



EDITORIAL

“Big Data”; aplicación y utilidad para el sistema sanitario

José Manuel Martínez Sesmero

Dr. por la UCM y Especialista en Farmacia Hospitalaria. Servicio de Farmacia. Complejo Hospitalario de Toledo. España.

“Big Data”; aplicación y utilidad para el sistema sanitario

Empecemos definiendo de manera sencilla el anglicismo, aparentemente difuso, que conocemos como *Big Data*: se trata de aquel conjunto de datos que, por su tamaño ingente, sobrepasa la capacidad de ser gestionado por bases de datos de integración tradicionales. A pesar de que muchos autores consideran esta definición demasiado dispersa¹, si profundizamos en las características que componen el *Big Data*, existe un mayor grado de acuerdo en aducir que se fundamenta en el paradigma de la 3 “V” (volumen, variedad y velocidad)². El elevado volumen de datos (más de un *petabyte*) precisa nuevas técnicas de almacenamiento a gran escala y enfoques distintos para recuperar la información; la variedad de las fuentes de datos (texto, audio, vídeo, etc.) hace que las redes relacionales sencillas sean difícilmente aplicables; y por último, el incesante incremento con que se generan los datos, hace que la velocidad sea un parámetro clave en su manejo³. Según el *Massachusetts Institute of Technology* podríamos añadir una cuarta “V”, que hace referencia al valor aportado por estos grandes conjuntos de datos.

El análisis del *Big Data* ha abierto la puerta a una nueva era para la mejora en la prestación de servicios y solución de problemas en el ámbito de los sistemas sanitarios. La gran mayoría de los agentes que participan en las estructuras de los servicios de salud reconocen que el análisis del *Big Data* puede ofrecer nuevas posibilidades en la elaboración de modelos predictivos, patrones de comportamiento, el descubrimiento de nuevas necesidades, reducir riesgos, así como proveer servicios más personalizados, todo ello en tiempo real y teniendo en cuenta toda la información relevante.

El mejor cuidado de la salud, impulsado por el tratamiento global y sistemático de los datos, debiera plantearse desde tres perspectivas⁴: la individual de cada paciente, la de las comunidades de pacientes y, por último, la de la sociedad en su conjunto.

Individualmente, en el área de dispositivos médicos y de los cuidados domiciliarios se están desarrollando aplicaciones de monitorización remota de electrocardiograma, seguimiento de los datos de los sensores de dispositivos como marcapasos o desfibriladores implantados, que han logrado grandes avances en lo que respecta a la gestión y la prevención de rehospitalizaciones por tales pacientes⁵. Estas experiencias, a través de la utilización de datos de múltiples sensores, van encaminadas a ayudar a la vida independiente de las personas mayores, con un gran potencial para mejorar los procesos asistenciales complejos y facilitar un cuidado individualizado y eficiente para crónicos y ancianos bajo estas circunstancias.

En las comunidades de pacientes, los sujetos se influyen mutuamente, y parte de esta influencia hoy fluye a través de las redes sociales y en la web 2.0 (Facebook, Twitter, etc.). Este hecho es relevante tanto para aprovecharlo como instrumento encaminado a las cuestiones relacionadas con la salud, incluso como modelo para entender cómo los patrones de comportamiento pueden emerger o involucionar. Las redes sociales pueden ser aprovechadas directamente como una herramienta para ayudar a los pacientes a vivir más y con mejor calidad. Por ejemplo, ya existen varios servicios en línea dirigidos a los pacientes con condiciones específicas, para unirse a determinadas comunidades en las que se interrelacionan, intercambian información u opiniones, y ofrecer y recibir apoyo (por ejemplo

* Autor para correspondencia.

Correo electrónico: jmanuelms@sescam.jccm.es (Dr. José Manuel Martínez Sesmero).

<http://www.patientslikeme.com/>). Estas redes pueden convertirse en excelentes fuentes de datos si cada vez es más grande el número de pacientes que participan en ellas, y pueden ayudar a descubrir nuevos e inesperados patrones de salud o aspectos relacionados con la misma.

Una tercera área donde existe un claro potencial significativo para las grandes tecnologías de *Big Data*, desde un prisma más global, es la de gestión de enfermedades infecciosas. Las epidemias se extienden a través de personas y de los contactos persona-persona o por el entorno, y ello puede suponer un gran peligro para la sociedad, sobre todo en los países en vías de desarrollo. Sin embargo, debido a la gran movilidad de personas que se produce actualmente, esta idea ha de extenderse a los países con sistemas de salud más evolucionados. Por ejemplo, la Comisión Europea ha desarrollado el sistema "MediSys"⁶, una herramienta para escanear y buscar información con el objeto de reforzar la red de vigilancia de enfermedades transmisibles y la detección temprana de las actividades bioterroristas. A través del algoritmo de "MediSys" se pueden obtener noticias de última hora utilizando más de 20.000 artículos de Internet analizados al día y producidos por la *Europa Media Monitor*, que se pueden enviar a las personas clave, gestores, decisores, etc., por correo electrónico y SMS. Algunas instituciones sanitarias, e incluso autoridades en esta materia de determinados países, están potenciando el análisis del *Big Data* para resolver los problemas tradicionales de salud, como la reducción de los reingresos⁵, el aumento de la eficacia y eficiencia de la asistencia sanitaria, la mejora de la calidad de la atención, y las demandas de predicción para los servicios de salud.

En el ámbito de la farmacoterapia, la acumulación de datos sanitarios (clínicos, genéticos, etc.) que relacionan las estructuras químicas de sustancias a una enfermedad por su efecto terapéutico o adverso, ha permitido desarrollar modelos predictivos que asocian una estructura química con el efecto que tendría en humanos. Como muestra de ello, recientemente se ha publicado un estudio⁷ que incluye moléculas relacionadas con la etiología de 934 condiciones que amenazan la salud y se utilizan para tratar 835 enfermedades diferentes. En primer lugar, se identifican restos químicos que podrían estar asociados de forma independiente con cada efecto fenotípico. Con el uso de estos fragmentos, se construyen predictores precisos para aproximadamente 400 fenotipos clínicos, encontrando muchas estructuras relacionadas y responsables de estos efectos.

Existen varios factores que deben ser tenidos en cuenta a la hora de utilizar el análisis del *Big Data* en el ámbito sanitario⁸. En primer lugar, los nuevos componentes y relaciones entre los datos han de trabajar de forma sistemática y deben garantizar altas cotas de calidad, sólo así obtendremos resultados e informaciones útiles y valiosas para la comprensión de los diferentes contextos de los

estados de salud y pronosticar el futuro de la asistencia sanitaria. En segundo lugar, tenemos que manejar la información que genere el *Big Data* con cierto grado de prudencia, incluso después de las diferentes consideraciones estratégicas que se puedan plantear, ya que a pesar de que tenga mucho potencial para mejorar nuestra comprensión de los fenómenos asistenciales y fortalecer nuestra capacidad de predecir con cierto grado de exactitud el futuro, siempre se han de barajar otros aspectos coyunturales, ético-morales, etc. Por lo tanto, no tenemos que pensar en los resultados del análisis del *Big Data* como única fuente de información, por muy poderosa que fuera, ni tampoco se debe incurrir en ideas preconcebidas ni tendenciosas. En tercer lugar, se habrá de garantizar la seguridad y la protección de los datos, que en el caso del sector sanitario son de especial sensibilidad (datos de carácter personal relacionados con la salud, información clínica, datos genéticos, etc.) y que según la legislación vigente siempre habrán de ser preservados.

Finalmente, sería recomendable el estudio académico formal del análisis del *Big Data* para poder ampliar y validar la aplicación práctica del mismo, tanto en el ámbito de la asistencia hospitalaria y, desde luego, puertas afuera de su estructura (atención primaria, centros sociosanitarios, etc.). Dicho estudio debería centrarse en los criterios más significativos y objetivos que podamos disponer, tales como la eficacia, efectividad clínica y social, eficiencia, utilidad, para adoptar el análisis del *Big Data* en la práctica real como una herramienta más en la toma de decisiones.

Conflicto de intereses

El autor declara no tener ningún conflicto de interés.

Bibliografía

1. Y. Genovese and S. Prentice, "Pattern-based strategy: getting value from big data," Gartner Special Report G00214032, 2011.
2. The Big Data Research and Development Initiative, http://www.whitehouse.gov/sites/default/files/microsites/ostp/bigdata_press_release_final_2.pdf.
3. Merelli I, Pérez-Sánchez H, Gesing S, D'Agostino D. Managing, analysing, and integrating big data in medical bioinformatics: open problems and future perspectives. *Biomed Res Int*. 2014;2014:134023.
4. Grossglauser M, Saner H. Data-driven healthcare: from patterns to actions. *Eur J Prev Cardiol*. 2014 Nov;21(2 Suppl):14-7.
5. IBM. Reducing readmissions to improve care [Internet]. Armonk (NY): IBM; c2014 [cited at 2014 Oct 20]. Available from: <http://www-01.ibm.com/software/sg/industry/healthcare/pdf/setonCaseStudy.pdf>.
6. European Commission. Medical intelligence in Europe [Internet]. Brussels: European Commission; c2014 [cited at 2014 Oct 20]. Available from: http://ec.europa.eu/health/preparedness_response/generic_preparedness/planning/medical_intelligence_en.htm.
7. Duran-Frigola M, Rossell D, Aloy P. A chemo-centric view of human health and disease. *Nat Commun*. 2014 Dec 1;5:5676.
8. Song TM. Efficient utilization of big data on healthcare and welfare area. *Healthc Welf Forum* 2012;193:68-76.