

Abordaje editorial y ético en la revisión y calificación de investigación médica basada en modelos de inteligencia artificial

Editorial and ethical approach in reviewing and rating medical research based on artificial intelligence models

Luis Pino^{1,2*} e Iván Triana^{1,2}

¹O^xLER SAS Medicina Exponencial; ²Universidad de los Andes-Fundación Santa Fe de Bogotá. Bogotá, Colombia

La inteligencia artificial (IA) ha emergido como una herramienta revolucionaria en la investigación médica permitiendo avances significativos en el diagnóstico, tratamiento y gestión de enfermedades. A medida que aumenta el número de estudios que utilizan modelos y desarrollos basados en IA, los comités editoriales de revistas médicas y los comités éticos de investigación en universidades y hospitales enfrentan nuevos desafíos para evaluar la calidad, la ética y la relevancia de estas investigaciones, lo cual ha sido poco mencionado al compararlo con la gran revolución de nuevos estudios clínicos basados en estas herramientas¹. Entre las consideraciones éticas y editoriales que se proponen están la privacidad de los datos, la transparencia de los algoritmos, la calidad y validez de los datos, los posibles sesgos en los sistemas de IA y la responsabilidad sobre el uso de los resultados obtenidos¹. En este editorial, la Revista Colombiana de Cardiología explorará estos aspectos éticos, proponiendo un enfoque equilibrado que maximice los beneficios de la IA mientras se minimizan sus riesgos. De igual forma, propenderá por una mejor interpretación y análisis de los trabajos sometidos a publicación que se basen en modelos de IA, para el contexto colombiano.

Calidad de los datos

Los algoritmos de IA dependen de grandes volúmenes de datos para «aprender» y hacer predicciones precisas. Por lo tanto, es crucial que los comités editoriales verifiquen la calidad de los datos utilizados. Esto involucra evaluar la fuente de los datos, su representatividad y las técnicas ejecutadas en la metodología para solucionar problemas comunes, como datos faltantes y datos extremos o poco balanceados (preprocesamiento e ingeniería de variables). Adicionalmente, consideramos que este proceso debe tener una vigilancia centralizada por el autor o responsable del producto final del trabajo al momento de tratarse de sistemas independientes o sin revisión humana, como lo pueden ser los registros hospitalarios basados en IA. Por lo tanto, creemos que los comités editoriales deben exigir a los autores que proporcionen detalles sobre la procedencia de los datos y las medidas tomadas para asegurar su calidad, incluyendo una descripción de los métodos de preprocesamiento e ingeniería de variables en modelos de IA analítica^{2,3}.

*Correspondencia:

Luis Pino
E-mail: luispino@oxler.me

Fecha de recepción: 20-09-2024
Fecha de aceptación: 02-10-2024
DOI: 10.24875/RCCAR.M24000231

Disponible en internet: 28-11-2024
Rev Colomb Cardiol. 2024;31(5):265-267
www.rccardiologia.com

0120-5633 / © 2024 Sociedad Colombiana de Cardiología y Cirugía Cardiovascular. Publicado por Permanyer. Este es un artículo *open access* bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Transparencia y explicabilidad de los algoritmos

La transparencia en el desarrollo y el funcionamiento de los algoritmos de IA es crucial para la confianza y validez de los resultados de la investigación. Los comités editoriales deben insistir en que los autores expliquen claramente cómo se han desarrollado y entrenado sus modelos de IA, incluyendo las arquitecturas utilizadas, los parámetros ajustados y los criterios de evaluación. Esto es fundamental para abordar el problema de la «caja negra» asociado a la IA, como los algoritmos de entrenamiento profundo (*deep learning*, DL), en los cuales el comité evaluador o el usuario final pueden no entender la estructuralidad del funcionamiento del algoritmo, lo que, a su vez, genera un uso inadecuado de las salidas del mismo, de alto impacto en la toma de decisiones clínicas avaladas por el algoritmo¹. Lipton destaca la importancia de la explicación correcta y de la descripción en los sistemas de IA para asegurar que los resultados sean interpretables y verificables⁴. En consecuencia, los comités editoriales deben promover la publicación de modelos transparentes y explicables, así como exigir a los autores que proporcionen explicaciones comprensibles y claras del cómo sus algoritmos alcanzan determinados resultados usando métricas de desempeño validadas en los modelos de IA analítica y descripción de las técnicas de ajuste fino utilizadas en la IA generativa^{4,5}.

Validación de resultados

Este es un aspecto que hemos conocido previamente en los diferentes estudios de pruebas diagnósticas, o en los ensayos clínicos, y es que toda prueba diagnóstica, o ensayo clínico requiere de una adecuada validez interna y externa para que sus resultados sean aplicables a la población general. En este sentido, los estudios basados en IA no deben ser excluidos de este aspecto y, por consiguiente, requieren de una validación rigurosa de los resultados y de sus desenlaces, especialmente en su capacidad para ser reproducidos en la población objeto. En IA a este concepto se le conoce como capacidad de generalización o industrialización del modelo. Se considera que los comités editoriales deben evaluar si los estudios han utilizado métodos adecuados para validar sus modelos, como la validación cruzada y el uso de controles únicos o poblaciones para comparar sus resultados, así como la forma de aleatorización, muestreo o estratificación utilizada.

Así mismo, se debe fomentar la publicación de código y conjuntos de datos, permitiendo a otros investigadores reproducir y verificar los resultados³. Existen diferentes matrices de evaluación para sistemas de IA, como DECIDE-AI, que pretenden ser generalizados y que pueden apoyar a los comités editoriales en esta evaluación objetiva. Esto se comentará más adelante.

Consideraciones éticas (consentimiento informado, privacidad de los datos)

Los aspectos éticos son especialmente relevantes en cualquier investigación médica; por lo tanto, siguen siendo fundamentales en los estudios basados en IA. Uno de los principales desafíos es la protección de la privacidad y la confidencialidad de los datos de los pacientes. Como se mencionó previamente, la IA requiere grandes volúmenes de datos para entrenar sus algoritmos, y estos datos a menudo contienen información sensible y personal, por lo que requieren de un sistema de protección de datos adecuado, el cual debe ser específico y descrito para cada tipo de estudio^{1,5}.

Por lo anterior consideramos que los comités editoriales deben evaluar si los estudios han abordado correctamente las cuestiones éticas relacionadas con la privacidad y anonimización de los datos, así como con el consentimiento informado, dado que deben asegurarse de que los pacientes hayan sido informados de cómo se utilizarán sus datos, y de que entiendan los riesgos y beneficios en un ensayo clínico basado en IA. Adicionalmente, los comités deben estar atentos a los posibles sesgos en los algoritmos de IA, debido a que los modelos entrenados con datos sesgados pueden arrojar resultados erróneos, por lo que se debe solicitar a los autores específicamente el plan de mitigación de sesgos del estudio^{1,5}.

En este mismo apartado ético, en especial para los comités de ética hospitalarios, consideramos importante mencionar que la IA no tiene una representación legal para responder por los resultados obtenidos, y el único responsable de un error o un beneficio será la institución o el personal médico que use la IA, así como el grupo investigador/desarrollador, de modo que, en los estudios de IA donde se tomarán decisiones terapéuticas con base en resultados de analítica, se deben desarrollar marcos regulatorios que establezcan responsabilidades claras para los desarrolladores y usuarios de las herramienta de IA^{6,7}.

Para el investigador es muy importante recordar que la IA no es considerada dentro del componente de autor, dado que no completa autoridad legal ni ética, por lo que los autores de un manuscrito deben reportar

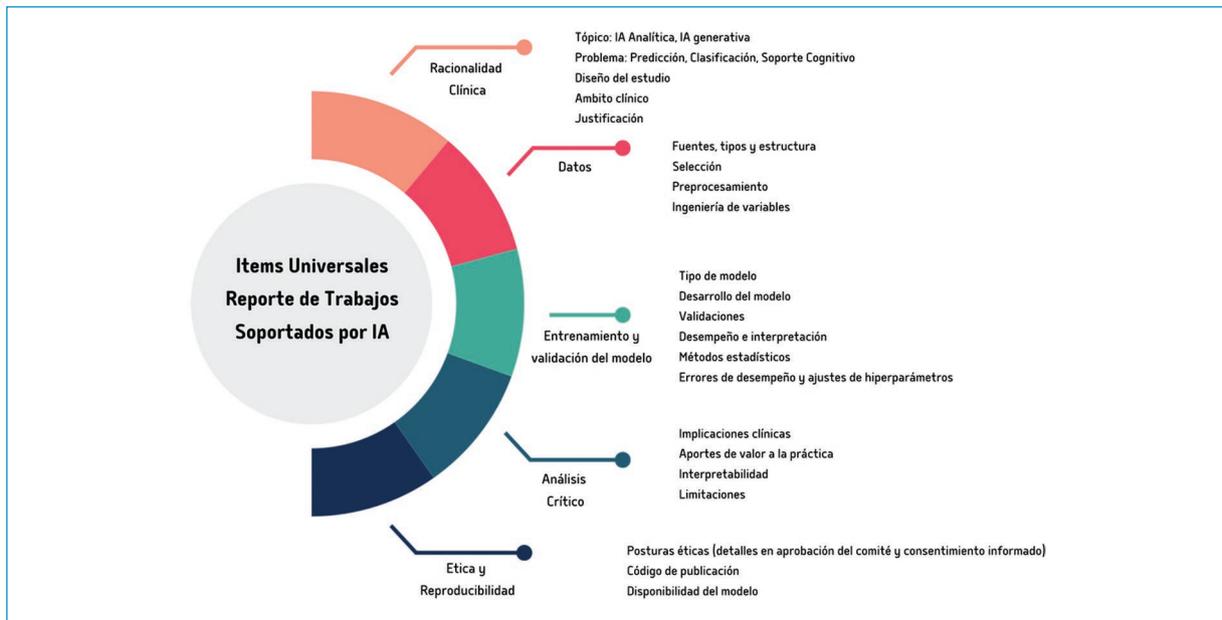


Figura 1. Ítems universales para el reporte de investigación basada en IA. Adaptada de Kolbinger et al.⁶

siempre su uso. En dado caso que se trate de una investigación que utilizó IA para su desarrollo, tanto en el proceso de escritura como de selección de pacientes, o en el almacenamiento de datos, siempre debe describirse detalladamente en qué parte del trabajo se utilizó la IA, el nombre de la herramienta utilizada, la fecha y la confirmación legal de que el autor asume la responsabilidad del producto final^{6,7}.

Finalmente, varias guías de escritura y reporte de evidencia científica han trabajado mediante consensos de expertos en ajustar cierta normatividad para la escritura de estos trabajos y optimizar la calidad de la investigación científica basada en IA. Dependiendo del tipo de estudio y de la fase clínica de investigación se pueden encontrar varias de ellas, como lo son PRISMA-I, para ensayos clínicos aleatorizados; PRIME, para fase preclínica en imágenes cardiovasculares; CONSORT-AI, para ensayos clínicos aleatorizados; DECIDE-AI, para sistemas tempranos de IA en soporte a la decisión clínica; R-AI-DIOLOGY, para estudios preclínicos de neuroradiología; CLEAR para estudios de medicina nuclear o radiología⁶.

Con el fin de aportar un breve resumen en este editorial sobre las guías mencionadas anteriormente, Kolbinger et al., en una revisión sistemática de la literatura sobre guías de reporte de evidencia científica basada en IA, resumieron los diferentes ítems universales para el reporte de los productos de investigación basados en modelos de IA, los cuales consideraron como universales al incluirse en más del 50% de las guías⁶ (Fig. 1).

A manera de conclusión, consideramos que la gran revolución de investigación científica basada en IA, también propone el reto de una revolución en los comités de ética y en los comités editoriales, dado que la mayoría carecen de la infraestructura y del conocimiento para abordar esta nueva fuente de investigación. Esta revolución propone un cambio en las matrices de calificación de estos tipos de estudios, en la rigurosidad de la escritura, en la revisión de transparencia y calidad de los algoritmos, en el detalle legal de la responsabilidad de los resultados y en la verificación de riesgos y beneficios que deben contemplarse antes de publicar o de implementar una nueva tecnología basada en IA en la población colombiana.

Bibliografía

1. Bouhouita-Guermech S, Gogognon P, Bélisle-Pipon JC. Specific challenges posed by artificial intelligence in research ethics. *Front Artif Intell [Internet]*. 2023 [cited 1 Sep 2024];6. <https://www.frontiersin.org/journals/artificialintelligence/articles/10.3389/frai.2023.1149082/full>.
2. Beam A, Kohane I. Big data and machine learning in health care. *JAMA*. 2018;319(13):1317-8.
3. Gundersen OE, Kjensmo S. State of the art: reproducibility in artificial intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence [Internet]*. 2018 [cited 1 Sep 2024];32(1). <https://ojs.aaai.org/index.php/AAAI/article/view/11503>.
4. Lipton ZC. The mythos of model interpretability. *Commun ACM*. 2018;61(10):36-43.
5. Blumenthal D, Patel B. The regulation of clinical artificial intelligence. *NEJM AI*. 2024;1(8):Alpc2400545.
6. Kolbinger F, Veldhuizen G, Zhu J, Truhn D, Kather JN. Reporting guidelines in medical artificial intelligence: a systematic review and meta-analysis. *Commun Med*. 2024;4:71.
7. JAMA cardiology. Instructions for authors. <https://jamanetwork.com/journals/jamacardiology/pages/instructions-for-authors>.