



Hacia un **cuidado de la salud** basado en los datos



Las siguientes páginas analizan con detalle de qué manera la explosión global del Big Data, la digitalización y el uso intensivo de modernas técnicas estadísticas de 'machine learning' pueden ayudar a transformar la industria de la salud, logrando que las personas vivamos vidas más largas y saludables con un menor coste para toda la sociedad

Juan José Casado

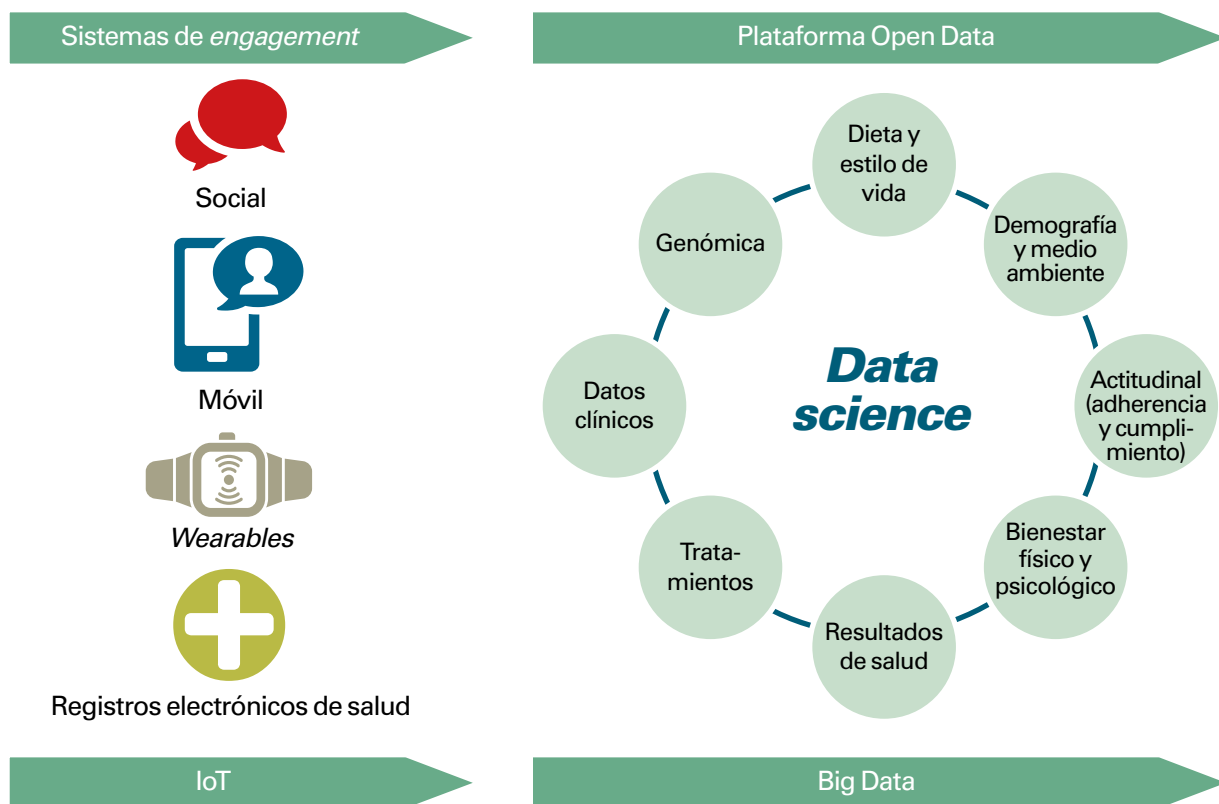
Academic Director del máster en Business Analytics & Big Data del IE Business School

El 31 de enero de 2015, Obama dio un discurso ante el mundo que abrió la puerta hacia un nuevo futuro de la industria de la salud, una nueva frontera marcada por el uso de las tecnologías digitales, el acceso a la información y nuevas políticas que permitan a los pacientes, los investigadores y los proveedores de salud trabajar juntos para el desarrollo de una medicina personalizada: “Quiero que mi país lidere una nueva Era de la Medicina, una que nos permita proporcionar el tratamiento adecuado, en el momento adecuado, al paciente adecuado. Esta noche voy a lanzar una nueva iniciativa llamada Medicina de Precisión para acercarnos más a la curación de enfermedades como el cáncer y la diabetes y para devolverle a cada americano el acceso a la información personalizada que necesitamos para mantener a nuestras familias saludables”. El mensaje de Obama

era claro: la industria de la salud debía abrirse a las posibilidades que las nuevas tecnologías brindan para conocer más en profundidad a los pacientes y ofrecer una medicina personalizada, garantizando la privacidad de la información, pero desafiando toda barrera cultural y legislativa alrededor de su uso. Para terminar de reforzar este mensaje, puso al frente de esta iniciativa, en lugar de a un médico, como cabría esperar, a un experto en Big Data, D. J. Patil, científico de datos, que hasta entonces había trabajado en compañías digitales como LinkedIn o eBay.

Pero, para hacer esta promesa realidad, tanto los gobiernos como las compañías relacionadas con la gestión de la salud, farmacéuticas, aseguradoras y sistemas de provisión médica tienen que adoptar una serie de iniciativas apoyadas por los profesionales sanitarios y la sociedad en su conjunto:

'FRAMEWORK' DE GESTIÓN DE LA SALUD BASADA EN DATOS



1. Uso de tecnologías digitales para lograr un mayor conocimiento del paciente.
2. Fomentar la accesibilidad de los datos en todo el ecosistema de salud para construir una visión 360° del paciente, única e integrada.
3. Aplicar el uso del Big Data en el proceso de investigación clínica y farmacéutica.
4. Aplicar medicina preventiva basada en datos, favoreciendo el automanejo del paciente.
5. Realizar una gestión poblacional que maximice los resultados de salud disminuyendo costes.
6. Aplicar modelos predictivos basados en datos digitales para el control epidemiológico.
7. Aplicar modelos predictivos en la práctica clínica diaria.
8. Aplicar inteligencia artificial para asistir a los profesionales en el diagnóstico.

USO DE TECNOLOGÍAS DIGITALES PARA LOGRAR UN MAYOR CONOCIMIENTO DEL PACIENTE

En palabras de Anne Wojcicki, CEO de 23&Me, la primera compañía genómica especializada en comercializar para el público en general estudios de ADN: “Es una locura que en este mundo digital donde tenemos registros médicos electrónicos, cualquier entidad financiera tenga toda la información acerca de lo que nos gusta comprar, pero nadie tenga toda la información sobre nuestra salud. ¿Cómo es posible que mi médico no pueda saber, basándose en mi estilo de vida y mis antecedentes clínicos, si me quedan cinco años para empezar a ser diabética y, sin embargo, Walmart sepa exactamente lo que voy a comprar ese día?”.

La realidad es que hoy, a través de los móviles, las redes sociales o el Internet de las Cosas, se podría conocer en profundidad a los pacientes y hacer realidad la promesa de una medicina completamente personalizada, basada en incorporar a la práctica clínica todos aquellos factores de las personas que pueden tener impacto en su salud y en el efecto que sobre ellos puede tener cada medicamento. Pero ¿cómo y qué dimensiones del ser humano se podrán conocer gracias al Big Data?

- **Datos genómicos.** No fue hasta que comenzaron a aplicarse las tecnologías del Big Data a la secuenciación del genoma humano que los costes pudieron reducirse hasta el precio objetivo de los mil dólares, que permite comenzar a aplicar cada vez con mayor frecuencia la información del ADN a la práctica clínica.

Hasta ahora, la mayoría de los tratamientos médicos habían sido diseñados para el “paciente medio”, siendo muy exitosos para algunos pacientes, pero no para otros. La comprensión genética de las personas es crucial para proporcionar una atención óptima para muchas enfermedades, ya que permite entender en qué variante genética tiene más éxito cada tratamiento.

Uno de los primeros laboratorios que comenzó a desarrollar fármacos genéticamente personalizados fue Genentech, con Herceptin, un fármaco que es eficaz en el tratamiento del cáncer de mama, pero solo en pacientes que tienen una composición genética en particular (HER2/neu). Para su ensayo, Genentech incluyó solo a 470 pacientes con esta composición genética, en vez de los 2.200 que se hubieran requerido en un ensayo habitual sobre el cáncer. Fue capaz de reducir la duración de la prueba de los habi- ➤➤➤

Medicina de precisión

Investigación clínica

Epidemiología digital

Salud preventiva

Salud predictiva

Diagnóstico asistido por ordenador

Data science



▶▶▶ tuales diez años a solo dos, demostrando su eficacia y permitiendo que 120.000 pacientes fueran capaces de obtener acceso a esta terapia, algo que, de otro modo, nunca se habría aprobado, por no haber obtenido las tasas de eficacia mínimas exigidas.

- **Datos de bienestar físico y psicológico.** Cada vez más datos personales de salud, como los de la actividad diaria, psicológicos y fisiológicos, pueden ser recogidos a través de dispositivos inalámbricos, sensores móviles, *wearables* o lo que, en general, denominamos el Internet de las Cosas. Estos aparatos se multiplican rápidamente, y dentro de cinco años podría haber un billón de dispositivos insertados en los cuerpos humanos. Hoy en día, los aparatos más conocidos son las pulseras de *fitness* y los relojes inteligentes, que monitorizan la cantidad de actividad física que realizamos cada jornada a través de un podómetro.

Pero ¿qué pueden medir y monitorizar estos nuevos dispositivos *wearables*, más allá de contar los pasos que damos cada día? Un oxímetro integrado en un anillo permitirá monitorizar la frecuencia cardíaca; la actividad muscular se puede recoger con un sensor electromiográfico incrustado en la ropa; el estrés se puede detectar con un sensor electrodermal incorporado en una pulsera, y los patrones de actividad o de sueño físico, a través de un acelerómetro en un reloj. Además, el período más fértil de una mujer puede ser identificado con un seguimiento detallado de la temperatura corporal, mientras que los niveles de

atención mental pueden ser monitorizados con un pequeño número de electrodos de electroencefalograma no gelificado (EEG). Los niveles de interacción social (también conocidos por afectar al bienestar general) pueden ser controlados usando detecciones de proximidad a los demás con Bluetooth o dispositivos con capacidad wifi... Estos son solo algunos ejemplos de los datos que hoy se pueden recoger a través de estos nuevos dispositivos y sensores generadores de Big Data.

- **Datos de hábitos, estilos de vida o relaciones sociales.** Las redes sociales pueden ser una fuente inagotable de conocimiento de las personas. El perfil de Twitter o de Facebook de cualquier persona revela toda una taxonomía de gustos, aficiones y estilos de vida. Sabiendo, simplemente, el perfil de Twitter de una persona se puede llamar a la API pública de la red y conocer a qué personajes famosos sigue, lo que dará una pista de sus gustos y aficiones. Además, algoritmos de reconocimiento de imágenes pueden ayudar a clasificar las fotos que cuelga una persona en su perfil de Instagram y deducir su estilo de vida: saber si le gusta viajar, hacer deporte o, simplemente, hacer vida familiar.

A través de analizar quiénes son los contactos de una persona en las redes sociales o, simplemente, sus llamadas telefónicas es sencillo determinar lo que se conoce como el grafo social de la persona o, lo que es lo mismo, con quién se interrelaciona y qué rol juega dentro de una comunidad. Esta información ha demostrado ser de gran importancia en la salud desde que, en los años 70, un importante estudio reveló que aquellos individuos con menor número de conexiones sociales tenían más probabilidad de verse afectados por enfermedades graves que los mejor conectados socialmente. Desde entonces, estas conexiones sociales se están utilizando para analizar la propagación de enfermedades epidemiológicas o para explicar la incidencia de patologías con un fuerte componente de hábito social, como se ha demostrado que es la obesidad. También se está utilizando para la detección geográfica de comunidades con actitudes de riesgo, como aquellas que fomentan la no vacunación de los hijos.

- **Datos de hábitos dietéticos.** Las enfermedades asociadas a la obesidad, la diabetes tipo 2, la hipertensión arterial, la enfermedad coronaria, el accidente cerebrovascular, la enfermedad de la vesícula biliar, la osteoartritis, la apnea del sueño, los problemas respiratorios y ciertos tipos de cáncer –incluyendo el cáncer de mama en las mujeres–, se ha demostrado que están fuertemente relacionados con nuestros hábitos alimenticios. Sabemos que somos lo que comemos y que

los alimentos actúan como medicina para mantener la salud o como catalizador para desencadenar determinadas patologías, pero lo cierto es que todavía no entendemos cómo, y mucho menos cuál, es la relación con otros factores como el estrés o la falta de ejercicio.

Compañías como Telspec han creado un dispositivo capaz de escanear los alimentos a nivel molecular, proporcionando información relevante y útil sobre lo que comemos para que en el día a día podamos tomar decisiones alimenticias que ayuden a mejorar la salud. Estas tecnologías permiten analizar alimentos, identificar calorías, macronutrientes, alérgenos, fibra y azúcares y también proporcionar información nutricional relevante. El éxito de estas soluciones no dependerá de la tecnología, sino de la sencillez con la que el usuario pueda llegar a desarrollar un cambio fundamental en la forma de comer diaria.

- **Datos del perfil psicológico.** El perfil psicológico de las personas es clave, ya no solo para monitorizar su estado de salud mental, sino para entender ciertos rasgos de la personalidad que puedan favorecer el desarrollo de determinadas patologías.

La práctica clínica habitual utiliza la evaluación de los rasgos de la personalidad del paciente para, por

para cada individuo desde sus valores hasta los cinco grandes rasgos de la personalidad, o *Big Five*: apertura, escrupulosidad, extroversión, simpatía y neurotismo.

- **Datos de adherencia a los medicamentos.** Algunos de los factores más importantes a la hora de que un tratamiento tenga éxito son la actitud y el compromiso del paciente con la medicación, lo que se conoce como adherencia al tratamiento, o, lo que es lo mismo, que el paciente tome su medicación siguiendo la pauta prescrita por el médico (que no abandone el tratamiento cuando lo considere oportuno o se salte tomas de manera habitual).

Durante años, los médicos y compañías farmacéuticas han tratado de medir la adherencia, pero nunca de una forma eficaz, pues es raro que un paciente reconozca este comportamiento.

Proteus, compañía de sensores dedicada al ramo de la salud, ha desarrollado un sensor ingerible que se puede introducir dentro de cada píldora y que, al ser consumido, detecta los jugos gástricos del estómago, enviando una señal a un parche adherido a la piel del paciente. Este parche recoge la hora exacta de cada ingesta, información que es volcada posteriormente a una app que puede ser compartida con familiares y

Cada vez más datos personales de salud, como los de la actividad diaria, psicológicos y fisiológicos, pueden ser recogidos a través de dispositivos inalámbricos, sensores móviles, 'wearables' o lo que, en general, denominamos el Internet de las Cosas

ejemplo, decidir entre distintos tratamientos en función de lo responsable y adherente al tratamiento que parezca dicho paciente o de su propia capacidad para afrontar el automanejo de una enfermedad crónica. Estos juicios de valor que se realizaban en función de la intuición o experiencia del profesional pronto podrán ser sustituidos por plataformas basadas en el Big Data, como la que IBM lanzó recientemente, System U, que es capaz de analizar datos generados a partir de la vida social digital de un individuo (ya sea en Facebook, Twitter, correos electrónicos o incluso chats) para determinar sus valores, creencias y rasgos de personalidad.

System U se basa en el estudio de la psicolingüística, una rama de la ciencia cognitiva que examina la forma en que se usa el lenguaje. Solo leyendo los tuits o los emails de una persona, la plataforma es capaz de definir

profesionales médicos para monitorizar la correcta toma de cada medicación.

- **Datos del contexto.** El auge de otro concepto muy pegado al Big Data, el de *smart cities*, ha permitido en muy poco tiempo que, aparte de los datos relativos a la persona, haya cada día una mayor cantidad de datos del contexto y del entorno geográfico en los que se sitúa ese individuo en cada momento. Estos nuevos datos pueden ser tremendamente útiles a la hora de analizar el impacto de determinadas variables medioambientales en la salud de las personas.

La capacidad de geolocalizar al individuo, sumada a la posibilidad de acceder a fuentes públicas de información, abre la oportunidad de entender cómo impactan en la salud factores del entorno tales como la temperatura, los niveles de polución, el >>>

➤➤➤ ruido, la presión ambiental, la fuerza del viento o, simplemente, el grado de humedad del ambiente.

FOMENTAR LA ACCESIBILIDAD DE LOS DATOS EN TODO EL ECOSISTEMA DE SALUD PARA CONSTRUIR UNA VISIÓN 360° DEL PACIENTE, ÚNICA E INTEGRADA

Para que todo lo anterior sea posible, los investigadores tienen que ser capaces de integrar los datos de los registros de salud con los de los ensayos clínicos y con datos fisiológicos genómicos. Y, para ello, es clave contar con la colaboración voluntaria del paciente. Cada uno de nosotros puede convertirse en un productor de Big Data, de manera que los datos que generamos lleguen a superar en el futuro los que se acumulan en la atención clínica. En temas como la nutrición, el ejercicio, el tabaquismo o el consumo de alcohol, los mejores datos son los que proporcionan los propios pacientes.

En Dinamarca, por ejemplo, a cada recién nacido se le asigna un número de identificación único que rastrea sus interacciones de toda la vida con el sistema de atención de salud gratuita del país y casi todos los otros servicios del gobierno. En 2002, los investigadores utilizaron datos recogidos a través de este sistema de identificación para analizar retrospectivamente el estado de vacunación y la salud de casi todos los niños nacidos entre 1991 y 1998 (537.000 en total). En ese momento era el mayor estudio jamás realizado sobre vacunación, y sirvió para desacreditar la corriente que relacionaba la vacunación para el sarampión y el autismo, demostrando que no existía tal relación.

APLICAR EL USO DEL BIG DATA EN EL PROCESO DE INVESTIGACIÓN CLÍNICA Y FARMACÉUTICA

Los dos grandes sistemas operativos móviles han lanzado sendas plataformas para ayudar a los investigadores a recopilar datos sobre la salud de un gran número de personas: Google FIT y ResearchKit, diseñados específicamente para favorecer la investigación médica.

A los quince días de lanzar Apple esta nueva plataforma, seis mil personas ya habían realizado el test de caminar a través de una nueva app llamada MyHeart Counts. Pese a ser uno de los test más populares y haberse realizado diversos estudios históricamente, ninguno contaba con más de mil pacientes. Cuatro meses después, más de cuarenta mil personas en Estados Unidos se habían inscrito para participar en el estudio.

Los médicos saben que la actividad física es un fuerte predictor de la salud del corazón a largo plazo, pero no está tan claro qué tipo de actividad es la mejor, o si



afecta por igual a los distintos subgrupos de personas. Sin embargo, el volumen de datos que facilitan estas nuevas tecnologías digitales hace posible este tipo de estudios, ya que en las investigaciones tradicionales puede que no haya suficientes datos para encontrar resultados estadísticamente significativos.

APLICAR MEDICINA PREVENTIVA BASADA EN DATOS, FAVORECIENDO EL AUTOMANEJO DEL PACIENTE

El uso de sensores, *wearables* y otros dispositivos móviles facilitará que cada paciente recoja datos relativos a su enfermedad de manera pasiva y sin apenas esfuerzo, lo que permitirá a los pacientes crónicos tener un control constante y un mayor conocimiento sobre el grado de desarrollo y los efectos de su patología.

Por ejemplo, es posible identificar la gravedad de los síntomas depresivos en función del número de conversaciones, la cantidad de actividad física y la duración del sueño utilizando tan solo una pulsera portátil y un smartphone. Y la apnea del sueño podría ser diagnosticada rápidamente –y la calidad del sueño mejorada– con un portátil ligero que mide el ritmo cardíaco, el volumen de la respiración y el ronquido (a través de la vibración del tejido), en lugar de tener que recurrir a realizar una polisomnografía en el hospital.

En esta misma línea, han surgido múltiples plataformas encaminadas a agregar los datos necesarios para facilitar el seguimiento y control de la diabetes, como, por ejemplo, Databetic. El funcionamiento de



este tipo de plataformas es relativamente sencillo, ya que permite importar automáticamente los datos sobre los niveles de azúcar en la sangre procedentes de los monitores de glucosa más usados, como, por ejemplo, DEXCOM CGM. También se podrán importar datos relativos al ejercicio físico, actividad e incluso alimentos ingeridos a través de un módulo que permite crear una memoria fotográfica de comidas y analizar el impacto en los niveles de glucosa en sangre de cada alimento que se ingiere. La idea es que el paciente pueda llevar una vida normal, dedicando el menor tiempo a la gestión de su enfermedad, pero ganando en control y capacidad de anticipación. Estos datos, además, pueden ser compartidos con el médico, que no necesita programar visitas periódicas, sino hacer un seguimiento reactivo, citando al paciente a consulta solo cuando es verdaderamente necesario.

REALIZAR UNA GESTIÓN POBLACIONAL QUE MAXIMICE LOS RESULTADOS DE SALUD DISMINUYENDO COSTES

La gestión poblacional de la salud hace referencia al proceso de optimizar la salud de un colectivo de individuos minimizando los costes asociados. Ello tiene un triple objetivo: mejorar la experiencia y resultados de salud de cada individuo, reducir el coste per cápita y mejorar la salud de la población en su conjunto.

Existen múltiples tipos de colectivos en los que puede aplicarse, ya sean geográficos, políticos, sociales, raciales, o, simplemente, el colectivo de empleados de una misma empresa, por ejemplo.

Para realizar correctamente esta gestión es imprescindible contar con un sistema unificado de historias clínicas electrónicos debidamente digitalizados y estructurados que permitan explotar todos los datos disponibles de cada individuo.

La gestión poblacional usa modelos analíticos y estadísticos avanzados para clasificar a cada persona en un segmento en función del riesgo asociado a su salud, y establece diferentes niveles de cuidado y atención según estos riesgos, desde la prevención, el autocuidado o la telemedicina hasta la gestión de casos personalizada.

Una de las organizaciones más reconocidas por su gestión poblacional basada en explotación de su sistema de Big Data es la compañía americana Kaiser Permanente, que opera en la costa oeste americana y que tiene uno de los mayores registros de pacientes, con nueve millones de personas y diez *petabytes* de información. La compañía es capaz, mediante analítica de datos, de diseñar una calculadora en tiempo real para reducir el riesgo de sepsis en sus neonatos, descubrir que el consumo de un popular medicamento como Vioxx está triplicando el riesgo de infarto en su población o encontrar una correlación entre un número excesivo de afecciones cardiovasculares en un área geográfica y la escasez de metros cuadrados de zonas verdes (y cofinanciar la construcción de parques junto con las autoridades locales para revertir este último problema).

APLICAR MODELOS PREDICTIVOS BASADOS EN DATOS DIGITALES PARA EL CONTROL EPIDEMIOLÓGICO

En un mundo digital e interconectado como el actual, se pueden integrar un gran número de fuentes de datos, incluyendo los datos de los operadores de telefonía móvil o de las redes sociales, para modelar y predecir la propagación de múltiples enfermedades infecciosas.

En los últimos años, siempre que tiene lugar un desastre natural –como el ocurrido tras el terremoto de Haití de 2010, en el que se produjo un brote de cólera en todo el país–, científicos analizan los registros telefónicos para predecir hacia qué zonas geográficas se van a desplazar las poblaciones afectadas y enviar allí la ayuda humanitaria y médica necesaria para minimizar el impacto de esos brotes.

Igualmente, las redes sociales han demostrado ser un predictor excelente para anticipar brotes de enfermedades como el asma o la gripe: monitorizando el número de personas que mencionan en cada momento estas enfermedades se puede anticipar cuándo va surgir un brote y en qué poblaciones, a la vez que se puede monitorizar su extensión. >>>

APLICAR MODELOS PREDICTIVOS EN LA PRÁCTICA CLÍNICA DIARIA

En el entorno hospitalario hay una enorme cantidad de datos críticos que se producen, generalmente, durante la monitorización de los pacientes. Las nuevas tecnologías de Big Data y las técnicas de computación de procesamiento en *streaming* permiten implantar algoritmos que detecten en tiempo real un posible riesgo para el paciente. La aplicación del análisis predictivo en tiempo real permite a los médicos tomar medidas para detectar los efectos adversos antes de que sucedan.

Es el caso del hospital de Ontario para niños enfermos, que desarrolló un sistema capaz de monitorizar en tiempo real más de mil doscientas cincuenta lecturas por segundo de cada uno de los neonatos bajo vigilancia. La aplicación de modelos predictivos sobre

conocimiento necesario para realizar diagnósticos y ofrecer recomendaciones de tratamiento a personas diagnosticadas con cáncer.

Tras incorporar más de medio millón de pruebas médicas, alrededor de dos millones de páginas de texto de investigación sobre cáncer de mama y de pulmón y los registros de un millón y medio de pacientes diagnosticados, los expertos enseñaron a Watson cómo procesar, analizar e interpretar el significado de toda esa información. Un esfuerzo de meses que, sin duda, ha merecido la pena y ha cambiado el rumbo de la medicina para siempre.

Actualmente, Watson ha demostrado ser capaz de diagnosticar correctamente el 90% de los casos de cáncer de pulmón, mientras que, de media, un especialista en oncología solo es capaz de diagnosticar correctamente el 50%.

Que una decisión correcta puede suponer la diferencia entre la vida y la muerte para un paciente es lo que quizá debió de pensar IBM cuando, tras desarrollar Watson y ganar el concurso de preguntas y respuestas más popular de la TV americana, decidió afrontar el reto de enseñarle oncología. Watson procesa, analiza e interpreta el significado de grandes cantidades de información

datos recolectados en períodos de varios días ha demostrado ser un importante predictor de enfermedades como la sepsis, y está suponiendo la diferencia entre la vida y la muerte para muchos niños canadienses.

FOMENTAR EL DIAGNÓSTICO ASISTIDO POR COMPUTACIÓN

Que una decisión correcta puede suponer la diferencia entre la vida y la muerte para un paciente es lo que quizá debió de pensar IBM cuando, tras desarrollar Watson y ganar el concurso de preguntas y respuestas más popular de la TV americana, decidió afrontar el reto de enseñarle oncología.

La capacidad de procesamiento de datos de Watson es de ochenta *teraflops*, y para igualar o superar la capacidad de respuesta humana se conecta a noventa servidores con información de doscientos millones de páginas, que procesa mediante seis millones de reglas lógicas. Todo en un espacio igual al que ocuparían diez neveras. Pero el poder de Watson radica en su capacidad de aprendizaje. Un año después, y en colaboración con el prestigioso Memorial Sloan Kettering Cancer Center, Watson ha desarrollado el

CONCLUSIÓN

Es razonable esperar que en un futuro muy cercano, y con la implantación masiva del Big Data y de las más modernas técnicas analíticas, los seres humanos puedan aprender a gestionar preventivamente su salud, encontrar tratamientos más eficaces para curar enfermedades que hoy en día no se pueden curar y, en general, vivir más tiempo y de manera más saludable.

Lo importante es entender que las grandes barreras ya no son tecnológicas o de falta de información, sino que son simplemente culturales, y que cuando los gobiernos y las organizaciones relacionadas con el mundo de la salud tomen la firme decisión de adoptar plenamente la ciencia de los datos, el futuro cambiará para todos. ■