000000000000304
6. Kaur N, Bhattacharya S, Butte AJ. Big Data in Nephrology. Nat Rev Nephrol [Internet]. 2021;17(10):676–87. Available from: doi:10.1038/s41581-021-00439-x
7. Dai H, Younis A, Kong JD, Puce L, Jabbour G, Yuan H, et al. Big Data in Cardiology: State-of-Art and Future Prospects. Front Cardiovasc Med [Internet]. 2022;9(April):1–13. Available from: doi:10.3389/fcvm.2022.844296

