

## Profesión

Ramón de Burgos Pol

Analista en Farmacoeconomía y Evaluación de Resultados en Salud en PORIB (Pharmacoeconomics & Outcomes Research Iberia). Evaluador externo de aplicaciones móviles en salud (Distintivo AppSaludable) de la Agencia de Calidad Sanitaria de Andalucía

“Las aplicaciones analíticas de *big data* y su capacidad de generar evidencia a nivel de paciente pueden arrojar luz sobre el beneficio de un tratamiento para un paciente particular»

# «Big data» y salud: preparados, listos... ¿ya?

No es extraño sentirse a la vez atraídos e intimidados por el fenómeno del *big data*: la utilización masiva de datos. Prácticamente cada semana aparecen referencias a ello en los medios, vaticinando una suerte de revolución irrefrenable basada en la utilización de datos, ingentes y masivas bases de datos: patrones de predicción de enfermedades, modelos predictivos de comportamiento, individualización de tratamientos, identificación de factores de riesgo a nivel de persona y paciente...

Mi sensación es que los profesionales sanitarios asistimos un poco incrédulos ante tanta maravilla, en cierto modo porque no entendemos muy bien una parte técnica muy relevante relacionada con el tratamiento y gestión de los datos. Lo cierto es que actualmente se generan infinidad de datos relacionados con la salud, aunque su análisis y aplicación no ha sido ni está siendo nada fácil para el sector. La Organización Mundial de la Salud (OMS) publicaba en 2015 en su boletín algunas de las grandes ventajas que podría traer esta tecnología disruptiva a la salud global, pero también algunas de las amenazas si no se realiza correctamente<sup>1</sup>.

En general, el término *big data* se refiere a un volumen masivo de datos electrónicos no estructurado. Reúne cuatro propiedades: volumen, velocidad, variedad y valor<sup>2</sup>. Requiere un gran volumen de datos que se recogen a gran veloci-



©zenbor/123RF

dad, e integra múltiples fuentes<sup>1</sup> como registros, historia clínica, redes sociales, aplicaciones móviles, geolocalización, *wearables*, etc. Puede integrar, además, información sobre contaminación, tráfico, calefacción, tiendas de comestibles y mercados o insumos alimenticios para mejorar la precisión de la estratificación de la población por riesgo<sup>3</sup>. Mucha de esta información ni siquiera existía hace años o no estaba digitalizada como hasta ahora. Ante tal ingente cantidad de información, el *big data* tiene que generar valor y, en consecuencia, la capacidad de transformar<sup>4</sup>.

Hasta ahora, los ensayos clínicos aleatorizados, el *gold standard* de la investigación en el campo de la salud y por supuesto en el sector farmacéutico, se basan en la utilización de datos homogéneos y estructurados. Esta información clínica es la que se usa en el proceso natural de autorización de un fármaco, y la que se utiliza en el Informe de Posicionamiento Terapéutico en la negociación de reembolso y precio. Y no es para menos ya que, al utilizar técnicas como la aleatorización o el cegamiento de pacientes y profesionales de forma prospectiva, proporciona mayor calidad a esa evidencia por minimización de sesgos (errores sistemáticos). No obstante, este proceso implica un gran consumo de recursos: son costosos y requieren la implicación de profesionales sanitarios y organismos públicos, y por supuesto hay que reclutar pacientes. En un ensayo clínico avanzado o en fase III, los pacientes ofrecen sus DATOS sin reparos, ya que al fin y al cabo se está investigando la enfermedad que les afecta y, por lo tanto, obtendrán *a priori* algo a cambio. El resultado de la investigación, en la cual hay muchos intereses y profesionales implicados, beneficiaría al paciente directamente.

### «Big data»: aplicaciones en la práctica clínica

El *big data* se relaciona directamente con otros dos conceptos: *Real World Data* (RWD) y *Real World Evidence* (RWE). Esto en realidad no es otra cosa que datos en contexto real, que analizados de forma adecuada pueden llegar a generar evidencia científica, es decir, RWE. Los datos homogéneos y bien estructurados de los que se nutre la investigación clínica «clásica» (p. ej., ensayos clínicos aleatorizados) representan sólo una pequeña parte de los datos disponibles. Existe otra ingente cantidad de datos no estructurados «de larga cola»<sup>5</sup>. Su análisis tendría diferentes aplicaciones, como ayudar a comprender mejor cómo funcionan determinados fármacos en contexto real, factores relacionados con la no adherencia, comorbilidades del paciente, interacciones y efectos secundarios...<sup>6</sup>.

### Herramientas diagnósticas

Una aplicación puede ser la mejora de las herramientas y criterios diagnósticos. En la práctica habitual, se usan herramientas que presentan ciertas limitaciones y que se utilizan como criterio de selección para instaurar un determinado tratamiento (p. ej., Framingham para el riesgo



©rawpixel/123RF

“ La utilización de técnicas de datos que permitiese conocer mejor a la población cercana a un establecimiento sanitario como una farmacia daría la posibilidad de ofrecer servicios farmacéuticos profesionales que realmente se ajustaran a las necesidades de la población de referencia»

cardiovascular en la prescripción de estatinas). Las aplicaciones analíticas de *big data* y su capacidad de generar evidencia a nivel de paciente (medicina personalizada) puede arrojar luz sobre el beneficio de un tratamiento para un paciente particular, ayudando a estimar modelos predictivos más precisos para valorar el riesgo del paciente, especialmente si se encuentra cercano al umbral establecido por una determinada herramienta<sup>7</sup>.

### Adherencia y farmacia comunitaria

Según un estudio estadounidense<sup>8</sup>, el desarrollo de herramientas clínicas basadas en datos de varias fuentes (datos demográficos, perfil de paciente, datos fisiológicos de la historia clínica digital, datos sobre adherencia desde farmacia, medidas de salud comunicadas por el paciente [*patient reported outcomes*]) se ha mostrado ya útil para mejorar la adherencia entre un 6 y un 20% en pacientes con antidiabéticos y medicamentos cardiovasculares. Otro estudio permitió identificar distintos factores de riesgo (baja presión arterial o tipo de cobertura médica) y que un tercio de los

pacientes a los que se les había prescrito un nuevo antihipertensivo no recogía la medicación en los 30 días siguientes a su prescripción<sup>9</sup>.

### **Perspectiva económica: generar valor y generar ingresos**

La utilización de técnicas de datos que permitiesen conocer mejor a la población cercana a un establecimiento sanitario como una farmacia daría la posibilidad de ofrecer servicios farmacéuticos profesionales que realmente se ajustaran a las necesidades de la población de referencia. Se crearía una demanda que, además de ser un vehículo para la monitorización de patologías gestionadas desde el ámbito ambulatorio (p. ej., diabetes, hipercolesterolemia...), podría generar una fuente de ingresos adicional.

Otra de las ventajas del *big data* es la capacidad de ajustarse al consumidor. Al poderse recoger datos relacionados con el bienestar, datos sanitarios y hábitos de consumo, se dan las condiciones para analizar esa información y ofrecer servicios ajustados a las necesidades del cliente. En este sentido, la farmacia comunitaria podría posicionarse de forma privilegiada, especialmente teniendo en cuenta la confianza que genera el farmacéutico como profesional sanitario entre los ciudadanos. Desde el punto de vista de gestión, el análisis de la ingente cantidad de datos de proveedores farmacéuticos orientada al *stock* de la farmacia repercutiría, además, en una gestión más eficiente de la farmacia comunitaria.



## **El análisis masivo de datos hace aflorar ciertos riesgos, como el de la privacidad de las personas»**

### **«Warning»: esto es privado y otros problemas del «big data»**

Pese al enorme potencial del *big data*, hoy en día existen fuertes limitaciones para su expansión. Las diferentes fuentes utilizadas necesitan de interoperabilidad<sup>5</sup>, de códigos de estandarización de la información que permitan agilizar su integración. Además de los aspectos relacionados con la seguridad de la información y de los aspectos legales y éticos, es necesario desarrollar herramientas que permitan integrar información de diferentes formatos y vocabularios para explotar la posibilidad del *big data* de forma eficiente<sup>10</sup>. Pero sin duda la seguridad de la información es uno de los aspectos que más preocupa, ya que no se trata sólo de un problema técnico.

### **«Big data» y protección de datos**

El análisis masivo de datos hace aflorar ciertos riesgos, como el de la privacidad de las personas. Las técnicas de seu-

dominización o disociación (los datos de una persona se asocian a un código para evitar su identificación, aunque se sabe que el conjunto de datos pertenece a una determinada persona) cobran un papel muy importante en la fase de recolección y procesamiento automatizado de datos para la creación de modelos predictivos propia de una primera fase del *big data*<sup>11</sup>. Un ejemplo a escala nacional es el proyecto VISC+ (*big data* sanitario con fines científicos que centraliza los datos sanitarios en Cataluña), que peligró durante unos años por las dudas de la Autoridad Catalana de Protección de Datos acerca de la cesión a terceros de la información que aglutinaría. El centro de la polémica versaba sobre la confidencialidad de los datos. Finalmente, la Agencia de Calidad y Evaluaciones Sanitarias (AQuAS), un organismo público, se encargará de la gestión, distribución y «anonimización» de los datos sin que haya posibilidad de venta a entidades privadas.

Lo ocurrido con este ambicioso proyecto debería hacernos reflexionar sobre el papel de las autoridades sanitarias y organizaciones en la gestión de los datos de los usuarios, ya que éstos demandan cada vez más información sobre qué se va a hacer con ellos.

### **Perdone, ¿qué va a hacer usted con mis datos?**

Un amigo, «A», me comentó hace poco que, en una revisión periódica de salud laboral, lo primero que le pidieron fue que cumplimentara un formulario sobre cesión de los datos de salud a la aseguradora que proveía los servicios. Pero «A» decidió no hacerlo. Y no me dio más explicaciones. Le estuve dando vueltas a los posibles motivos y lo llamé. Me dijo que nadie le explicó qué iban a hacer con sus datos, que fue como una especie de imposición. Le preocupaban aspectos como la cesión a terceros o quién se responsabiliza de la custodia de los datos. Me llamó la atención que no pondría ningún problema si fuera el sistema regional de salud (el público) el que se lo pidiera, ya que entiende que no va a hacer un uso lucrativo. Pero insistió: «Nadie me explicó qué iban a hacer con ellos ni cómo iban a custodiar sus datos personales». ¿Les suena? Asimetría de información y privacidad/confidencialidad de los datos.

Sin duda, alguien podría haber informado a mi amigo, que es bastante observador y curioso, de que sus datos serían disociados, y si se van a ceder o no a terceros o si le pedirán permisos, etc. Todas ellas en realidad son limitaciones ya identificadas por la Agencia Española del Medicamento recientemente<sup>4</sup>. Quizá lo único que quería escuchar «A», como luego me confirmó, era que su uso estaba destinado a poder integrarlos en bases de datos más grandes con las que poder identificar algún patrón o riesgo, y así contribuir de alguna manera a mejorar la atención a la salud. El hecho es que «A», de alguna forma, exige que se le preste atención a algo que considera tiene «valor» *per se*, y que además trata de algo tan importante como su «privacidad».



©rawpixel/123RF

## Desde el punto de vista sanitario, la farmacia comunitaria podría adquirir un papel relevante en la mejora de la adherencia con la integración de las técnicas de *data mining* y *big data*»

### ¿Valen realmente dinero mis datos?

Valor y monetización no tienen por qué ir de la mano. El responsable del proyecto 4T, Chema Alonso, es muy claro cuando asegura que los servicios de Internet que son gratuitos en realidad los paga el usuario con sus datos (la famosa ley de Internet de que, si no lo pagas, no eres el cliente, sino el producto). Esto abre el campo de la monetización de los datos: ¿cuánto dinero podrías darme por mis datos? Me ha sorprendido encontrar en *Financial Times* una herramienta que permite estimar el coste de tus datos en función de tus características financieras, personales, estado de salud, gustos y planes de futuro. La verdad es que, desde un punto de vista individual, no significan apenas nada en cuanto a ingresos, pero, claro, a gran escala es otra cosa.

### Consideraciones finales

Desde el punto de vista sanitario, la farmacia comunitaria podría adquirir un papel relevante en la mejora de la adherencia con la integración de las técnicas de *data mining* y *big data*. El hecho de compartir datos y colaborar con organismos públicos, cuidando mucho los aspectos de privacidad y confidencialidad de la información necesaria, podría permitir generar evidencia que no sólo mejoraría las condiciones de salud de la población, sino que además contribuiría a un uso más eficiente de los recursos. El análisis de

metadatos destinado a crear modelos predictivos de comportamiento que detecten baja adherencia basados en datos disociados de usuarios (demográficos, estados de salud, prescripción, ingresos hospitalarios...) permitirá identificar el nivel de adherencia de cada uno de los pacientes o subgrupos específicos que no se encuentran identificados por la literatura científica. Se abre camino, así, a la creación de servicios farmacéuticos profesionales individualizados o programas que, basados en la evidencia específica de su entorno, estén centrados en determinados subgrupos, de forma que resuelvan de manera práctica y asequible problemas globales actuando localmente. A pesar de las limitaciones todavía existentes (como la privacidad/confidencialidad de la información y la interoperabilidad de los sistemas), las potencialidades que ofrece el *big data* para generar valor y transformar el sector sanitario son enormes. Y todo versa sobre lo mismo: tus datos. ●

### Bibliografía

1. Wyber R, Vaillancourt S, Perry W, Mannava P, Folaranmi T, Celi LA. Big data in global health: improving health in low- and middle-income countries. *Bull World Health Organ.* 2015; 93(3): 203-208.
2. Ward JS, Barker A. Undefined by data: a survey of Big data definitions. Ithaca: Cornell University Library, 2013. Disponible en: <http://arxiv.org/pdf/1309.5821v1.pdf> (último acceso: 10/01/2017).
3. Parra Calderón CL. Big data en Sanidad en España: la oportunidad para una estrategia nacional. *Gac Sanit.* 2016; 30(1): 63-72.
4. Onukwughu E. Big Data and its role in health economics and outcomes research: a collection of perspectives on data sources, measurement and analysis. *Pharmacoeconomics.* 2016; 34(2): 91-93.
5. Macías Saint-Gerons D, de la Fuente Honrubia C, de Andrés Trelles F, Catalá-López F. Perspectiva futura de la farmacoepidemiología en la era del «Big Data» y la expansión de las fuentes de información. *Rev Esp Salud Pública.* 2016; 0: e1-e7.
6. Collins B. Big Data and Health Economics: strengths, weaknesses, opportunities and threats. *Pharmacoeconomics.* 2016; 34(2): 101-106.
7. Rumsfeld JS, Joynt KE, Maddox TM. Big data analytics to improve cardiovascular care: promise and challenges. *Nat Rev Cardiol.* 2016; 13(6): 350-359.
8. Dixon B, Alzeer AH, Phillips EO, Marrero DG. Integration of provider, pharmacy, and patient-reported data to improve medication adherence for type 2 diabetes: a controlled before-after pilot study. *JMIR Med Inform.* 2016; 4(1): e4.
9. Comer D, Couto J, Aguiar R, Wu P, Elliott D. Using aggregated pharmacy claims to identify primary nonadherence. *Am J Manag Care.* 2015; 21(12): e655-e660.
10. Fernández-Luque L, Mejova Y, Mayer MA, Hasvold PE, Joshi S. Panel: big data & social media for empowering patients with diabetes. *Stud Health Technol Inform.* 2016; 225: 607-609.
11. Gil E. Big data, privacidad y protección de datos. Agencia Española de Protección de Datos. Madrid: Agencia Estatal Boletín Oficial del Estado, 2015.