



Medicina de Familia SEMERGEN

www.elsevier.es/semergen



EDITORIAL

Dilemas éticos del *big data* en asistencia sanitaria primaria

Ethical dilemma of big data in primary healthcare



Actualmente, en la era de la informatización, el incremento de datos que podríamos denominar como macrodatos o «*big data*», adquiere un crecimiento exponencial con múltiples aplicaciones prácticas, de las que la asistencia sanitaria en cuidados primarios no se encuentra exenta. Dichos datos se caracterizan no solamente por su volumen y variabilidad, sino por la complejidad en su gestión y la velocidad en la adquisición de los mismos (pudiéndose obtener en tiempo real) lo que implicará, a nivel de la información clínica, la aparición de una nueva serie de herramientas inteligentes en la toma de decisiones en materia de salud. A nivel de la asistencia sanitaria se precisarán, por tanto, novedosas iniciativas basadas en la generación de algoritmos para mejorar el diagnóstico y tratamiento de las enfermedades, así como de la digitalización de los datos clínicos. A través del desarrollo de procedimientos de extracción y empleo de macrodatos a gran escala, será posible desarrollar un análisis matizado de la salud del paciente por medio de todos los datos clínicos disponibles, elaborando un enfoque más completo de todas aquellas características que influyen en el diagnóstico y el tratamiento de las enfermedades¹.

A diferencia de la Asistencia Hospitalaria, la mencionada tecnología disruptiva adquiere especial preponderancia en el control y seguimiento de pacientes con alteraciones crónicas (enfermedades respiratoria, diabetes...), en los que la Asistencia Primaria adquiere una marcada importancia en la detección y seguimiento de los mismos, tal y como se pone de manifiesto en las numerosas asistencias de dichos pacientes en centros de Salud, produciendo en numerosos casos el colapso de la misma si no se incorporan mecanismos de gestión acorde con esta nueva tecnología.

En este sentido, hemos de citar la publicación pionera de Ginsberg et al.², centrada en la localización de registros de consultas de Google para detectar la prevalencia de la gripe en regiones específicas de Inglaterra con gran número de población. Dicho informe marcaría un punto de inflexión, asistiendo en la actualidad a un progresivo

incremento de estudios similares y desarrollo de nuevas aplicaciones tecnológicas para el procesamiento de los «*big data*» en investigación traslacional, como *hadoop* en secuenciación NGS (*Next Generation Sequencing*), así como en la fenotificación de pacientes a través de información extraída de los historiales médicos. Para proporcionar soporte a la gestión de los mismos, constituye una realidad actual la existencia de nuevos perfiles profesionales como los «*data manager*» especializados en datos de naturaleza biológica, y el empleo de nuevas herramientas de gestión, fundamentalmente algoritmos digitales, para el procesamiento de la información, que sean capaces de la detección precoz de la enfermedad o situaciones de riesgo, como ya sucede con la aplicación «Google DeepMind» en Inglaterra, que se emplea a nivel ambulatorio en el diagnóstico de la retinopatía en pacientes con diabetes.

Ahora bien, surge la siguiente interrogante: ¿Cuál es su procedencia? Hemos de precisar que la adquisición de dicho conocimiento no procede exclusivamente de los historiales médicos electrónicos, dispositivos de telemedicina o incluso de aplicaciones individualizadas (p. ej., *wearables*), como sería de esperar al garantizar la autonomía del paciente en la gestión de los mismos, sino que los mismos se difunden en Internet de forma heterogénea a través de múltiples canales, a menudo informales, sobre los que el paciente, en ausencia de una reglamentación específica, se encuentra indefenso en lo que concierne a su propia privacidad³. En investigación sanitaria, existe una clara jurisprudencia en relación al derecho a la privacidad de los datos personales, expresados en la Ley 14/2007 de Investigación Biomédica, que hace referencia a la protección del sujeto no solamente en su seguridad individual, sino en lo concerniente al tratamiento de las muestras biológicas y en su derecho al anonimato de sus datos personales. Este último aspecto es recogido más específicamente en la Ley Orgánica 2/2018 de Protección de Datos Personales y Garantía de los Derechos Digitales, de tal forma que en una investigación que

<https://doi.org/10.1016/j.semerg.2022.01.011>

1138-3593/© 2022 Sociedad Española de Médicos de Atención Primaria (SEMERGEN). Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

requiera el manejo de un gran número de datos, existen dos posibilidades para preservar la identidad del paciente: la anonimización, proceso por el cual se elimina de forma irreversible cualquier vinculación con su identidad, y la codificación, por el que los datos personales se identifican a través de un código («*encriptación*»), de tal forma que si se respetan dichos principios, se podrán realizar investigaciones sobre grupos poblacionales amplios sin necesidad de la obtención del preceptivo Consentimiento Informado de los mismos, siempre y cuando pueda acreditarse la imposibilidad de recabar el mismo o que su adquisición represente un trabajo tan desmesurado que pudiera llegar a paralizar el proceso de investigación

El reto que se plantea en la gestión de «*big data*» en su aplicación, tanto en asistencia sanitaria como en investigación, es asegurar, por una parte, la privacidad de los usuarios y, por otra, el respeto por el principio de autonomía de los mismos en cuanto a la cesión de datos referentes a su estado de salud. A tal fin, los gestores de «*big data*» habrán de asegurar que los datos que emplean son anónimos, esto es, que no contienen el nombre de la persona que los generó, y agregarse estadísticamente, lo que implica que no suponen un grupo de datos concretos sino un conjunto global de información, de ahí que sea imperativo el desarrollo de estrategias institucionales en el ámbito de la investigación biosanitaria que preserve el mencionado derecho del paciente sin renunciar a la explotación masiva de la información digital disponible de diversa procedencia, como puedan ser el empleo de distintas aplicaciones informáticas (redes sociales, *e-books*), compras *on-line*, facturación de determinadas fuentes energéticas, etc., que permitirían mejorar la precisión en la estratificación por riesgo de la población. Ejemplos relevantes de estas iniciativas se están desarrollando actualmente en el Reino Unido a través de la iniciativa «*care data*» del National Health Service para la apertura del acceso a los registros clínicos para fines de investigación, o la estrategia «*Big Data to Knowledge BD2K*» de infraestructura de los *National Institutes of Health* en los Estados Unidos, con la finalidad de evitar la adquisición y posterior comercialización de dichos datos por empresas privadas con ánimo de lucro.

Una vez preservado el derecho a la privacidad del paciente, este tipo de tecnologías representarán un

cambio trascendente no sólo en la calidad asistencial, sino en la relación médico-paciente al permitir realizar inferencias con medicina poblacional, preventiva, personalización de la asistencia... Partiendo de este modelo de referencia, el análisis y evaluación del «*big data*» permitirá la creación de sistemas algorítmicos en salud que permitan la gestión de toda la información para poder predecir el estado de salud de los pacientes y sus necesidades, constituyendo un apoyo en la práctica de los profesionales sanitarios, de tal forma que el paciente acudirá a la consulta con algún tipo de dispositivo que incluirá una base de datos coordinada en tiempo real con el historial clínico, y conectado con un sistema de cálculo logarítmico se generará respuestas individualizadas, base de una medicina personalizada⁴.

Bibliografía

1. Schneider F, Weiller C. Big Data und künstliche Intelligenz. *Nervenarzt*. 2018;89:859–60.
2. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MK, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature*. 2009;457:1012–4.
3. Price WN 2nd, Cohen IG. Privacy in the age of medical big data. *Nat Med*. 2019;25:37–43.
4. Benke K, Benke G. Artificial Intelligence and Big Data in Public Health. *Int J Environ Res Public Health*. 2018;15:2796.

J.D. Sánchez López^{a,*} y F. Luque Martínez^b

^a *Facultativo Especialista de Área de Cirugía Oral y Maxilofacial, H.U. Virgen de las Nieves, Granada (España), Presidente del Comité Ético de Investigación de Granada (España), Responsable de Calidad asistencial del Servicio de Cirugía Maxilofacial, España*

^b *Doctor en Farmacia, Responsable de Formación del H.U. Virgen de las Nieves de Granada (España), Vicepresidente del Comité Ético de Investigación de Granada, España*

* Autor para correspondencia.

Correo electrónico:
josed.sanchez.sspa@juntadeandalucia.es
(J.D. Sánchez López).