



Limitaciones y desafíos de la inteligencia artificial en salud

Limitations and challenges of artificial intelligence in health

Iñigo Gabilondo Cuéllar

Servicio de Neurología, Hospital Universitario de Cruces, IIS Biocruces Bizkaia, Barakaldo, Bizkaia, España

Las potenciales aplicaciones de la inteligencia artificial (IA) en la investigación y la práctica clínica neurológicas están creciendo de manera exponencial. La IA puede ayudar a mejorar la clasificación diagnóstica y pronóstica, la monitorización y la toma de decisiones en neurología, y puede facilitar la identificación de mecanismos fisiopatológicos y el modelado del funcionamiento cerebral, así como la individualización terapéutica y la medicina personalizada¹. Para ello, los algoritmos de IA están utilizando datos multimodales que incluyen información de historias clínicas electrónicas, escalas clínicas, imágenes médicas, registros neurofisiológicos, datos procedentes de sensores portátiles o ambientales y datos de muestras biológicas (proteómica, genómica, metabolitos o microbioma)². La IA también se está utilizando para el desarrollo de prótesis neurales e interfaces cerebro-computadora que pueden restaurar la función en personas que han sufrido daño cerebral o medular³. Por último, la IA también se puede utilizar con el objetivo de optimizar la gestión de pacientes y los procesos asistenciales, permitiendo, por ejemplo, mejorar los protocolos de adquisición de datos como las imágenes médicas (reducción de tiempos de adquisición, detección automática de artefactos y anomalías, y reducción de dosis de contraste y de radiación)⁴.

Además de las ventajas evidentes, todos estos avances y aplicaciones de la IA en neurología y en salud tienen importantes limitaciones y riesgos a diferentes niveles: éticos, clínicos, socioeconómicos, científicos y tecnológicos. En primer lugar, la IA en salud no puede reemplazar la ética médica ni los juicios clínicos basados en la experiencia del médico en situaciones complejas o impredecibles⁵. La IA tampoco puede proporcionar una respuesta emocional o empática a los pacientes⁶, aspecto clave en la relación de éstos con sus médicos. Además, dado que el rendimiento de la IA depende en gran medida de la cantidad y la calidad de los datos de entrenamiento que utiliza, si éstos no son representativos del ámbito clínico, socioeconómico o epidemiológico de interés, si son de mala calidad o si son incompletos, sus resultados pueden estar sesgados y ser poco fiables. Esto puede implicar incluso sesgos discriminatorios desde el punto de vista socioeconómico, racial, religioso o de género⁷.

Por otro lado, la IA en salud también se enfrenta a desafíos relacionados con la interpretación de sus resultados. Aunque los algoritmos de IA pueden analizar grandes cantidades de datos y generar patrones, la propia complejidad de los datos y la naturaleza opaca de algunos algoritmos de IA, especialmente los de *deep learning*, pueden hacer difícil para los expertos humanos comprender cómo se llegó a una determinada conclusión, convirtiendo los resultados de la IA en una «caja negra»⁸. Esto hace que sea casi imposible aprender de las decisiones de la IA y además dificulta que

Autor de correspondencia:
Iñigo Gabilondo Cuéllar
E-mail: igabilon@gmail.com

Fecha de recepción: 01-04-2023
Fecha de aceptación: 11-04-2023
DOI: 10.24875/KRANION.M23000052

Disponible en internet: 13-06-2023
Kranion. 2023;18:43-4
www.kranion.es

se genere confianza en el sistema mismo, lo que puede ser particularmente problemático en situaciones de salud críticas, cuando las decisiones deben tomarse rápido⁹. Para evitar estas situaciones, será necesario comparar los resultados de los algoritmos de IA más complejos con modelos interpretables, realizar estudios *post-hoc* para mejorar dicha interpretabilidad, hacer controles de calidad y documentar los detalles relevantes para el desarrollo de los modelos de IA en tarjetas estándar que incluyan información sobre conjuntos de datos, algoritmos, estudios de validación y pruebas de sesgo¹⁰. Otra de las limitaciones metodológicas de la IA son los errores relacionados con el sobreajuste (*overfitting*), que surgen como consecuencia de usar más variables de las necesarias, a menudo como parte de un modelo que es más flexible de lo necesario, ocasionando predicciones falsamente elevadas que probablemente no se puedan replicar ni implementar¹¹.

La IA en salud requiere tamaños muestrales sumamente grandes, de más de 100.000 sujetos. Obtener este volumen de datos no será posible si no se fomentan estándares en la industria sanitaria y el intercambio de datos¹². Del mismo modo, otro desafío importante de la IA en salud es la interoperabilidad de los sistemas. La IA puede requerir el acceso a datos de múltiples sistemas, incluyendo registros médicos electrónicos, imágenes médicas y datos de sensores. Si estos sistemas no son interoperables, es decir, si no pueden intercambiar datos de manera efectiva, puede ser difícil para los algoritmos de IA analizarlos y generar resultados precisos¹³. Por ello, los algoritmos de IA deberían diseñarse con una visión amplia, utilizando principios de diseño técnico sólidos centrados en su generalización, eficiencia, reutilización, modularidad, escalabilidad y solidez. Por otro lado, la IA no puede garantizar la privacidad ni la seguridad de los datos de salud del paciente si se utiliza de manera inapropiada o si hay una brecha en la seguridad¹⁴. En este sentido, el primer conjunto de preocupaciones incluye el acceso, el uso y el control de los datos de los pacientes en manos privadas. Otro conjunto de preocupaciones se relaciona con el riesgo externo de violaciones de la privacidad mediante métodos impulsados por IA. La capacidad de desidentificar o anonimizar los datos de salud de los pacientes puede verse comprometida e incluso anulada a la luz de los nuevos algoritmos de IA, capaces de reidentificar dichos datos.

Un componente fundamental para lograr un despliegue seguro y eficaz de los algoritmos de IA es el desarrollo de los marcos normativos necesarios. Esto plantea un desafío único dado el ritmo actual de innovación, los riesgos significativos involucrados y la naturaleza potencialmente fluida de los modelos de aprendizaje automático. También es importante considerar el impacto regulatorio de las mejoras y las actualizaciones que los proveedores de productos de IA probablemente desarrollen durante la vida útil del producto. Algunos sistemas de IA se diseñarán para mejorar con el tiempo, lo que representa un reto para los procesos de evaluación y regulación tradicionales.

Incluso con algoritmos de IA altamente efectivos que superen todos los obstáculos anteriores, las barreras humanas para la adopción de la IA en la práctica clínica y en la investigación en neurología son sustanciales. Para garantizar que esta tecnología pueda llegar a los pacientes y beneficiarles será importante mantener un enfoque en la aplicabilidad clínica y en los resultados de los pacientes, desarrollar métodos avanzados para la interpretabilidad e interoperabilidad algorítmica, y lograr una mejor comprensión de las interacciones humano-computadora.

BIBLIOGRAFÍA

1. Hillis JM, Bizzo BC. Use of artificial intelligence in clinical neurology. *Semin Neurol.* 2022;42:39-47.
2. Acosta JN, Falcone GJ, Rajpurkar P, Topol EJ. Multimodal biomedical AI. *Nat Med.* 2022;28:1773-84.
3. Zhang X, Ma Z, Zheng H, Li T, Chen K, Wang X, et al. The combination of brain-computer interfaces and artificial intelligence: applications and challenges. *Ann Transl Med.* 2020;8:712.
4. Ranschaert E, Topff L, Panykh O. Optimization of radiology workflow with artificial intelligence. *Radiol Clin North Am.* 2021;59:955-66.
5. Harish V, Morgado F, Stern AD, Das S. Artificial intelligence and clinical decision making: the new nature of medical uncertainty. *Acad Med.* 2021;96:31-6.
6. Kerasidou A. Artificial intelligence and the ongoing need for empathy, compassion and trust in healthcare. *Bull World Health Organ.* 2020;98:245-50.
7. Zou J, Schiebinger L. AI can be sexist and racist — it's time to make it fair. *Nature.* 2018;559:324-6.
8. Wang F, Kaushal R, Khullar D. Should health care demand interpretable artificial intelligence or accept "black box" medicine? *Ann Intern Med.* 2020;172:59-60.
9. Chang H, Cha WC. Artificial intelligence decision points in an emergency department. *Clin Exp Emerg Med.* 2022;9:165-8.
10. Savage N. Breaking into the black box of artificial intelligence. *Nature.* 2022 Mar 29. doi: 10.1038/d41586-022-00858-1. Online ahead of print.
11. Volovici V, Syn NL, Ercole A, Zhao JJ, Liu N. Steps to avoid overuse and misuse of machine learning in clinical research. *Nat Med.* 2022;28:1996-9.
12. Wang SY, Pershing S, Lee AY; AAO Taskforce on AI and AAO Medical Information Technology Committee. Big data requirements for artificial intelligence. *Curr Opin Ophthalmol.* 2020;31:318-23.
13. Lehne M, Sass J, Essenwanger A, Schepers J, Thun S. Why digital medicine depends on interoperability. *NPJ Digit Med.* 2019;2:79.
14. Murdoch B. Privacy and artificial intelligence: challenges for protecting health information in a new era. *BMC Med Ethics.* 2021;22:122.