



EDITORIAL

Dr. Javier Mora
Editor invitado

En salud, hablar de ciencias resulta cotidiano. Gran parte de la práctica clínica sienta sus bases en la “medicina basada en la evidencia”, consolidada recién a fines de los años 90. La formación de pre y post grado en salud contempla siempre, en mayor o menor medida, cátedras de metodologías de la investigación, bioestadística, epidemiología clínica, lectura crítica, entre otros. Lo que sustenta una recomendación “tipo A”, mediante la revisión sistemática de literatura de alta calidad, es la evidencia a través de ensayos clínicos aleatorizados y ciegos. En ellos, para obtener significancia estadística que explique un fenómeno, se requiere de un elemento en común: los datos.

El método científico aplicado en medicina contempla el plantearse primero una pregunta de investigación que demuestre o rechace una hipótesis, para luego buscar los datos que metodológicamente puedan responder a aquella pregunta de forma adecuada. Actualmente, con la cantidad de datos acumulados y la capacidad computacional moderna, podemos obtener nueva y valiosa información mediante una metodología inversa, es decir, “desde los datos”. Dejar que los datos hablen, encontrar patrones de distribución o asociación no conocidos previamente, automatizar procesos complejos que a una persona le tardaría una excesiva cantidad de tiempo y recursos poder realizar, obtener sistemas de clasificación, predicción y recomendación con un alto nivel de asertividad, son algunas de las tareas que se logra a través de la Ciencia de Datos.

Pero hablar de Ciencia de Datos o de Ciencias de la Computación no resulta familiar en el ámbito clínico. Incluso en ambientes académicos es aún difícil o poco difundido mirar a la computación desde una perspectiva científica. Acostumbrados a utilizar los computadores como una herramienta eminente-

mente de registro, y de visualización de exámenes e imágenes médicas, se suele pasar por alto el inmenso valor de un repositorio o *corpus* de datos clínicos; es decir, de un gran volumen de datos. Y en nuestra región, más aún, en español.

Quienes de alguna u otra manera hemos podido darnos cuenta y estudiar el valor agregado del “procesamiento de datos”, estamos llamados a hacer público el trabajo, investigación y resultados que con las técnicas de la Ciencia de Datos podemos obtener. De esta forma, en conjunto con todos los autores invitados, presentamos este número de la Revista Médica Clínica Las Condes dedicado a la Ciencia de Datos Clínicos, tema inusual en una revista científica médica en español.

Revisamos cómo hemos ido evolucionando en algunos aspectos de la tecnología en salud hasta la Inteligencia Artificial de nuestros días, abarcando aspectos técnicos, prácticos y éticos. Recorreremos también cómo la salud llega a estar inmersa en la era digital que estamos viviendo a través de importantes avances que van siendo cada vez más habituales en nuestro quehacer diario. Al hablar de aplicaciones de aprendizaje automático en salud no sólo se muestran relevantes investigaciones en distintos campos clínicos, sino que además se explican conceptos generales para comprender mejor la Ciencia de Datos Clínicos, en un lenguaje académico y abordable para los profesionales de la salud no habituados a estas terminologías ni formas de ver los datos y técnicas computacionales.

Dado su alto impacto, dedicamos un apartado especial al procesamiento de lenguaje natural (PLN) para texto clínico en español; técnica que en otras áreas del conocimiento está logrando excelentes resultados en traducción de idiomas,

aplicativos de mensajería automática, reconocimiento de voz, construcción de texto nuevo con sentido en un contexto a partir del ingreso de pocas palabras, etc, y que en salud ha permitido rescatar la inmensa cantidad de texto libre y datos altamente no estructurados de las fichas clínicas que otrora no se utilizaban por no estar ordenados en tablas. De tal modo se logra adjudicar valor local y regional a las investigaciones que podamos realizar con nuestros propios datos.

Para lograr automatizar procesos prescriptivos de salud desde el procesamiento descriptivo de datos, es necesario pasar por modelos predictivos. Mostramos como las técnicas de aprendizaje de máquina (o aprendizaje automatizado), al procesar grandes volúmenes de datos y con funciones no lineales, logran mejor rendimiento que la regresión lineal tradicionalmente utilizada para predecir resultados. Si bien en las ciencias de la salud el objetivo final de la investigación es poder llegar a la causalidad de los fenómenos, el lograr realizar tareas de clasificación y predicción con alto nivel de acierto mediante la asociación (extracción) de las características de todos los datos permite disminuir los sesgos que conlleva la elección arbitraria de variables que se utilizan para el cálculo de puntajes de riesgo. Paralelamente, esto conlleva la posibilidad de contar con modelos predictivos con datos locales, y no necesariamente aplicar los modelos elaborados en otras regiones, provenientes de países desarrollados con condiciones muy diferentes a nuestras realidades.

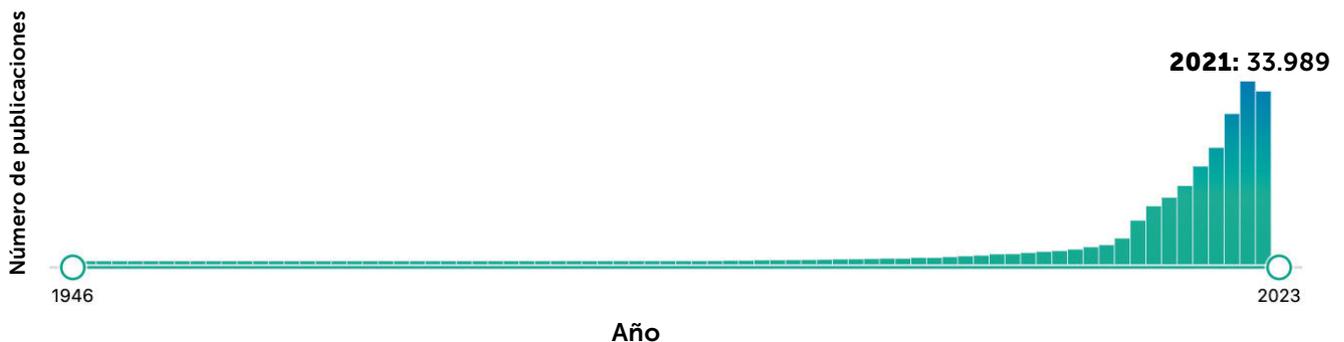
Estas tecnologías, y otras no incluidas en estos apartados, no tendrían sentido si no llegamos a aplicarlas en la población de nuestros países latinoamericanos. Por eso abordamos también,

los desafíos y oportunidades que enfrentamos en nuestra región, con las características propias de nuestras culturas y los escasos recursos disponibles, donde juega un rol preponderante las aplicaciones y librerías de programación de código abierto. Para este objetivo también, es vital incorporar a todas las profesiones, disciplinas y áreas de la salud en este esfuerzo; en esta oportunidad revisamos el rol de la enfermería en la salud digital desde la transformación del cuidado a través de la tecnología centrada en la persona. Es así como podemos llegar a proyectar la inteligencia artificial al servicio de la medicina del futuro.

Para todo esto indudablemente se requiere alto poder computacional. Afortunadamente, estamos transitando desde el paradigma de centrarse en la infraestructura, el hardware y el *software*, a enfocarse en lo verdaderamente importante: lo que se hace con los datos utilizando dichos recursos. Desde las bases teóricas de la máquina universal de Alan Turing (la primera concepción de un computador) en la primera mitad del siglo XX, a la computación actual en la nube, el cómputo de datos está simplificando y facilitando el camino para que más personas nos podamos acercar al desarrollo de la Ciencia de Datos Clínicos y que sus beneficios lleguen a la mayor cantidad de población posible.

Gracias al significativo aporte y nivel de trabajo de cada uno de los autores, revisores y al Comité Editorial, los artículos abordados en el presente número de la Revista Médica de Clínica Las Condes son el constructo inicial de una emergente área del conocimiento que en efecto debemos impulsar y desarrollar en forma conjunta y colaborativa en beneficio de la salud pública y privada de nuestras sociedades.

Resultado de búsqueda de palabras clave "clinical data science" en National Library of Medicine



<https://pubmed.ncbi.nlm.nih.gov/?term=clinical+data+science>