

La inteligencia artificial y la ética.

La inteligencia artificial y la ética.

Artículo de investigación

Ethics and artificial intelligence

Celia Rivilla-Piñango

Estudiante de Grado en Ingeniería de la Salud por la Universidad de Málaga y la Universidad de Sevilla. celia.rivilla@gmail.com

María-Jesús Tamayo-López

Cirujana y profesora en la Universidad de Sevilla. Universidad de Sevilla – Hospital Universitario Virgen del Rocío de Sevilla. mtamayolopez03@gmail.com

Resumen

La IA puede tener en el ámbito clínico muchas y muy variadas aplicaciones de gran utilidad, sin embargo, su utilización plantea varios desafíos éticos que surgen de la combinación de nuevas tecnologías, nuevas partes interesadas, una gran cantidad de datos y nuevas técnicas computacionales y analíticas. Frente a ello descubrimos una falta de controles reglamentarios y estándares comunes para guiar su creciente aplicación en el ámbito de la salud.

Entre las inquietudes planteadas se incluyen abordar los riesgos para la privacidad y confidencialidad de los pacientes, los límites entre el papel del médico y el de la máquina en la atención al paciente y toma de decisiones, la certeza de que las personas tendrán la información necesaria sobre dichas herramientas para asegurar una decisión informada y que los riesgos de daños frente a posibles beneficios sean siempre evaluados. También las amenazas en cuanto a las preferencias del paciente, la interpretabilidad e imparcialidad de algunos algoritmos y la accesibilidad.

La inteligencia artificial y la ética.

Palabras clave: inteligencia artificial; bioética; aprendizaje automático; toma de decisiones; big data

Abstract

AI can have many and very diverse applications than can be highly useful in the clinical setting. However, its use raises several ethical challenges that emerge from the combination of new technologies, new stakeholders, a large amount of data and new computational and analytical techniques. Furthermore, we lack regulatory controls and common standards to guide its growing application in the field of health.

Concerns raised include addressing risks to privacy and confidentiality of patients, the boundaries between the role of the physician and that of the machine in patient care and decision-making, the certainty that people will have the necessary information on these tools to ensure an informed decision and that the risks of harm versus possible benefits are always evaluated. Also threats regarding patient preferences, interpretability and fairness of some algorithms and accessibility.

Key words: artificial intelligence; bioethics; machine learning; decision-making; big data

Introducción

El concepto de inteligencia artificial (IA) se refiere a la disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico, y el crecimiento de su aplicación en salud y medicina es una realidad, especialmente en la última década.

La IA puede tener en el ámbito clínico muchas y muy variadas aplicaciones de gran utilidad, entre ellas algunas como facilitar y mejorar la capacidad de diagnóstico, predecir reingresos hospitalarios, ayudar en el descubrimiento de nuevos fármacos o procesar y obtener información de historiales médicos electrónicos. Sin embargo, su utilización también plantea varios desafíos y preocupaciones éticas.

Las cuestiones éticas relacionadas con la IA empezaron a investigarse hace más de 20 años,

La inteligencia artificial y la ética.

cuando se conceptualizó la tecnología de los historiales médicos electrónicos (EHR), y a medida que nuevas formas de tecnologías generalizadas producen datos, están surgiendo inquietudes sobre el respeto de los principios éticos biomédicos en la aplicación de estas tecnologías. El respeto a la autonomía, la beneficencia, la no maleficencia y la justicia, siguen siendo relevantes y deben priorizarse para garantizar que las personas estén protegidas ante posibles daños.

Los desafíos éticos surgen de la combinación de nuevas tecnologías que evolucionan rápidamente, nuevas partes interesadas (por ejemplo, grandes empresas tecnológicas), una gran cantidad de datos y nuevas técnicas computacionales y analíticas; y frente a ellos descubrimos una falta de controles reglamentarios o estándares comunes para guiar esta convergencia en el ámbito de la salud.

Muchas de las inquietudes planteadas en este tema incluyen abordar el riesgo adicional para la privacidad y confidencialidad del paciente, el análisis de los límites entre el papel del médico y el de la máquina en la atención al paciente, la certeza de que las personas tendrán la información que necesitan respecto a dichas herramientas para tomar una decisión informada y que los riesgos de daños frente a posibles beneficios se evaluarán siempre.

Algunas iniciativas que están buscando estas implicaciones legales, sociales y éticas (ELSI por sus siglas en inglés) en la investigación digital son, por ejemplo: la iniciativa ReCODE Health de la Universidad de California en San Diego, el programa de PERVADE en la Universidad de Maryland, y en el proyecto mHealthELSI de Sage Bionetworks y la Universidad de Louisville, todas ellas con el objetivo de estudiar posibles políticas y formas de gobernanza de un espacio en gran medida no regulado como es la aplicación de IA en nuevos ámbitos como la medicina.

Objetivos

El objetivo de este trabajo es estudiar las posibles implicaciones éticas de la incorporación de herramientas basadas en inteligencia artificial en medicina y en el cuidado de la salud, y valorar la importancia de la bioética a la hora de abordar los nuevos desafíos éticos que plantea este campo para quienes toman decisiones sobre la selección, prueba, implementación y evaluación de estas tecnologías para su uso en la atención médica.

Se pretende identificar las cuestiones éticas más relevantes que surgen según va cobrando

La inteligencia artificial y la ética.

importancia este ámbito mediante la realización de una revisión de los principales estudios publicados sobre el tema.

Nos planteamos las siguientes preguntas:

¿La IA está pensada para sustituir a los profesionales en el diagnóstico de los pacientes?

¿La IA se propone como ayuda al diagnóstico de los pacientes pero la decisión final es del profesional?

Material y método

Material:

La revisión de la bibliografía incluye las publicaciones de los últimos 5 años, en español e inglés, en las siguientes bases de datos: Google académico, Pubmed, Medline y la plataforma Web of Science.

Método:

La búsqueda bibliográfica inicial se realiza mediante el título o tema usando las siguientes palabras clave y la combinación de ellas: ética, problemas éticos, inteligencia artificial, *machine learning*, *deep learning*, *natural language process*, medicina, biomedicina y salud.

La estrategia de búsqueda electrónica puede resumirse en la siguiente secuencia: (*ethic** OR *ethical issues* OR *ethical challenges*) AND (*artificial intelligence* OR *machine learning* OR *deep learning* OR *synthetic intelligence* OR *natural language process*) AND (*medicine* OR *biomedicine* OR *healthcare*).

Criterios de inclusión:

Artículos publicados en español, inglés

Desde 2015-2020

Documentos de acceso abierto

Criterios de exclusión:

Otros idiomas que no sean inglés y español

Artículos anteriores 2015

Si no aportan criterios claros en cuanto a problemas éticos

La inteligencia artificial y la ética.

Se hace una revisión manual de los artículos para descartar artículos duplicados y verificar los criterios especificados a través de su resumen y lectura integral en caso de artículos cuya lectura del resumen no es suficientemente clara en cuanto al cumplimiento de los criterios definidos. Posteriormente se clasificaron según la temática con el fin de identificar aquellos artículos que pudiesen resultar más útiles para nuestro análisis y para obtener una idea general de los posibles problemas éticos en discusión. Esta clasificación se muestra reflejada en la Tabla 1.

Resultados

Tras la búsqueda se identifican 26 artículos de los cuales se eligen 9 que incluyen las palabras claves y aportan criterios y consenso para exponer el tema elegido.

Temática	Número de artículos
Implicaciones éticas de la creciente aplicación de IA en medicina	8
Ética de IA en diagnóstico por imagen, radiología y medicina nuclear	7
Ética sobre la IA integrada en psiquiatría y salud mental	3
Confianza en los DSS (Decision Support System) basados en IA	1
Implicaciones éticas de la aplicación de AI en medicina de laboratorio	1
Consideraciones éticas sobre IA en pronósticos en cuidados intensivos	1
Ética en la utilización de sistemas de IA en la atención sanitaria al cáncer de mama	1
Consideraciones éticas de la IA en otorrinolaringología	1
Ética de la inteligencia artificial en los cuidados perioperatorios	1
Retos éticos en el uso de IA en tecnología asistencial	1
Problemas éticos sobre IA y aplicaciones móviles	1

Tabla 1. Artículos clasificados por temática tras la búsqueda inicial

Comprobamos que efectivamente es un tema que está cobrando mucha importancia analizando el número de artículos por año de publicación: entre 2 y 3 artículos en 2015 y

La inteligencia artificial y la ética.
2016, aumentando notablemente en los siguientes años y concentrándose la gran mayoría de los artículos publicados en los años 2019 y 2020.

Discusión

La selección e implementación de tecnologías basadas en IA en el ámbito de la salud requiere la consideración de principios éticos, riesgos y beneficios, cuestiones de respeto de la privacidad, confidencialidad, acceso y usabilidad, así como implicaciones sobre la gestión de datos, el consentimiento informado y la autonomía del paciente.

Debe asegurarse un uso responsable de la tecnología de IA y los daños potenciales relacionados con su uso deben ser transparentes para todos los involucrados.

En los últimos años el campo de la IA está experimentando un intenso proceso de industrialización y mercantilización fuera del entorno académico en el que se ha desarrollado tradicionalmente, y sucede entonces que los límites médicos e industriales de la investigación son cada vez más confusos y surgen ciertas preocupaciones éticas.

Esta industrialización también significa que los productos de IA llegan con más facilidad a la sociedad en general y, por lo tanto, que los problemas éticos relacionados con el uso de la IA especialmente en la medicina y dominios sanitarios deben ser abordados.

También es fundamental que los requisitos mínimos utilizados para hacer que una tecnología de salud digital esté disponible para el público general no se confundan con un producto que ha pasado pruebas rigurosas o ha demostrado un valor terapéutico real. Por ejemplo, una aplicación que no pretende proporcionar un diagnóstico o tratamiento clínico, no requeriría de aprobación y podría ser fácilmente accesible desde dispositivos como teléfonos móviles, por ello es importante que los pacientes comprendan hasta qué punto la tecnología respaldará sus necesidades de atención médica ya que podrían confiar falsamente en la tecnología para proporcionar la atención que creen que necesitan y resultar en un perjuicio.

Además de las amenazas a la privacidad, debemos cuestionar el proceso de consentimiento informado, si existe un cálculo objetivo del riesgo de daños contra los potenciales beneficios y si las personas incluidas en la fase de prueba del producto son aquellas que tienen más probabilidades de beneficiarse.

Cuando la tecnología tiene implicaciones para la salud, los estándares adquieren una

La inteligencia artificial y la ética.

importancia crítica, por ello se hace necesario desarrollar una infraestructura para examinar las herramientas de IA antes de implementarlas en un sistema de salud.

En el uso de IA como una herramienta de apoyo a la decisión clínica, se hace necesario distinguir entre el asesoramiento basado en evidencia y la toma de decisiones en la atención médica. Determinar cuánto y qué tareas deberíamos delegar al aprendizaje automático, requiere analizar y distinguir específicamente los diferentes sistemas y diferentes tareas delegadas.

Surgen así conflictos de políticas legales como negligencia médica y responsabilidad que aparecen de forma más compleja asociados a algunos algoritmos que a menudo se caracterizan como “cajas negras”, los cuales no pueden proporcionar a sus usuarios una interpretación de cómo el algoritmo ha llegado a un determinado resultado ya que la forma funcional que relaciona los datos disponibles (entrada) con un resultado dado (salida) resulta demasiado compleja.

Estas cuestiones de explicabilidad e interpretabilidad son primordiales cuando la toma de decisiones en medicina (diagnóstico, pronóstico, etc.) debe transmitirse a los humanos.

El deseo de crear y comercializar nuevas aplicaciones de IA en medicina ha creado una demanda y un mercado para datos derivados de pacientes cuya propiedad y derechos de uso son muy complejos y varían según la jurisdicción, y dependen de su grado de anonimización.

El uso de grandes bases de datos, un elemento básico de la IA ilustra cómo entre la IA y la protección de la privacidad pueden aparecer motivos enfrentados. Se han desarrollado métodos para asegurar la anonimización pero incluso algunos supuestamente muy precisos parecen incapaces de eliminar todos los tipos de información de salud protegida ya que puede requerir eliminar o sobrescribir metadatos que se encuentran ocultos o en etiquetas privadas. Así mismo también se deben evitar imágenes que podrían reconstruirse para crear una visualización facial o ser aplicadas a tecnologías de reconocimiento facial. En caso de no ser posible lograr una anonimización real debe exigirse que los pacientes sean informados al respecto. Asimismo grandes cantidades de datos de capacitación para las aplicaciones de IA se recopilan de múltiples fuentes como registros médicos y de seguros, datos farmacéuticos, datos genéticos y redes sociales lo cual hace más fácil rastrear esos datos hasta los pacientes.

En cuanto a casos raros y generalización deben considerarse mecanismos para evitar daños

La inteligencia artificial y la ética.

al paciente que pueden surgir de una excesiva dependencia a herramientas que podrían ser precisas en la mayoría de los casos, pero podrían ser susceptibles de fallar por completo en otras situaciones, ya que la aplicación de IA puede acercarse o incluso superar el rendimiento humano en el diagnóstico de las enfermedades más comunes o para cuya detección está diseñada, sin embargo pueden existir condiciones raras presentes en una imagen que un médico especialista podría diagnosticar fácilmente pero que puede no haber estado presente o correctamente etiquetado en la capacitación del modelo datos.

Asimismo pueden existir sesgos relacionados con algoritmos de IA que podrían no proporcionar predicciones igualmente precisas de resultados por raza, género o estatus socioeconómico. Esto es debido a que las bases de datos de investigación de ensayos clínicos de las cuales se obtienen los conjuntos de datos de aprendizaje automático (ML) se derivan en gran medida de poblaciones mayoritarias. Por lo tanto, los algoritmos resultantes pueden ser más propensos a fallar cuando se aplica a grupos de pacientes subrepresentados, por ejemplo podrían generalizarse mal en otros entornos como otros lugares con nuevos pacientes.

Así han de usarse modelos basados en datos de entrenamiento que sean representativos de la población, combinación de casos, modalidades y adquisición, y deben llevarse a cabo reevaluaciones del rendimiento del algoritmo en nuevas configuraciones, lo cual puede requerir la adición de fuentes locales datos de entrenamiento.

La aplicación de algoritmos sesgados y opacos podría aumentar las disparidades en materia de salud y poner en peligro la confianza pública, por ello para su aplicación se requiere una comprensión práctica de los posibles peligros y límites, y el desarrollo de un diseño bien pensado, de alta calidad y validado clínicamente para garantizar una implementación responsable.

Para garantizar aplicaciones de IA confiables, sería conveniente la creación de un marco y un sistema ético global, así como pautas especiales para aplicaciones de IA en medicina.

Conclusión

Se reconoce el potencial de la IA para mejorar enormemente la atención médica, sin embargo esta tecnología conlleva un nuevo conjunto de desafíos éticos que debe ser identificado y mitigado ya que la tecnología de IA tiene también capacidad para amenazar la

La inteligencia artificial y la ética.

preferencia del paciente, la interpretabilidad y la explicabilidad, la imparcialidad, la seguridad, la privacidad y el anonimato, la equidad y la accesibilidad y el beneficio para todos; y la política actual y las pautas para la aplicación de IA están rezagadas con respecto al progreso que esta tecnología está haciendo en el campo de la medicina.

Abordar estos problemas requiere la colaboración entre las comunidades de AI-ML y medicina- salud y autoridades reguladoras que contribuyan al desarrollo de métodos, protocolos, pautas y líneas de análisis de datos, un código de ética uniforme al que adherirse que tenga en cuenta explícitamente todas estas cuestiones éticas y proporcione un marco ético confiable a medida que la tecnología avanza rápidamente.

Bibliografía

1. Guan J. Artificial Intelligence in Healthcare and Medicine: Promises, Ethical Challenges and Governance. *Chin Med Sci J*. 2019;34(2):76-83.
2. Car J, Sheikh A, Wicks P, Williams MS. Beyond the hype of big data and artificial intelligence: building foundations for knowledge and wisdom. *BMC Med*. 2019;17(1):143.
3. Nebeker C, Torous J, Bartlett Ellis RJ. Building the case for actionable ethics in digital health research supported by artificial intelligence. *BMC Med*. 2019;17(1):137.
4. Vellido A. Societal Issues Concerning the Application of Artificial Intelligence in Medicine. *Kidney Dis (Basel)*. 2019;5(1):11-17.
5. Buruk B, Ekmekci PE, Arda B. A critical perspective on guidelines for responsible and trustworthy artificial intelligence. *Med Health Care Philos*. 2020;10.1007/s11019-020-09948-1.
6. Safdar NM, Banja JD, Meltzer CC. Ethical considerations in artificial intelligence. *Eur J Radiol*. 2020;122:108768.
7. Braun M, Hummel P, Beck S, Dabrock P. Primer on an ethics of AI-based decision support systems in the clinic. *J Med Ethics*. 2020;medethics-2019-105860.
8. Luxton DD. Recommendations for the ethical use and design of artificial intelligent care

La inteligencia artificial y la ética.

providers. Artif Intell Med. 2014;62(1):1-10.

9. Ho CWL, Soon D, Caals K, Kapur J. Governance of automated image analysis and artificial intelligence analytics in healthcare. Clin Radiol. 2019;74(5):329-337.

Cómo citar: Rivilla-Piñango, C. Tamayo-López, MJ. La inteligencia artificial y la ética. Bioética y Ciencias de la salud [internet] Enero-Junio Vol 9 (1) 2021