

EDITORIAL

Big data, creación de valor en nutrición clínica**Big data: Value creation in clinical nutrition**

Julia Alvarez Hernández

Sección de Endocrinología y Nutrición, Hospital Universitario Príncipe de Asturias, Universidad de Alcalá, Alcalá de Henares, Madrid, España

La inteligencia artificial (IA) ha irrumpido en nuestras vidas de forma imparable. Su origen se remonta a los trabajos de lógica matemática y computación de Alan Turing¹ y McCulloch y Pitts, publicados hace más de 80 años². IA es un término acuñado por Minsky y McCarthy en 1956, en la conferencia de Dartmouth³. Se define como el campo científico de la informática que se centra en la creación de programas y mecanismos que pueden mostrar comportamientos considerados inteligentes. En otras palabras, «las máquinas piensan como seres humanos».

Pero para poder conseguir su desarrollo, la IA se nutre de una gran cantidad de datos (*big data* [BD]) y los emplea para desarrollar algoritmos y para constituir su propia lógica. En definitiva, utiliza los datos para obtener información del entorno e interactuar con él en consecuencia. Así, IA y BD son términos estrechamente vinculados que ofrecen oportunidades de mejora en todas las disciplinas si sabemos aprovechar los sistemas de información a nuestra disposición.

BD es la voz inglesa que se emplea en el sector de las tecnologías de la información y de la comunicación, para aludir a un conjunto de datos que, por su volumen y variedad, y por la velocidad a la que necesitan ser procesados, supera las capacidades de los sistemas informáticos habituales. En español, se ha propuesto traducir BD por «macrodatos», presentándola como «una alternativa válida pues aporta, como *big*, el significado de «grande»; es una solución breve y no tiene, como ocurriría con «megadatos», un posible riesgo de confusión con el término mega, muy frecuente también en los mismos ámbitos⁴.

El BD involucra todos los aspectos de la vida humana, incluida la biología y la medicina⁵. Los avances de las últimas décadas, el mundo de las ómicas (genómica, proteómica, metabolómica) y otros tipos de tecnologías, así como la implantación de la historia clínica electrónica (HCE), han condicionado un exponencial crecimiento de los datos contribuyendo a la realidad del BD en el ámbito sanitario. Hoy en día es una demanda real que puede permitir aumentar el conocimiento para innovar y mejorar la calidad y eficiencia asistencial. Según Margolis, «El *big data* es no solo una nueva realidad para el científico biomédico, sino un imperativo que debe entenderse y utilizarse eficazmente en la búsqueda de nuevos conocimientos»⁶.

Algunos autores señalan que el término BD no tiene una definición adecuada en el tesoro de Medline (MeSH). Tras una revisión en profundidad, destaca que en las publicaciones sanitarias el término que mejor define el BD es el volumen⁷.

Es evidente que en el campo de la medicina es dominante el enorme volumen de datos de asistencia sanitaria existente. Esta incluye registros médicos personales, imágenes médicas, datos a nivel genético, secuencias genómicas de datos de población, datos de investigación clínica (estudios observacionales, ensayos clínicos etc.) y un largo etcétera. Más recientemente, también están alimentando este crecimiento exponencial las imágenes en 3D, así como las lecturas de los sensores biométricos o los dispositivos «vestibles» (*wearables*’), es decir, dispositivos incorporados a la vestimenta o usados corporalmente como implantes o accesorios que pueden actuar como extensión del cuerpo o mente del usuario, algunos de ellos muy utilizados en el campo de la Endocrinología y la nutrición clínica.

Si bien es cierto que el concepto de volumen de datos en medicina, como en otras disciplinas, es relevante para

Correo electrónico: julia.alvarez@movistar.es

hablar de BD, cada vez se considera más que el concepto de BD está asociado a las 5 «V»: volumen, velocidad, variedad, veracidad y valor. Podríamos decir que estamos hablando del conjunto de datos o combinaciones de conjuntos de datos de gran tamaño (volumen), diversos y complejos (variedad), que crecen rápidamente y necesitan ser procesados (velocidad), reales, auténticos y de calidad (veracidad) que analizados adecuadamente aportan valor.

Por todo esto, no podemos quedarnos en la superficie conceptual del término y el gran reto, en mi opinión, está en el continuo desarrollo, implementación real, extensa y colaborativa de sistemas de análisis de datos más potentes que los métodos tradicionales. Hay que poner el foco en el desarrollo de plataformas para la captura más eficaz, almacenamiento y manipulación de estos grandes volúmenes de datos que nos permitan añadir valor a nuestra actividad sanitaria.

Un buen ejemplo de esto es la herramienta Savana Manager, utilizada por Ballesteros et al. en su estudio⁸. Este sistema innovador, mediante la tecnología EHRead, es capaz de analizar y extraer de forma automática la información clínica relevante contenida en el texto libre de las HCE mediante técnicas de procesamiento del lenguaje natural y de BD, y transformarla en información ordenada con fines de investigación⁹.

Ballesteros et al. prueban en su estudio una de las utilidades del BD en la investigación clínica y gestión sanitaria. El análisis de un voluminoso conjunto de datos de las HCE ha identificado que el infradiagnóstico de la desnutrición relacionada con la enfermedad (DRE) sigue siendo una realidad. Esta circunstancia ampliamente destacada por distintas iniciativas en la lucha contra la desnutrición en una preocupación constante de los endocrinólogos dedicados a la nutrición clínica¹⁰.

El BD del estudio en cuestión valora más de 180.000 registros de hospitalización y, con un menor esfuerzo que el empleado en los estudios de prevalencia clásicos, permite hacer una aproximación a las características de la población estudiada. Los pacientes identificados como desnutridos eran mayoritariamente pacientes con insuficiencia cardíaca (35%), infección respiratoria (23%), infección urinaria (20%) y enfermedad renal crónica (15%). Además, ha permitido conocer que estos pacientes tenían una mayor edad (75 vs. 59 años), mortalidad (7,08% vs. 2,98%) y estancia (8 vs. 5 días, $p < 0,0001$) que los pacientes sin diagnóstico de desnutrición. El estudio establece que el 2,47% de los episodios incluían el diagnóstico de DRE, cifra alejada de más del 23% encontrada en el estudio Predyces¹¹.

El verdadero valor de este ejercicio de análisis de BD reside en que, con un relativo menor esfuerzo, los autores tienen información relevante para gestionar el abordaje de la DRE en su centro. Pueden definir mejor dónde poner el foco de atención para implantar acciones de mejora que consigan sensibilizar a las unidades asistenciales seleccionadas por los datos de infradiagnóstico, para hacer una detección precoz de pacientes en riesgo nutricional o desnutridos que se beneficien de una intervención específica que mejore la calidad asistencial y reduzca los costes.

Tenemos suficientes evidencias de cómo los avances en BD e IA, junto a la inteligencia humana, están siendo capaces de hacer una medicina de alto rendimiento. Las utilidades a día de hoy de la IA van desde la selección de embriones

en procesos de fecundación in vitro, el control médico con dispositivos que actúan mediante lenguaje oral tipo Alexa, hasta el control de salud mental, la monitorización de parámetros de interés (presión arterial, frecuencia cardíaca, electrocardiograma, glucemia, etc.) o de adherencia terapéutica, acciones paramédicas en el ámbito cardiológico o neurológico (infartos de miocardio, ictus etc.), asistentes para la lectura de imágenes radiológicas y la interpretación de las mismas, prevención de ceguera (lectura de retinografías), identificar mutaciones responsables de cáncer, promover la seguridad del paciente e incluso prever la muerte en el entorno hospitalario¹². Es esperable que, en su progresión, los avances permitan seguir mejorando la prevención, el diagnóstico y el tratamiento de enfermedades, así como una mejor y más eficiente gestión de la atención sanitaria con un impacto positivo en la calidad y la eficiencia.

Pero para alcanzar estos logros que transportarían el sistema sanitario a una nueva era es necesario aunar esfuerzos para conseguir involucrar a los distintos protagonistas del desarrollo del proceso de digitalización (administraciones públicas, empresas privadas, hospitales, médicos, centros de investigación, universidades, etc.). Uno de los retos más importantes en todos los países en este desarrollo es integrar la tecnología con las políticas de privacidad y confidencialidad, las infraestructuras y una cultura de uso compartido de los datos¹³. Aunque somos conscientes de que la transformación digital en salud en España ya ha comenzado, voces autorizadas proponen, desde hace algunos años, el desarrollo de una Estrategia Nacional de Salud Digital en la que se tengan en cuenta todos estos factores, se incorpore un marco claro de cooperación, se priorice la implementación de casos de uso de valor compartido y se haga viable la medición del impacto de la misma¹⁴.

Bibliografía

1. Turing AM. Computing machinery and intelligence. *Mind*. 1950;59:433–60.
2. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys*. 1943;5:115–33.
3. Conferencia de Dartmouth 1956. Disponible en: <https://dartmouthconference.wordpress.com/>.
4. Barea Mendoza J, Rovira Forns J, Quecedo Gutiérrez L, Gol Montserrat J, del Llano Señarís J. Oportunidades y retos de los macrodatos (big data) en la toma de decisiones sanitarias. Pinto, Madrid: Fundación Gaspar Casal ed.; 2018.
5. Zhang Z. Big data and clinical research: Focusing on the area of critical care medicine in mainland China. *Quantitative Imaging in Medicine and Surgery*. 2014;4(5):426–9.
6. Margolis R, Derr L, Dunn M, et al. The National Institutes of Health's Big Data to Knowledge (BD2K) initiative: Capitalizing on biomedical big data. *Journal of the American Medical Informatics Association*. 2014;21(6):957–8.
7. Baro E, Degoul S, Beuscart R, Chazard E. Toward a literature-driven definition of big data in healthcare. *BioMed Research International*. 2015 [Article ID 639021], 9 p. Disponible en: <https://doi.org/10.1155/2015/639021>.
8. Ballesteros MD, Pintor de la Maza B, Barajas Galindo D, Cano Rodríguez I. Persiguiendo la desnutrición relacionada con la enfermedad mediante herramientas de big data. *End Nutr y Diab*. 2020.
9. Espinosa-Anke L, Tello J, Pardo A, Medrano I, Ureña A, Salcedo I, et al. Savana: A global information extraction and

- terminology expansion framework in the medical domain. *Procesamiento del Lenguaje Natural*. 2016;57:23–30. Disponible en: <http://hdl.handle.net/10230/33531>.
10. Fight Against Malnutrition. Disponible en: <https://www.espen.org/fight-against-malnutrition>.
 11. Álvarez-Hernández J, Planas Vila M, León-Sanz M, García de Lorenzo A, Celaya Pérez S, García Lorda P, et al. Prevalence and costs of malnutrition in hospitalized patients; the PREDyCES Study. *Nutr Hosp*. 2012;27(4):1049–59, <http://dx.doi.org/10.3305/nh.2015.32.4.9700.Predyces>.
 12. Topol EJ. High-performance medicine: The convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44–56.
 13. Menasalvas E, Gonzalo C, Rodríguez-González A. Big data en salud: retos y Oportunidades. *Economía Industrial*. Disponible en: <https://www.mincotur.gob.es/Publicaciones/Publicacionesperiodicas/EconomiaIndustrial/RevistaEconomiaIndustrial/405/MENASALVAS,%20GONZALO%20Y%20RODR%C3%8DGUEZ.pdf>.
 14. Parra Calderón CL. Big data en sanidad en España: la oportunidad de una estrategia nacional. *Gac Sanit*. 2016;30(1):63–5.