



AULA DE EDUCACIÓN MÉDICA

Aspectos éticos para la educación médica en inteligencia artificial

Pilar Leon-Sanz

Facultad de Medicina, Universidad de Navarra, Pamplona, España

Recibido el 4 de agosto de 2025; aceptado el 26 de agosto de 2025

Disponible en Internet el xxxx

PALABRAS CLAVE

Inteligencia artificial en medicina;
 Datos de salud;
 Equidad algorítmica;
 Privacidad y confidencialidad;
 Consentimiento informado;
 Educación médica

Resumen

Introducción: se revisan las cuestiones ético-médicas específicas que se han contemplar en la formación de las competencias relacionadas con la realización o aplicación de proyectos de inteligencia artificial (IA) en medicina.

Método: se analizan los principales documentos de consenso y otros estudios.

Resultados: se abordan las características éticas generales de la IA en medicina y a raíz de ellas, se consideran 3 aspectos relevantes en el debate ético-médico sobre la IA en medicina: el uso y tipos de datos de salud, teniendo en cuenta los usos secundarios y la reutilización de los datos; y se analiza cómo se ha previsto el acceso a los datos de salud en el Espacio Europeo de Datos sanitarios (tanto para un uso primario como secundario). En segundo lugar, se aborda la equidad algorítmica y, en tercer lugar, la responsabilidad de los profesionales de la salud en el diseño y aplicación de la IA en general y también de la IA generativa.

Conclusiones: la IA no es éticamente neutra, las competencias formativas han de contemplar las buenas prácticas en los proyectos de IA para que sean fiables también desde el punto de vista ético, de modo que las mejoras que se introduzcan garanticen el respeto a los valores y derechos de las personas y de la sociedad.

© 2025 El Autor/Los Autores. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

KEYWORDS

Artificial intelligence in medicine;
 Health data;
 Algorithmic equity;
 Privacy and confidentiality;
 Informed consent;
 Medical education

Ethical aspects for medical education in artificial intelligence

Abstract

Introduction: The article reviews the specific ethical issues that should be considered in developing competencies related to the implementation or application of Artificial Intelligence projects in medicine.

Method: The article analyzes the main consensus documents and other studies.

Results: The general ethical characteristics of AI in medicine is described. And also, three relevant aspects of the ethical and medical debate on AI in medicine are considered: the use and

Correo electrónico: mpleon@unav.es.

<https://doi.org/10.1016/j.edumed.2025.101100>

1575-1813/© 2025 El Autor/Los Autores. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

types of health data, taking into account secondary uses and data reuse, and how access to health data has been provided in the European Health Data Space (both for primary and secondary use). Second, the article addresses equity and algorithmic fairness, and third, the responsibility of health professionals in the design and application of AI in general, as well as generative AI.

Conclusions: AI is not ethically neutral. Training competencies must consider best practices in AI projects so that they are also ethically sound, ensuring that the improvements introduced guarantee respect for the values and rights of individuals and society.

© 2025 The Author(s). Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción: la reflexión ética sobre la inteligencia artificial

En general, los proyectos de inteligencia artificial (IA) se refieren al desarrollo de aspectos como el razonamiento, el aprendizaje, la percepción, la predicción, la planificación o control, en entornos reales y virtuales. Se trata de programas informáticos capaces de ejecutar, de forma autónoma, operaciones como el aprendizaje o el razonamiento lógico.

La aplicación de la IA a la medicina ha adquirido un enorme interés debido al potencial que tiene para avanzar en la investigación y en las prácticas biomédicas. Hay un uso creciente de estos sistemas en la detección de enfermedades (imágenes, anatomía patológica y biomarcadores), en la predicción de resultados clínicos, en la personalización de los tratamientos o en los proyectos para descubrir nuevos fármacos¹⁻³, etc. Al mismo tiempo, existe una gran sensibilidad respecto a la inseguridad que su aplicación puede suponer para los pacientes y los profesionales⁴.

En estas aulas médicas vamos a referirnos a las características éticas de la IA en medicina incluidas en la mayoría de los documentos de consenso; y a 3 cuestiones que han de estar presentes en la formación continuada sobre la ética de la IA en medicina: en primer lugar, la utilización de los datos de salud para el diseño y aprendizaje de los proyectos, teniendo en cuenta los usos secundarios y la reutilización de estos datos. Abordaremos, en segundo lugar, la equidad algorítmica y, en tercer lugar, la responsabilidad de los profesionales de la salud en el diseño y aplicación de la IA.

Dada la próxima entrada en vigor (2027), revisaremos cómo está previsto el acceso a los datos de salud en el Espacio Europeo de Datos Sanitarios, tanto para un uso primario o asistencial, como secundario (investigación, usos para políticas sanitarias, etc.).

Nos referiremos también a la IA generativa, puesto que desde la aparición de los *Large Language Models* (LLM) o modelos de lenguaje de gran tamaño (ChatGPT, OpenAI 2022; Gemini [Google 2024]; Copilot [Microsoft 2025]; DeepSeek AI 2025, etc.), estos sistemas lideran la innovación en el acceso y procesamiento de información. Se trata de sistemas de predicción de texto avanzados que resultan de utilidad creciente para el trabajo profesional (redacción de comunicaciones, resúmenes e informes, tareas administrativas,

interpretación y traducción) y para crear plataformas con aplicación médica⁵.

El primer paso para desarrollar una buena práctica médica es la reflexión personal sobre lo que supone la IA, lo cual favorecerá el pensamiento crítico, imprescindible para usar bien la IA, para considerar cuándo y de qué forma conviene emplear la IA o sus resultados en la actividad clínica o investigadora.

El buen uso de la IA se ha de valorar en cada fase del ciclo de vida de la IA: desde la concepción, desarrollo y validación inicial, a la aplicación y utilización, al mantenimiento del proyecto, a su seguimiento, a su financiación y comercialización y, finalmente, al término de la actividad. Como apunta el Comité de Bioética de la UNESCO, conviene advertir que la ética de la IA tiene un componente dinámico por la evolución continua de estas tecnologías⁶.

Características éticas de la inteligencia artificial

A diferencia de otras metodologías estadísticas que operan con parámetros controlados, los sistemas basados en la IA no son herramientas neutras desde un punto de vista ético, puesto que, desde el inicio, requieren decisiones humanas. Es el caso de la selección de los datos, del diseño de los objetivos, etc. Las aplicaciones basadas en la IA no se reducen a códigos abstractos, sino que sus principios operativos tienen implicaciones éticas que han de guiar su impacto real. Es el caso de la justicia algorítmica, que conlleva que los datos de entrenamiento no tengan sesgos históricos o estructurales, datos desiguales o que puedan introducir o perpetuar algún tipo de discriminación. Como veremos a lo largo de las páginas siguientes, se ha de facilitar la explicabilidad, es decir, que los algoritmos sean comprensibles a los usuarios en los procesos de toma de decisiones, se ha de procurar que nuestras decisiones no perjudiquen a nadie y menos aún a personas o grupos vulnerables, y también se ha de garantizar la privacidad y seguridad de los sistemas.

Con independencia de las regulaciones, como el Reglamento de la UE sobre la IA⁷ o el Espacio Europeo de Datos Sanitarios⁸, en el ámbito de la ética biomédica, existen numerosos e importantes documentos de consenso con recomendaciones para la evaluación ética de las tecnologías de la IA^{6,9-12}. Una revisión sistemática de más de 50 documentos publicados entre 2018 y 2023, realizada por las

Academias Nacionales de Medicina, Ciencias e Ingeniería de EE. UU.¹³, identificó 5 características éticas comunes. Los sistemas de IA han de tener en cuenta: a) la centralidad de la persona, que requiere preservar su dignidad, la autonomía (consentimiento informado), así como la privacidad y la confidencialidad; b) la transparencia, claridad e inteligibilidad o explicabilidad; c) la inclusividad y la equidad, para evitar sesgos algorítmicos; d) la seguridad de las personas, de los grupos de personas y de la sociedad; e) la promoción de una IA con capacidad de respuesta y sostenible, ya que los recursos que requiere la tecnología son cada vez más relevantes.

Respeto a los derechos y valores de las personas

La ética de la IA es común a la ética de la ciencia, de la medicina y de la ética de la tecnología, que mantienen entre sus principios generales la primacía de los derechos de las personas y de los pacientes por encima de la generación de nuevos conocimientos, de la mejora de las intervenciones preventivas, diagnósticas y terapéuticas, o del avance de la sociedad¹⁴.

En efecto, la ciencia no es simple experimentación y recolección de datos, sino que, sobre todo, es análisis e interpretación de datos y experiencias. Colaborar en la obtención de un algoritmo que mejore un diagnóstico o hacer un estudio epidemiológico, hallar nuevas moléculas para un tratamiento, etc., es algo muy importante, pero no justifica que se puedan conculcar derechos básicos de los pacientes. «Todos [los que participen] tenemos la responsabilidad de garantizar que nuestra investigación se utilice de forma ética» (Arts. 70, 85, 86)¹⁵.

La ética de la IA depende de las mejoras que introduzca en la atención al paciente individual y en la salud de la sociedad en general (Art. 82.2)¹⁵, pero también, del modo en que esos adelantos se han conseguido.

La inteligencia artificial y los datos de salud

La IA necesita datos para su aprendizaje y aplicación, y en el ámbito sanitario los datos son especiales porque afectan al núcleo de la intimidad de la persona¹⁶ y son obtenidos, generalmente, con una finalidad asistencial. Ese es su uso primario. Como se observa en la [tabla 1](#), los datos de salud tienen un origen múltiple: unos proceden de la asistencia médica, otros de la investigación, del área de la salud pública, del ámbito administrativo, o simplemente son incorporados como consecuencia del registro de actividades sociales. Es una información que se prolonga a lo largo del tiempo y puede hacer referencia a terceras personas, como los familiares. Frecuentemente, proceden de sujetos vulnerables, enfermos o menores de edad.

Tipos de datos: codificación, seudonimización y anonimización de datos

Tradicionalmente, se ha dicho que una información tiene carácter personal si está ligada a un nombre, es decir, si está identificada. Pero con finalidades secundarias o no asistenciales, como la investigación, la información puede

Tabla 1 Características del origen de los datos de salud

Diversidad de procedencia	Tipo de registros	Requisitos éticos
Fuentes privadas	Clínicos	Calidad de los datos
Fuentes públicas	Laboratorio	Confidencialidad y privacidad
	Investigación	Consentimiento informado (cesión de datos)
	Salud pública	Custodia de los datos
	Administrativos	Comunicación de resultados
	Actividades sociales/de consumo	Transparencia (origen/contextos/operaciones)
		Uso y finalidad

Elaboración propia.

estar codificada o seudonimizada, de manera que los datos no se puedan atribuir a una persona concreta sin recurrir a una información adicional o código. En el caso de los datos seudonimizados, la codificación es realizada por una persona ajena al proyecto, lo que proporciona una mayor garantía de confidencialidad. Tanto la información codificada como seudonimizada es personal, porque permite volver a identificar a las personas de quienes provienen esos datos, aunque facilita la explotación de los datos. Es un modo muy frecuente de utilizar los datos en proyectos de investigación clínica.

Los datos anónimos o anonimizados, en cambio, son aquellos que no se pueden relacionar con personas singulares porque, de forma irreversible, han perdido (o no han tenido) carácter personal. También se han desarrollado métodos de agregación de datos que permiten la alimentación de los algoritmos y facilitan la privacidad porque dificulta la identificación de los datos.

Como veremos a continuación, el uso de los datos de salud en proyectos de explotación de grandes bases de datos, o bien en proyectos de IA, plantea: la reutilización de los datos, el perfilado de las personas o grupos de personas, el modo de obtener el consentimiento para el uso de los datos de salud y el respeto a la confidencialidad de las personas¹⁷.

La reutilización de los datos

La reutilización de los datos de salud para fines secundarios como la planificación de servicios de salud, la investigación, proyectos de IA, etc., supone nuevas oportunidades, evita costes y los inconvenientes de recabar la misma información o semejante, para objetivos diversos ([tabla 2](#)), pero plantea también dificultades éticas por las consecuencias que implican el cambio del contexto para el que se han obtenido o la finalidad. De manera que los datos reutilizados pueden adquirir sentido y significado diferentes^{18,19}. Sería el caso, por ejemplo, de la reutilización de los datos proporcionados para revisar la cualificación de los cirujanos y las complicaciones quirúrgicas de sus intervenciones, para usos

Tabla 2 Uso secundario de los datos y reutilización de los datos

Utilidad	Consecuencias	Requisitos éticos
Nuevas oportunidades	Cambio de significado de los datos	Adecuación de los objetivos
Disminuye costes	Posible identificación de las fuentes	Precisión de los algoritmos
Evita dificultades obtención datos	Dificultad obtención del consentimiento	Justificación social y transparencia
Permite combinar bases de datos	Identificación de grupos	Confidencialidad y privacidad
Facilita acceso a datos de diversos ámbitos	Creación de perfiles	Consentimiento a la cesión de datos
	Hallazgos incidentales	Evitar discriminación y estigmatización
		Protección de poblaciones vulnerables
		Garantías en los usos comerciales

gerenciales no previstos inicialmente, o bien, si las fuerzas del orden o de seguridad tuvieran acceso a bases de datos clínicos, pues los indicadores de salud o enfermedad se podrían convertir en «indicadores de culpabilidad» (p. 18)¹⁰.

Del consentimiento a la cesión de datos

El consentimiento informado ha sido una de las principales garantías establecidas para proteger la autonomía y la responsabilidad de las personas en la asistencia médica y en la investigación²⁰. Así lo recuerda la Declaración de Helsinki: «Los médicos u otras personas cualificadas deben obtener el consentimiento libre e informado de los participantes en la investigación para la recopilación, el procesamiento, el almacenamiento y el uso secundario previsible de material biológico y datos identificables o reidentificables... un comité de ética de la investigación debe aprobar la creación y el seguimiento del uso continuo de esas bases de datos y biobancos» (p. 32)¹⁴.

En el ámbito específico de los grandes conjuntos de datos, y en proyectos de IA, el consentimiento informado no solo hace referencia a la autorización del uso de los datos, sino también al flujo de los datos. De ahí que se hayan desarrollado nuevos modos de prestar el consentimiento por parte de los pacientes¹⁷.

Se habla de «consentimiento dinámico» cuando se produce una actualización del consentimiento: supone volver a otorgar la conformidad si los datos se van a utilizar para otros fines o se van a incorporar nuevos datos propios. Este tipo de consentimiento se utiliza, por ejemplo, en cohortes de personas de estudios de salud pública que se prolongan en el tiempo.

El llamado «consentimiento amplio» es la autorización de las personas para el uso de la información personal para más de un proyecto, en una misma línea de investigación, sin tener que volver a contactar de nuevo con el paciente o el sujeto.

Otra manera de obtener datos de salud es la «cesión automatizada» de datos que se produce, si no hay una oposición, cuando se vuelcan los datos registrados en dispositivos electrónicos (respiradores, Smartwatch). Aunque la cesión automatizada de datos se ha generalizado, no siempre se conocen las implicaciones que puede tener y los nuevos fines comerciales, estudios de consumo, etc., para los que se pueden utilizar.

La «comunicación voluntaria» de datos implica que las personas ponen sus datos personales a disposición de

terceros. Se produce normalmente tras un llamamiento público de solicitud de datos para un proyecto determinado.

La regulación española de Protección de Datos (disposición adicional decimoséptima)²¹ admite que, con el informe previo favorable del comité de ética de la investigación, se puedan reutilizar los datos personales con fines de investigación en materia de salud y biomédica sin que se solicite un consentimiento informado, siempre que las personas no hayan expresado su negativa a su uso.

El Espacio Europeo de Datos Sanitarios

En marzo de 2025 fue publicado el Reglamento de la Unión Europea relativo al Espacio Europeo de Datos de Salud (EEDS), que regula la libre circulación de los datos de salud en los países de la Unión Europea (Artículo 1). El Reglamento prevé una puesta en marcha en marzo de 2027, tras un período transitorio, y la circulación de los datos de las historias clínicas electrónicas, en marzo de 2029⁸.

El EEDS pretende mejorar el acceso a los datos de salud electrónicos personales, con fines primarios (asistenciales) y secundarios, es decir, para investigación científica, para establecer políticas de salud (34)⁸, entre otros. Además, el Reglamento del EEDS incluye la comercialización de datos de salud, lo cual introduce nuevos objetivos, como la monetización de los datos. La Organización Médica Colegial ha analizado los retos y desafíos del EEDS en la Declaración sobre el Reglamento del Espacio Europeo de Datos Sanitarios, aprobada el 31 de mayo de 2025.

El Reglamento establece un consentimiento presunto para incorporar los datos de salud al EEDS, por lo que solo se excluirán los datos de quienes hayan manifestado expresamente su oposición (18)⁸. Con fines secundarios, los investigadores, las empresas o las instituciones públicas accederán, en general, a datos anonimizados y solo podrán acceder a datos seudonimizados si son necesarios para su finalidad. El desplazamiento de los datos tiene el riesgo de llevar a una pérdida de control. Un ejemplo es la compra de DeepMind, una compañía inglesa de investigación y desarrollo de IA, por una empresa norteamericana (matriz de Google), que implicó el acceso a los datos que se habían obtenido en un contexto diferente.

Se ha hecho hincapié en la formación y responsabilidad de los profesionales, puesto que son estos quienes incorporan los datos a las historias clínicas. La tecnología de procesamiento de lenguaje natural ha conseguido el acceso a los datos conservados en las historias clínicas, incluidos los

Tabla 3 Supervisión de los sistemas de inteligencia artificial

	Tipo de registros
Registros de trazabilidad	Decisiones técnicas y humanas en cada etapa del ciclo de vida del sistema basado en IA Auditorías Rendición de cuentas
Comités de supervisión legal y ética	Especialmente en aplicaciones críticas (salud, justicia)
Marcos legales adaptativos	Definición de responsabilidades y riesgo de la tecnología

desestructurados, a través del modelo estandarizado de datos comunes (OMOP, *Observational Medical Outcomes Partnership*) que facilita la recogida, estructuración, conservación, consulta y comunicación de los datos de las historias clínicas.

También se ha reivindicado una gobernanza ética con participación de autoridades, profesionales, comités de ética y representantes sociales, para garantizar la transparencia, la supervisión y la eficacia de los sistemas. En las [tablas 3 y 4](#) se muestran de forma esquematizada los modos de supervisión de los sistemas de IA y los objetivos éticos de la regulación de la IA.

Privacidad y confidencialidad: control de accesos y usos de los datos

El derecho a la confidencialidad implica que la información personal proporcionada por una persona no sea divulgada posteriormente sin su autorización, excepto en los casos establecidos ética y legalmente. La privacidad, el derecho a la confidencialidad y el modo de conservar la información han sido cuestiones importantes en la implantación de la tecnología de la información en medicina. Especial atención se ha prestado a la protección de los datos genéticos y a los datos psiquiátricos, por el riesgo de una posible estigmatización o discriminación¹⁷.

Tabla 4 Objetivos de la regulación

Objetivos éticos de la regulación de la inteligencia artificial

Garantizar la transferencia
Abordar el flujo de los datos (local, nacional e internacional)
Evitar desigualdades
Proteger poblaciones vulnerables
Promover el buen uso de los resultados
Compatibilizar intereses individuales vs. colectivos y transnacionales
Adecuar la distribución de los resultados
Establecer políticas de salud
Promover criterios éticos de gobernanza
Evitar regulaciones diversas sobre los mismos datos
Garantizar buenos usos comerciales

En el desarrollo de proyectos de IA es fundamental que el médico se comprometa a la protección de la confidencialidad y al control y propiedad de los datos del paciente (Art. 82.1 y 3)¹⁵.

Si bien es cierto que el derecho a la confidencialidad de los datos de salud no es absoluto, sin embargo, es la base para que se establezca la confianza profesional y social, por lo que se requiere una especial custodia y precaución, con el fin de evitar perjuicios con su uso, salvaguardar la buena fama, o bien respetar especialmente las decisiones de las personas.

Es más difícil preservar la privacidad conforme aumenta el número de datos almacenados y se diseñan algoritmos más sofisticados, aun cuando se trate de datos anonimizados, puesto que permiten correlacionar y extraer nuevos puntos de vista y nueva información^{20,22,23}. Un ejemplo de esto sucedió en 2012, con fines de investigación se combinaron una base de datos anonimizada proporcionada por la Group Insurance Commission de Boston (sin datos personales, ni información identificativa), con la base de datos de los votantes del Estado de dominio público (incluía nombre, código postal, dirección, sexo, fecha de nacimiento). Tras la combinación, fue posible identificar a políticos y ciudadanos concretos (p. 67)¹⁰. Por lo que, al establecer los criterios para salvaguardar la confidencialidad de las personas, hay que tener en cuenta que la combinación de bases de datos puede llevar a identificar a personas singulares o a grupos, según qué herramientas se utilicen y de qué otra información se disponga²⁴.

El derecho a no ser perfilado

La combinación de conjuntos de datos también puede dar lugar a enlaces más o menos temporales, o permanentes entre ellos ([tabla 2](#)) y crear así perfiles personales o de grupos. Por ejemplo, personas o grupos que tienen un riesgo más alto de desarrollar una enfermedad o la estimación de un menor efecto terapéutico de un tratamiento determinado. De ahí que el Consejo de Europa ha propuesto el reconocimiento del derecho de las personas a no ser perfiladas (p. 64)⁶. La denominada «identidad de grupo» ha de ser protegida en el ámbito del análisis de grandes bases de datos y de los proyectos de IA, con el fin de evitar la discriminación tanto de los individuos como de grupos de personas.

Respeto a la privacidad versus interés público

A través de encuestas y estudios de campo se percibe que la opinión de algunos sectores de la sociedad sobre la utilización de los datos personales es cambiante. Así, por ejemplo, los ciudadanos del Departamento de Western Australia apoyaron mayoritariamente la consulta que se formuló sobre la utilización, para fines de política sanitaria y otras investigaciones de la base de datos de morbilidad, de más de 3 decenios²⁴, y el informe del Nuffield Council describe ejemplos similares en Europa y en el Reino Unido (pp. 132–133)¹⁰. Lo que muestra que las personas tienen sensibilidades distintas o cambiantes respecto al uso de sus datos.

Equidad algorítmica

La equidad se define como: «la ausencia de prejuicio o favoritismo hacia un individuo o un grupo en función de sus características inherentes o adquiridas»²⁵. En el ámbito de la IA ha aparecido un nuevo campo de estudio denominado «equidad algorítmica», que pretende detectar sesgos y mitigarlos para garantizar decisiones justas y libres de discriminación²⁶. La equidad algorítmica es una parte importante de la corrección técnica y ética tanto en el diseño como en la aplicación de los proyectos de la IA. Los algoritmos de aprendizaje automático pueden amplificar sesgos históricos o estructurales presentes en los datos de entrenamiento.

Hay varias fuentes de sesgo algorítmico, entre las que podemos destacar el sesgo humano histórico, que aparece cuando se entrena un algoritmo con datos que ya estaban sesgados. Por ejemplo, se ha constatado que existen sesgos en el tratamiento del dolor según los datos provengan de una población masculina o femenina, lo cual podría perpetuar la discriminación.

Otra fuente de sesgo es la información incompleta o no representativa. Un ejemplo es el sistema de reconocimiento facial que utilizó principalmente fotos de hombres blancos para entrenar a la máquina, que llevó a resultados deficientes en el reconocimiento de mujeres y personas de otras razas²⁷.

Además, hay que evitar el sesgo del *experto* que lleva a que el diseño del algoritmo esté viciado por intentar ajustar el rendimiento del algoritmo al conjunto de datos o a los objetivos de la investigación, o simplemente porque se aplica la fórmula algorítmica que más se conoce, no tanto la más adecuada.

Además, los profesionales no pueden extrapolar los resultados más allá del alcance del estudio. Los sesgos introducidos adquieren más relevancia si se considera que, con frecuencia, los hallazgos son usados para nuevos análisis, lo cual puede perpetuarlos. De ahí que la transparencia se refiera tanto al origen de los datos, como a los contextos y a las operaciones realizadas, aun cuando, en ocasiones, la complejidad de los algoritmos no facilite la nitidez de la intervención realizada.

Sobre la responsabilidad profesional y el uso de la inteligencia artificial

La IA sirve de ayuda en la toma de decisiones clínicas y sanitarias, y cada vez lo hará de forma más eficaz y variada, pero ¿quién es responsable legal y éticamente de las decisiones tomadas por un sistema de IA: el desarrollador, el usuario, la empresa, todos ellos? La contestación es compleja y constituye un desafío jurídico, ético y social. Implica reconocer que la cadena de responsabilidad abarca a desarrolladores (por la integridad del diseño), a usuarios (por el contexto de implementación) y a empresas (por la gobernanza institucional). Se trata de una responsabilidad compartida que se articula a lo largo de toda la cadena de valor.

Es cierto que las personas no tenemos la capacidad de procesar de las máquinas, pero los sistemas de IA no entienden lo que procesan, por lo que la utilización de algoritmos para la toma de decisiones clínicas es una

estrategia importante, que no elimina la responsabilidad del médico²⁸.

Se insiste en que un profesional no puede actuar en contra de un protocolo contrastado y actualizado sin justificar esa actuación, pero los datos específicos de una persona o sus circunstancias pueden llevar a actuaciones diferenciadas, por lo que, de igual manera que un protocolo clínico no sustituye la responsabilidad profesional, tampoco lo hacen los datos o sugerencias proporcionados por los algoritmos. Además, se mantiene la incertidumbre respecto a la exactitud de los datos y respecto al poder estadístico de los análisis. Lo que sí proporciona una IA ética y sin sesgos, son evidencias mayores para la actuación profesional.

Por el contrario, sería mala praxis dejar las decisiones en manos de la IA, tanto en la asistencia, como en la investigación. Seguimos siendo responsables si hay un diagnóstico o un pronóstico erróneos, o un uso excesivo de recursos, violaciones de la confidencialidad, o mala gestión sanitaria, derivadas de la aplicación de sistemas de IA (86.1)¹⁵.

Reflexiones sobre inteligencia artificial generativa

Herramientas como ChatGPT, Gemini y Copilot han tenido un impacto inmediato porque son sistemas accesibles que pueden usarse de manera libre, aunque se mantenga la incertidumbre de no conocer con seguridad las implicaciones que su uso pueda tener. La IA generativa resulta útil para resumir el contenido de un artículo, identificar patrones en los datos y preparar presentaciones, pero ¿puede la tecnología desarrollar ideas verdaderamente innovadoras? ¿Qué autoría se le ha de reconocer?

En el debate bioético, un aspecto clave ha sido advertir la coherencia y la veracidad de las respuestas que producen estos sistemas. Cuando ChatGPT u otro sistema de IA generativa se confunde es porque no es capaz de conocer: el sistema procesa, relaciona datos reales o no, sesgados o no. Según la procedencia de los datos que utilice puede tener sesgos lingüísticos, de género, raciales, etc. Además, se insiste en que estos sistemas incurren en errores, por lo que los datos que aportan han de ser comprobados. Solo las personas somos capaces de conocer la verdad, por lo que se subraya la importancia del pensamiento crítico al interactuar con estas herramientas, ya que preocupa que se publiquen investigaciones fraudulentas, defectuosas o falsas.

De ahí que la declaración de los editores de revistas científicas sobre el uso responsable de tecnologías de IA generativa²⁹ establece que los LLM u otras herramientas de IA generativa no deben figurar como autores en los artículos. Solo los investigadores pueden asumir la responsabilidad de lo publicado. Todos, autores, revisores y editores, conservan la responsabilidad final sobre lo publicado, y cuando se aplica la IA persiste la responsabilidad ética y jurídica de cada profesional. Además, las principales revistas científicas requieren que los autores firmen un formulario declarando las contribuciones al trabajo y sean transparentes respecto al uso de la IA generativa.

Se propone considerar los modelos de IA generativa como herramientas colaborativas, con un enfoque secuencial. Son las personas las que generan ideas que pueden ser refinadas o evaluar la viabilidad a través de la IA. Al mismo tiempo, se ha de seguir mejorando constantemente los datos e

información que la IA proporciona para que se mantengan las verdaderas aportaciones a la ciencia médica.

Conclusiones

La IA no es éticamente neutra, las competencias formativas han de contemplar las buenas prácticas en los proyectos de IA para que sean fiables desde el punto de vista técnico y ético, de modo que las mejoras que se introduzcan garanticen el respeto a los valores y derechos de las personas y de la sociedad.

Con independencia de que exista una responsabilidad por parte de los Estados (regulaciones), de las instituciones (hospitales, industria, empresas informáticas), persiste la responsabilidad ética de cada profesional de la salud, de cada investigador, por lo que es necesaria la formación y la capacitación de los profesionales para tener sentido crítico y poder valorar los diseños o los resultados de la IA. De otra manera, hay riesgo de abrir una brecha digital entre los que diseñan, desarrollan y deciden la utilización de procedimientos de IA, y los profesionales que las utilizan o aquellos a los que se les aplica, lo cual puede producir indefensión.

El diseño de los programas educativos ha de tener en cuenta también la diferencia intergeneracional que existe tanto entre los profesionales, como entre la población en general. Se ha de facilitar la comprensión a los pacientes y sujetos incluidos en los proyectos, y, en general a los ciudadanos, el significado de los datos y de los procesos de IA (Art. 85)¹⁵.

Son las personas y no la IA quienes establecen la visión, quienes tienen que diseñar la estrategia diferencial. Se han de construir espacios de reflexión ética que permitan dar razón de su uso, resolver posibles contradicciones y promover procedimientos basados en la búsqueda de soluciones buenas, no solo aceptables. En último término, el buen uso y las buenas prácticas estarán en función de las mejoras que los proyectos de IA puedan introducir en la atención al paciente individual y en la salud de la sociedad en general.

Financiación

La autora declara que no ha recibido financiación para la realización de este trabajo.

Puntos clave

- La IA no es éticamente neutra, como podría serlo cualquier metodología estadística. El diseño, el mantenimiento, la explotación de los algoritmos o de los proyectos de la IA son decisiones humanas.
- La IA ha de ser consistente y fiable, tanto desde el punto de vista técnico como social.
- Los resultados de la IA dependerán de los datos utilizados, del método de clasificación utilizado; de la aplicación del algoritmo que mejor se adapte a los objetivos del análisis que se va a realizar.
- Las características éticas más relevantes de la IA se refieren a la centralidad de la persona, a la equidad de los algoritmos (por los posibles sesgos de los resultados), a

la transparencia, a la seguridad, también al respeto de la privacidad y a la sostenibilidad.

- Es necesario incorporar procedimientos que contemplen la voluntariedad y respeto a los pacientes y a los profesionales respecto a la cesión, flujo y análisis de los datos. Se ha desarrollado el derecho a no ser perfilado.
- El buen uso y las buenas prácticas dependerán de las mejoras que los proyectos de IA puedan introducir en la atención al paciente individual y en la salud de la sociedad en general.
- Cuando se aplica la IA, persiste la responsabilidad ética y jurídica de cada profesional que interviene en el proceso.
- Los resultados de la IA no eliminan totalmente la incertidumbre de la decisión clínica. El razonamiento médico ha de tener en cuenta los valores, las necesidades y las prioridades de los pacientes individuales, lo cual no es una habilidad computable.
- Solo las personas pueden ser autoras. Los sistemas de IA no pueden conocer la realidad, lo que hacen es procesar, relacionar datos reales o no, sesgados o no.
- Los médicos y los profesionales que participen en la elaboración de algoritmos clínicos han de estar identificados y no deben colaborar en la manipulación intencionada de datos o de resultados obtenidos de las grandes bases de datos sanitarias.
- Es necesaria la formación y la capacitación de los profesionales para tener sentido crítico y poder valorar los diseños o los resultados de la IA.
- Sin formación hay riesgo de abrir una brecha digital entre los que diseñan, desarrollan y deciden la utilización de procedimientos de IA, y los profesionales que las utilizan o aquellos a los que se les aplica.

Conflicto de intereses

La autora declara no tener ningún conflicto de intereses.

Bibliografía

1. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med*. 2019;25:954–61. doi:[10.1038/s41591-019-0447-x](https://doi.org/10.1038/s41591-019-0447-x).
2. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402–10. doi:[10.1001/jama.2016.17216](https://doi.org/10.1001/jama.2016.17216).
3. Huang K, Chandak P, Wang Q, Havaladar S, Vaid A, Leskovec J, et al. A foundation model for clinician-centered drug repurposing. *Nat Med*. 2024;30:3601–13. doi:[10.1038/s41591-024-03233-x](https://doi.org/10.1038/s41591-024-03233-x).
4. Mittelstadt BD, Floridi L. The ethics of big data: current and foreseeable issues in biomedical contexts. *Sci Eng Ethics*. 2016;22:303–41. doi:[10.1007/s11948-015-9652-2](https://doi.org/10.1007/s11948-015-9652-2), 304.
5. Wang X, Zhao J, Marostica E, Yuan W, Jin J, Zhang J, et al. A pathology foundation model for cancer diagnosis and prognosis prediction. *Nature*. 2024;634:970–8. doi:[10.1038/s41586-024-07894-z](https://doi.org/10.1038/s41586-024-07894-z).
6. UNESCO. Report of the international bioethics committee of UNESCO on big data and health. Paris; 2017. [consultado 21 Jul 2025]. Disponible en: <https://unesdoc.unesco.org/ark:/48223/pf0000248724>

7. Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial. [consultado 21 Jul 2025]. Disponible en: <https://www.boe.es/buscar/doc.php?id=DOUE-L-2024-81079>
8. Reglamento (UE) 2025/327 del Parlamento Europeo y del Consejo, de 11 de febrero de 2025, relativo al Espacio Europeo de Datos de Salud, y por el que se modifican la Directiva 2011/24/UE y el Reglamento (UE) 2024/2847. [consultado 21 Jul 2025]. Disponible en: <https://www.boe.es/buscar/doc.php?id=DOUE-L-2025-80382>
9. Executive Office of the President. Big data: seizing opportunities, preserving values. Washington D.C., USA; 2014. [consultado 21 Jul 2025]. Disponible en: https://obamawhitehouse.archives.gov/sites/default/files/docs/20150204_Big_Data_Seizing_Opportunities_Preserving_Values_Memo.pdf
10. Nuffield Council on Bioethics. The collection, linking and use of data in biomedical research and health care: ethical issues; 2015. [consultado 21 Jul 2025]. Disponible en: https://cdn.nuffieldbioethics.org/wp-content/uploads/Biological_and_health_data-report.pdf
11. International Medical Informatics Association. The IMIA Code of Ethics for Health Information Professionals; 2011. [consultado 21 Jul 2025]. Disponible en: <https://imia-medinfo.org/wp-content/uploads/2015/07/IMIA-Code-of-Ethics-2016.pdf>
12. World Medical Association. Declaration of Taipei on ethical considerations regarding health databases and biobanks; 2016. [consultado 21 Jul 2025]. Disponible en: <https://www.wma.net/policies-post/wma-declaration-of-taipei-on-ethical-considerations-regarding-health-databases-and-biobanks/>
13. National Academy of Medicine and The Learning Health System Series, Whicher D., Ahmed M., Israni S.T. and Matheny M., *Artificial intelligence in health care: the hope, the hype, the promise, the peril*, 2023, National Academies Press (US); Washington (DC). [consultado 21 Jul 2025]. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK605955/>
14. World Medical Association. Declaration of Helsinki on ethical principles for medical research involving human participants; 2024, nn. 3 y 4. [consultado 21 Jul 2025]. Disponible en: <https://www.wma.net/policies-post/wma-declaration-of-helsinki/>
15. Organización Médica Colegial, Código de Deontología Médica. Madrid; 2022. [consultado 21 Jul 2025]. Disponible en: https://www.cgcom.es/sites/main/files/minisite/static/41a75446-a6e8-4df0-af58-7ca9cb625f1e/Codigo_Deontologia_2022/index.html
16. Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE, artículo 9. [consultado 21 Jul 2025]. Disponible en: <https://www.boe.es/buscar/doc.php?id=DOUE-L-2016-80807>
17. Leon-Sanz P. Key points for an ethical evaluation of healthcare big data. *Processes*. 2019;7(8):493. doi:10.3390/pr7080493.
18. Goodman K.W., *Ethics, medicine, and information technology: intelligent machines and the transformation of health care*, 2015, Cambridge University Press; Cambridge, UK. [consultado 21 Jul 2025]. Disponible en: <https://www.cambridge.org/core/books/ethics-medicine-and-information-technology/6E2D1B7B1A0AFA5C94C09CEFD58DFFC>
19. Laurie G, Jones KH, Stevens L, Dobbs C. A review of evidence relating to harm resulting from uses of health and biomedical data; 2014. [consultado 21 Jul 2025]. Disponible en: https://www.pure.ed.ac.uk/ws/portalfiles/portal/19402878/Review_of_Evidence_Relating_to_Harms_Resulting_from_Uses_of_Health_and_Biomedical_Data_FINAL.pdf
20. León P. Bioética y explotación de grandes conjuntos de datos. En: Carnicero J, Rojas D, editores. *La explotación de datos de salud: Retos, oportunidades y límites*. Pamplona, España: Sociedad Española de Informática de la Salud; 2016. p. 25–41.
21. Ley Orgánica 3/2018, de 5 de diciembre, de Protección de datos personales y garantía de los derechos digitales. [consultado 21 Jul 2025]. Disponible en: <https://www.boe.es/buscar/act.php?id=BOE-A-2018-16673>
22. Rumbold J, Pierscione B. Contextual anonymization for secondary use of big data in biomedical research: proposal for an anonymization matrix. *JMIR Med Inform*. 2018;6(4):e47. doi:10.2196/medinform.7096.
23. Sweeney L. k-anonymity: a model for protecting privacy. *Int J Uncertain Fuzz Knowl-based Syst*. 2002;10(5):557–70. doi:10.1142/S0218488502001648.
24. Anderson JG, Aydin CE. Evaluating the impact of health care information systems. *Int J Technol Assess Health Care*. 1997;13(2):380–93.
25. Zhao C, Li C, Li J, Chen F. Fair meta-learning for few-shot classification. 2020 IEEE International Conference on Knowledge Graph, Online, August 9–11, 2020. IEEE; 2020. p. 275–82.
26. González-Sendino R, Serrano E, Bajo J, Novais P. A review of bias and fairness in artificial intelligence. *Int J Interact Multimedia Artif Intell*. 2024;9(1):5–17. doi:10.9781/ijimai.2023.11.001.
27. Hill K. *Your face belongs to us: a secretive startup's quest to end privacy as we know it*. New York: Random House, an imprint of Penguin Random House LLC; 2023. p. p. 330.
28. Holm S. Fair by chance? On the use of algorithms in therapeutic decisions. *J Med Ethics*. 2025;jme-2025-110819. doi:10.1136/jme-2025-110819.
29. Kaebnick GE, Magnus DC, Kao A, Hosseini M, Resnik D, Dubljević V. Editors' statement on the responsible use of generative AI technologies in scholarly journal publishing. *Hastings Cent Rep*. 2023;53(5):3–6. doi:10.1002/hast.1507. Erratum in: *Hastings Cent Rep*. 2023 Nov;53(6):10. doi: <https://doi.org/10.1002/hast.1534>.