

Inteligencia artificial en salud y sus implicaciones bioéticas: una revisión sistemática

Mariele Abadia Elias¹, Luciana Arruda Favarsani¹, Josiane Aparecida Vieira Moreira¹, Anelise Viapiana Masieiro², Natalia Veronez da Cunha Bellinati¹

1. Universidade do Planalto Catarinense, Lages/SC, Brasil. 2. University of Iowa, Iowa City, United States of America.

Resumen

El uso de la inteligencia artificial en salud va en aumento por facilitar el diagnóstico y la toma de decisiones, pero sus implicaciones plantean dudas relacionadas con la ética. Esta revisión sistemática desde las directrices Prisma identificó los usos de la inteligencia artificial en salud y sus implicaciones bioéticas. Las búsquedas se realizaron en Science Direct y Scopus utilizando los descriptores “*artificial intelligence*”, “*bioethics*” y “*health*”. De los trabajos en inglés publicados entre 2017 y 2021, se obtuvo 102 artículos. Aplicados los criterios, quedaron 11. Los estudios abordaron los principios bioéticos de beneficencia, no maleficencia, autonomía y justicia, añadiendo el elemento explicabilidad. La inteligencia artificial se correlacionó con la imprevisibilidad, previsibilidad, confianza, papel de los médicos, desarrollo de sistemas, privacidad, seguridad de los datos y aspectos financieros y sociales. Los desarrolladores, los profesionales sanitarios y los pacientes deben maximizar los beneficios y limitar los riesgos que involucra esta tecnología.

Palabras clave: Atención a la salud. Aprendizaje automático. Macrodatos. Ética.

Resumo

Inteligência artificial em saúde e implicações bioéticas: uma revisão sistemática

A presença de inteligência artificial na saúde vem crescendo, ajudando em diagnósticos e tomadas de decisão, mas suas implicações geram dúvidas relacionadas à ética. Esta revisão sistemática, baseada nas diretrizes Prisma, identificou os usos de inteligência artificial na saúde e suas implicações bioéticas. Foi realizada busca nas bases de dados Science Direct e Scopus usando os descritores “*artificial intelligence*”, “*bioethics*” e “*health*”. Trabalhos em inglês, publicados entre 2017 e 2021 foram considerados, resultando em 102 artigos. Após aplicação dos critérios estabelecidos, 11 foram selecionados. Os estudos discutiram os princípios bioéticos da beneficência, não maleficência, autonomia e justiça, adicionando o elemento explicabilidade. Inteligência artificial mostrou correlação com imprevisibilidade, previsibilidade, confiança, papel do médico, desenvolvimento de sistemas, privacidade, segurança de dados, e aspectos sociais e financeiros. Desenvolvedores, profissionais da saúde e pacientes devem maximizar os benefícios e limitar os riscos das ferramentas que usam essa tecnologia.

Palavras-chave: Atenção à saúde. Aprendizado de máquina. Big data. Ética.

Abstract

Artificial intelligence in health and bioethical implications: a systematic review

The presence of artificial intelligence in healthcare is growing, helping in diagnosis and decision making. However, its application raises doubts, mostly related to ethics. This study aimed to identify its uses in health and its bioethical implications from a systematic literature review using the PRISMA guidelines. The ScienceDirect and Scopus databases were searched, using the descriptors “*artificial intelligence*,” “*bioethics*” and “*health*.” Works in English, published between 2017 and 2021 were considered, resulting in 102 articles found and, after applying the established criteria, 11 were selected. The studies reported on the bioethical principles of beneficence, non-maleficence, autonomy and justice, adding an element, explainability. Relationships were found between artificial intelligence in health and unpredictability, predictability, trust, physicians’ role, systems development, privacy, data security, financial and social aspects. Developers, healthcare professionals and patients must maximize the benefits and limit the risks of tools that use this technology.

Keywords: Health care. Machine learning. Big data. Ethics.

Las autoras declaran que no existe ningún conflicto de interés.

La inteligencia artificial (IA) está siendo cada vez más adoptada en diferentes ámbitos. El término en sí es difícil de definir, ya que este fenómeno depende de diferentes factores¹. A pesar de esta dificultad y de sus diversos conceptos, se entiende que la IA se asocia a máquinas y computadoras con el fin de ayudar a la humanidad a resolver problemas y facilitar procesos de trabajo².

Los sistemas de IA funcionan utilizando algoritmos complejos y grandes conjuntos de datos que generan conclusiones, reemplazando el razonamiento humano por el análisis de rutina³. Para lograr una inteligencia a nivel humano, la IA necesita orientación como modelo de realidad¹. El *machine learning* (ML), traducido como aprendizaje automático, es un sistema de inteligencia artificial que puede aprender de modelos y eventualmente volverse autónomo, tomando decisiones y generando conclusiones que antes se consideraban restringidas a la competencia de la mente humana³.

La expansión de la informática para almacenar, gestionar, acceder y procesar datos mediante una red de servidores remotos —como el almacenamiento en la nube— ha llevado a la expansión de las aplicaciones de IA al sector de la salud⁴. La IA y el ML tienen el potencial de revolucionar la prestación de servicios en esta área³. Estas tecnologías se pueden utilizar en la informática en salud, que describe la adquisición, el almacenamiento, la recuperación y el uso de información para mejorar la atención al paciente mediante interacciones con el sistema⁴. Las herramientas de *big data*, es decir, el almacenamiento y el análisis de datos voluminosos, como los utilizados en el ámbito de la salud, también se pueden utilizar en asociación con la IA⁵.

La IA ha mejorado el diagnóstico clínico y el desempeño en la toma de decisiones en múltiples dominios del trabajo médico⁶. Estas herramientas pueden ayudar a adaptar los programas de salud pública, asegurando que la información relevante esté disponible para políticas sólidas y para la toma de decisiones⁴. En la actualidad, el diagnóstico automatizado por imágenes es sin duda el dominio más exitoso del uso de la IA en el ámbito médico. Muchas especialidades médicas, incluida la radiología, oftalmología, dermatología y patología, dependen de diagnósticos basados en imágenes⁶.

Sin embargo, los pacientes reconocen y están empezando a enfrentar las muchas cuestiones que plantean las aplicaciones de la IA en el ámbito de la salud⁷. Estas exigencias son legales, comerciales, sociales y, sobre todo, éticas⁸. Diseñar estos sistemas

y utilizarlos no es un desafío meramente técnico, ya que requiere atención a principios bioéticos³.

La bioética se centra en la relación entre los seres vivos, y cuando surge la IA, los seres humanos deben comprometerse éticamente con algo que no es natural en sí mismo, es decir, con su propia creación². Crecen las preocupaciones por la posible pérdida de control en la relación humano-IA, como, por ejemplo, hasta qué punto la IA puede o debe apoyar decisiones médicas o incluso tomarlas¹.

A pesar de ser una tecnología clave hoy en día, en muchos casos será necesario entender cómo se tomó la decisión de una máquina y evaluar la explicación de tal elección¹. A medida que avanza la IA, los marcos bioéticos deben adaptarse para incluir los problemas que estos sistemas pueden presentar, del mismo modo que el desarrollo de estas tecnologías debe adaptarse para incorporar principios bioéticos³. Así, esta revisión sistemática de la literatura buscó identificar las aplicaciones de la inteligencia artificial en la salud y sus implicaciones bioéticas.

Método

Esta revisión sistemática siguió el Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA Statement)⁹, y utilizó la estrategia Población, Intervención, Comparación y Resultados (PICO). Este método se utiliza en la práctica basada en evidencias y se recomienda para estructurar la búsqueda bibliográfica de evidencias para revisiones¹⁰. La Tabla 1 describe la estrategia PICO utilizada en esta revisión para responder la pregunta: ¿Cuáles son las aplicaciones de la inteligencia artificial en salud y sus implicaciones bioéticas?

Tabla 1. Descripción de la estrategia PICO

P	individuos, personas
I	el uso de la inteligencia artificial en la salud
C	no corresponde
O	implicaciones bioéticas

La búsqueda de artículos se realizó de forma manual en las bases de datos ScienceDirect y Scopus en octubre del 2021. Se utilizaron los descriptores en ciencias de la salud (DeCS/MeSH) “inteligencia artificial”, “bioética” y “saúde”, interrelacionados mediante el operador booleano “AND”.

Los criterios de inclusión establecidos consistieron en artículos científicos publicados entre el 2017 y el 2021, en inglés y con acceso en línea.

Se excluyeron libros, capítulos de libros, tesis, disertaciones, trabajos presentados y publicados en eventos, artículos de revisión y notas editoriales.

Los títulos y resúmenes de los artículos encontrados se leyeron en una primera selección. Se leyeron íntegramente los artículos seleccionados, observando estrictamente los criterios de inclusión y exclusión, verificando la duplicidad y la adecuación al tema de la investigación. Los artículos seleccionados para la revisión se sistematizaron en un cuadro para el análisis de los resultados. Las etapas de selección y lectura completa de los artículos fueron realizadas por tres revisores independientes.

Resultados

En total, se identificaron 102 artículos en las bases de datos siguiendo los criterios de búsqueda. De ellos, 72 se encontraron en la base de datos ScienceDirect y 30 en Scopus. Tras la lectura de los títulos y resúmenes, se excluyeron 86 artículos por no abordar el tema propuesto. Un artículo se excluyó por que estaba duplicado. Así, seleccio-

namos 15 artículos potencialmente elegibles, de los cuales 11 se conservaron después de su lectura completa. La Figura 1 ilustra el proceso de selección.

Del total seleccionado para su inclusión en la revisión, siete (63,64%) publicaciones se recuperaron en ScienceDirect y cuatro (36,36%) en Scopus. En cuanto al año de publicación, siete (63,64%) de los artículos se publicaron en el 2021, y un artículo (9,09%) se publicó cada año en lo que resta del período establecido, es decir, entre el 2017 y el 2020. Al analizar el origen de los estudios, la mayoría (45,45%) se desarrolló en América del Norte o en colaboración con instituciones de ese continente.

En cuanto al tipo de estudio, la mayoría es transversal (45,45%), pero también se encontraron estudios prospectivos (9,09%), estudios de casos (9,09%) y estudios descriptivos (9,09%), estudios exploratorios (9,09%), ensayos clínicos aleatorizados (9,09%) y estudio de enfoque multidimensional (9,09%).

Con base en los 11 trabajos seleccionados, elaboramos una tabla para análisis de los objetivos, tipo de estudio, principales resultados y conclusiones (Tabla 2).

Figura 1. Diagrama de flujo del proceso de selección de estudios para la revisión sistemática

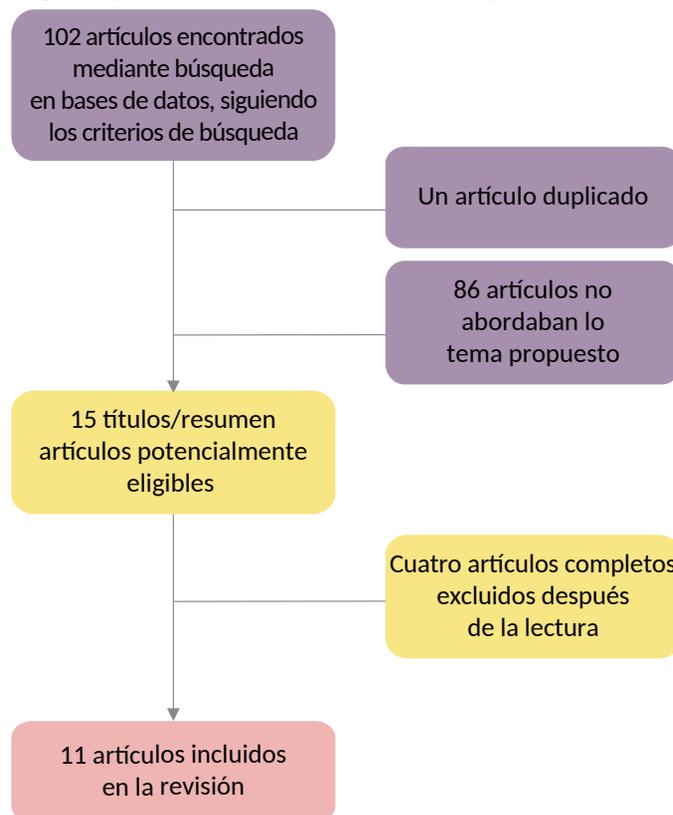


Tabla 2. Análisis de los principales datos presentados en las publicaciones: objetivo principal, tipo de estudio, principales resultados y conclusiones.

Autores; año	Objetivo principal	Tipo de estudio	Principales resultados	Conclusiones
Fernandes y colaboradores; 2019 ¹¹	Proporcionar resultados que puedan sustentar el estudio de la investigación médica en el proceso de toma de decisiones en bioética clínica, particularmente en casos de eutanasia.	Transversal	<p>Los datos procesados mediante métodos de selección de características se utilizaron para crear modelos capaces de predecir la decisión de eutanasia utilizando ML y un rastreador ocular.</p> <p>Los experimentos estadísticos demostraron que el modelo predictivo resultante del algoritmo <i>perceptron</i> multicapas (MLP) condujo a un mejor rendimiento.</p> <p>A partir de simulaciones con modelos MLP se extrajeron resultados interesantes (normas y reglas) para la toma de decisiones bioéticas.</p> <p>Algunos participantes tomaron una decisión racional, respetando el código de ética de los profesionales de enfermería y el código penal brasileño, que considera la eutanasia como homicidio. Otros consideraron el aspecto emocional, vinculando esta decisión al sufrimiento del paciente.</p>	<p>El buen desempeño presentado por el modelo predictivo demuestra que el enfoque de investigación propuesto puede usarse para probar hipótesis científicas relacionadas con la atención visual y la toma de decisiones, verificando en qué medida la visión es un factor determinante en la toma de decisiones, particularmente en la bioética clínica cuando se trata de cuestiones terminalidad de la vida.</p>
Silva, Lehoux, Hagemeister; 2018 ¹²	Evaluar si una innovación se califica como innovación responsable en salud mediante una herramienta desarrollada en tres etapas: triaje, evaluación y clasificación; y discutir los aspectos políticos del uso de la herramienta.	Delphi (prospectivo)	<p>El instrumento de triaje y evaluación de la innovación responsable en salud desarrollado fue juzgado por expertos y, tras la segunda ronda de comentarios sobre el tema, se alcanzó un consenso en 16 de las 20 preguntas sobre la importancia, claridad y adecuación de la estructura del instrumento.</p> <p>La sostenibilidad de los sistemas de salud se ve perjudicada por la forma actual en que se conciben y colocan en el mercado las innovaciones en salud. Hubo consenso en la mayoría de los criterios, atributos y escalas del instrumento.</p> <p>El uso futuro del instrumento puede contribuir al desarrollo de innovaciones que aporten mayor valor social.</p>	<p>El desarrollo de esta herramienta ayudará a llenar una importante laguna de conocimientos y políticas, aclarando las decisiones tomadas en una etapa inicial por las partes interesadas en la innovación, tales como inversionistas, creadores de tecnología, agencias de financiación de investigación y decisores políticos.</p>
Lysaght y colaboradores; 2019 ¹³	Abordar y analizar el marco ético para <i>big data</i> en salud e investigación para demostrar cómo puede fundamentar la toma de decisiones para el desarrollo y la implementación de sistemas de apoyo para el desarrollo y salud de forma ética y responsable ⁵ .	Estudio de caso	<p>Los sistemas de apoyo a la decisión clínica (CDSS) son programas que generan información en salud.</p> <p>Los que usan ML e IA son complejos y es por eso que los médicos deben capacitarse para mejorar sus habilidades de toma de decisiones clínicas al utilizar este tipo de recursos.</p> <p>Los algoritmos de IA del CDSS pueden reforzar los prejuicios sociales, pero también pueden aportar beneficios como sistemas de salud pública más eficientes.</p> <p>La inclusión y el análisis de datos se da de forma continua, aportando información sobre prácticas futuras adecuadas.</p> <p>Alimentar al CDSS con datos de pacientes puede causar conflictos sobre el doble papel de los médicos: atención e investigación.</p> <p>Los valores sustantivos enumerados son la integridad profesional y equidad; y los valores procesales son la transparencia y la rendición de cuentas.</p> <p>El CDSS asistido por IA deben ser explicable.</p> <p>La decisión final la debe tomar el profesional de salud y se debe considerar que pueden existir juicios morales que el programa no es capaz de realizar.</p> <p>El estudio de caso aborda el uso de estos programas en una unidad de cuidados intensivos. A pesar de aportar beneficios económicos, esta herramienta podría plantear cuestiones éticas como el ahorro en salud, la confianza en las recomendaciones y la preocupación por la responsabilidad.</p>	<p>Dado el aumento de los costos de los cuidados en salud, el desarrollo y la implementación de la asistencia basada en IA en la toma de decisiones clínicas tal vez sea inevitable.</p> <p>Los valores de integridad y responsabilidad profesional desempeñarán un papel más destacado en el ámbito de los cuidados brindados a los enfermos, mientras que los valores de justicia y el potencial de daño al grupo deben equilibrarse con imperativos de beneficio público a nivel social. Los imperativos del beneficio público a nivel social deben utilizarse para lograr el equilibrio.</p> <p>La transparencia afecta tanto a la confianza en la profesión médica como en los sistemas de salud.</p>

continúa...

Tabla 2. Continuación

Autores; año	Objetivo principal	Tipo de estudio	Principales resultados	Conclusiones
Cawthorne, Robbins-Van Wynsberghe; 2020 ¹⁴	Crear un marco ético para aplicarlo durante el diseño, desarrollo, implementación y evaluación de drones en salud pública.	Descriptivo	<p>La jerarquía de valores utilizada consiste en principios éticos, valores humanos, estándares y requisitos de diseño. La creación de la estructura consideró los cuatro principios bioéticos, además de un quinto principio de ética en IA: la explicabilidad. La benevolencia en el campo de la salud puede traducirse en valores de bienestar humano (y animal no humano), empleos y habilidades humanas y sostenibilidad ambiental. La no maleficencia abarca la privacidad, la seguridad, la protección, la tranquilidad, el empleo y las habilidades humanas, y la sostenibilidad ambiental. La autonomía incluye el libre albedrío, los valores humanos, la responsabilidad y la confianza. La justicia incluye la distribución equitativa de beneficios e indemnizaciones. La adopción de drones en la salud podría conducir a la reducción de la infraestructura de salud local, reduciendo la asistencia presencial. Sin embargo, también pueden conectar a personas en lugares remotos con servicios modernos. La explicabilidad se refiere a la facilidad con la que se pueden comprender los sistemas. El uso de un marco ético es especialmente útil para quienes tienen experiencia limitada en ética tecnológica.</p>	<p>Los principios éticos son abstractos y requieren una mayor contextualización y especificación para la reflexión. La creación de este marco ético refuerza el valor de integrar la ética en la práctica y sirve como modelo para el diseño y desarrollo en dominios de drones y no drones. El marco ayudó a identificar y refinar los posibles beneficios y mitigar los riesgos.</p>
Antes y colaboradores; 2021 ¹⁵	Desarrollar una nueva medida y evaluar la apertura y el alcance de las preocupaciones y los beneficios percibidos con respecto a las tecnologías sanitarias basadas en IA en una muestra de adultos en los Estados Unidos.	Transversal	<p>Los participantes se mostraron moderadamente abiertos a las tecnologías de salud basadas en IA, pero hubo variaciones según el tipo de aplicación. La confianza en el sistema de salud y en la tecnología fueron los correlatos más fuertes y consistentes de apertura, preocupación y beneficio percibido. Los participantes de mayor edad estaban menos abiertos a las tecnologías; y los hombres, más abiertos que las mujeres. El empleo a tiempo completo se asoció con una mayor apertura y con una menor preocupación. Las dos tecnologías que hicieron predicciones sobre enfermedades graves —riesgo de ataque cardíaco y probabilidad de supervivencia al cáncer— fueron las mejor calificadas. Un grupo de trabajo identificó cinco amplios dominios de actividad importantes para la colaboración utilizando datos de pacientes: privacidad, consentimiento informado, estandarización de elementos de datos, contratos con proveedores y evaluación de datos. Se presentaron los métodos y la comprensión ética de los marcos legales comúnmente utilizados para estos fines, así como el diseño del flujo de datos que puede ayudar a informar cómo se crean y se revocan los permisos. Cuando se discute el anonimato y la desidentificación para proteger la privacidad, es necesaria una descripción de la cuidadosa preparación y anotación de los conjuntos de datos, señalando técnicamente las dificultades. El volumen de datos requerido para la preparación del algoritmo de IA es muy alto, por lo que la preparación de estos datos debe ocurrir de manera segura y compartible con sus propietarios (pacientes).</p>	<p>La apertura de los participantes parece tenue, lo que sugiere que las estrategias de promoción temprana y los experimentos con nuevas tecnologías de IA podrían influir fuertemente en las opiniones sobre el tema. Puede que sea necesario abordar la confianza para impulsar la aceptación de estas innovaciones en el área de la salud.</p>
Batlle y colaboradores; 2021 ¹⁶	Comprender las mejores prácticas para compartir datos de pacientes en instituciones de salud.	Exploratorio	<p>Un grupo de trabajo identificó cinco amplios dominios de actividad importantes para la colaboración utilizando datos de pacientes: privacidad, consentimiento informado, estandarización de elementos de datos, contratos con proveedores y evaluación de datos. Se presentaron los métodos y la comprensión ética de los marcos legales comúnmente utilizados para estos fines, así como el diseño del flujo de datos que puede ayudar a informar cómo se crean y se revocan los permisos. Cuando se discute el anonimato y la desidentificación para proteger la privacidad, es necesaria una descripción de la cuidadosa preparación y anotación de los conjuntos de datos, señalando técnicamente las dificultades. El volumen de datos requerido para la preparación del algoritmo de IA es muy alto, por lo que la preparación de estos datos debe ocurrir de manera segura y compartible con sus propietarios (pacientes).</p>	<p>Crear una relación de intercambio de datos implica complejidad ética y de tecnología de la información. El anonimato y la privacidad del paciente mantienen la confianza y protegen a las entidades que buscan compartir datos de forma segura.</p>

continúa...



Tabla 2. Continuación

Autores; año	Objetivo principal	Tipo de estudio	Principales resultados	Conclusiones
Green y colaboradores; 2021 ¹⁷	Desarrollar herramientas integradas en sistemas digitales de salud para respaldar la toma de decisiones compartida y optimizar la preparación para el tratamiento de la enfermedad renal crónica.	Ensayo clínico aleatorizado	Al utilizar las herramientas, 243 (24%) de 1.032 pacientes en cuatro clínicas de nefrología se identificaron como de alto riesgo con progresión a insuficiencia renal en dos años. Los especialistas en transición renal inscribieron a 117 (48%) pacientes de alto riesgo hasta el final del primer año de investigación. Los enfermeros utilizaron la aplicación en el 100% de los pacientes para documentar 287 pasos de planificación de la terapia de reemplazo renal. Todos los especialistas en transición renal (100%) evaluaron la facilidad de uso y la utilidad del instrumento, estando de acuerdo o completamente de acuerdo con todos los ítems.	Los enfermeros informaron que las herramientas desarrolladas facilitaron la identificación de pacientes que necesitaban apoyo y sus actividades de navegación. Y la rápida identificación de pacientes que necesitan una toma de decisiones compartida e informada y su preparación para tratamientos de reemplazo renal.
Martinho, Kroesen, Chorus; 2021 ¹⁸	Obtener información sobre patrones de razonamiento y opiniones morales sobre la IA en salud de personas involucradas en la práctica médica.	Transversal	Con base en las preguntas de los médicos sobre la ética en torno a la IA en salud, se identificaron cuatro perspectivas principales: 1. La IA es una herramienta útil: permite a los médicos hacer aquello para lo que están capacitados. 2. Las reglas y reglamentos son cruciales: las empresas privadas solo se preocupan por el dinero. 3. La ética es suficiente: las empresas privadas son dignas de confianza. 4. Herramientas de IA explicables: el aprendizaje es necesario e inevitable. Todas las perspectivas consideran que los médicos deben participar en el proceso de creación de tecnologías de IA en salud, contribuyendo a la explicabilidad. Los médicos están más preocupados por el papel de las grandes empresas en el ámbito de la salud y menos concentrados o preocupados por cuestiones como la equidad, el prejuicio y las desigualdades en salud.	Cada perspectiva proporciona <i>insights</i> valiosos y, a menudo, contrastantes sobre cuestiones éticas que se deben operarizar y tener en cuenta al diseñar y desarrollar IA en salud.
Shen y colaboradores; 2021 ¹⁹	Analizar investigaciones de resonancia magnética (RM) altamente portátiles en entornos internacionales remotos y con recursos limitados para crear orientación ética y legal en un escenario mundial complejo.	Transversal	Es necesario asegurar que las comunidades locales colaboren en el trabajo de investigación y garantizar el valor social local del estudio. Los estudios de campo sobre RM deben asumir la responsabilidad de la seguridad de los participantes y de todos a su alrededor. Es importante prestar atención a las normas de privacidad y seguridad de datos (locales e internacionales). Es necesario identificar si la muestra en la que se entrenó el modelo de IA fue diversa para que las predicciones sean más precisas, teniendo en cuenta una diversidad de factores. Los resultados de las pruebas deben comunicarse a los participantes de forma aclaradora y, cuando se trata de hallazgos incidentales, una cuestión desafiante es cómo brindar apoyo clínico y derivación en estas comunidades remotas.	Los escáneres de resonancia magnética más asequibles y portátiles ofrecen oportunidades para abordar las necesidades de investigación insatisfechas y las iniquidades en salud en entornos internacionales remotos y con recursos limitados. Las comunidades locales deben ser colaboradoras permanentes en la cocreación de conocimientos. La investigación debe producir valor local para justificar los riesgos y minimizar la posibilidad de abuso.

continúa...

Tabla 2. Continuación

Autores; año	Objetivo principal	Tipo de estudio	Principales resultados	Conclusiones
Spiegel, Barker, Kistnasamy; 2021 ²⁰	<p>Describir y evaluar la aplicación de la IA en el desarrollo de diagnósticos asistidos por computadora para respaldar una adjudicación de reivindicaciones más eficiente para exmí- neros de oro con enfermedad pulmonar ocupacional en África Austral.</p>	Transversal	<p>Los resultados se correlacionaron con principios bioéticos. Beneficencia: la IA podría proporcionar un juicio más consistente que varios profesionales con diferentes niveles de habilidad. No maleficencia: mantener la privacidad y la seguridad de los datos y evitar la infiltración de ciertas tendencias en los sistemas de toma de decisiones. Autonomía: la IA podría provocar pérdidas en la cualificación profesional, a medida que los profesionales comiencen a confiar más en la tecnología. Se deben establecer protocolos y formar profesionales para identificar falsos negativos o falsos positivos. Justicia: cuando la demanda del mercado de inversión en IA es débil y donde las instituciones públicas no han respondido a su uso, el fracaso al abordar la tecnología puede ser una forma de mantener las desigualdades. Esto destaca el uso oportuno de las innovaciones para beneficiar a quienes las necesitan.</p>	<p>Se deben seguir desde el principio los esfuerzos para superar los desafíos técnicos en la aplicación de la IA para garantizar su uso ético.</p>
Stahl y colaboradores; 2021 ²¹	<p>Captar teóricamente y medir empíricamente los beneficios y desventajas de la IA para el progreso humano, más allá de los principios del aprendizaje automático, contrapesar los beneficios técnicos y económicos de la IA y sus aspectos legales, sociales y éticos.</p>	Enfoque multidimensional	<p>El discurso ético de la IA se discute en tres corrientes: 1. Problemas relacionados con la aplicación del ML: es difícil predecir en qué medida los datos se utilizarán para el fin determinado, ya que un perfil personal conduce a una clasificación para fines distintos al inicial. 2. Cuestiones sociales y políticas emergentes en una sociedad digital: estos sistemas requieren acceso a grandes cantidades de datos con fines de entrenamiento y validación, lo que genera desconfianza relacionada con la autonomía de las máquinas, sustitución de humanos por máquinas, injusticia en la distribución de costos y beneficios y control de datos, además de las consecuencias. 3. Cuestiones metafísicas sobre la naturaleza de la realidad y de la humanidad: se refieren a qué máquinas deberían poder decidir de forma autónoma. Las consecuencias económicas, el empleo, la justicia, la libertad, el contacto humano, la autonomía individual, la desigualdad, la integridad, la propiedad, el uso militar, la asimetría de poder, la responsabilidad y la sostenibilidad se encuadran en esta categoría.</p>	<p>Actualmente, no existe consenso entre los diferentes enfoques de gobernanza y seguridad de la información. La legislación de derechos humanos puede resolver muchas cuestiones sociales y éticas. La complejidad de los contextos y escenarios son factores que pluran las formas de abordarlos y permiten a los estudiosos o profesionales mantener una visión general. Atención al uso actual de IA y ML y, hasta cierto punto, a los sistemas sociotécnicos más amplios.</p>

continúa...

Discusión

Si bien son muy prometedoras para la salud, la ética y los derechos humanos de las nuevas tecnologías que utilizan la IA deben estar en el centro de su concepción, implementación y uso⁸. La incorporación de principios bioéticos en el desarrollo de estas tecnologías puede ayudar a proteger los derechos de los pacientes, minimizar riesgos, establecer responsabilidades e instituir métricas robustas para estudiar su efectividad y beneficios³.

Desde este punto de vista, algunos estudios seleccionados en esta revisión establecieron una relación directa entre sus resultados y los principios bioéticos de beneficencia, no maleficencia, autonomía y justicia^{14,20}. Otro principio considerado en la ética de la IA es la explicabilidad, que se ocupa del entendimiento de las herramientas que utilizan esta tecnología^{13,14,18}. El desarrollo de estos sistemas debe considerar los principios de beneficencia, mantenimiento de valores, responsabilidad y transparencia, lo que incluye la explicabilidad².

La explicabilidad es fundamental, ya que a partir de ella se puede generar la confianza necesaria y duradera para promover la aceptación de la IA entre los futuros usuarios¹. En un estudio, la confianza en el sistema de salud y en la tecnología fueron los correlatos más fuertes y consistentes de apertura, preocupación y beneficio percibido¹⁵. En otra investigación, los participantes se sintieron incómodos al confiar únicamente en las recomendaciones de una herramienta de IA, sin poder evaluar directamente su razonamiento⁷. Aunque los sistemas son importantes en el proceso de toma de decisiones, los expertos deben ser capaces de entender y de rehacer este proceso¹.

Incluso los expertos en desarrollo de IA no logran determinar la manera en que los insumos se transforman en resultados, es decir, cómo un perfil personal genera una decisión²¹. Por lo tanto, la imprevisibilidad en ML se vuelve clave para la discusión, ya que es un método para automatizar el análisis de datos utilizando algoritmos que identifican de forma iterativa patrones en los datos y aprenden de ellos⁴. Las consecuencias son imposibles de medir cuando la máquina está programada para aprender por sí sola²¹.

¿Qué deberían poder decidir las máquinas de forma autónoma y quién es el responsable de la decisión? Aún no hay respuestas a estas preguntas.

Por ahora, los médicos deben participar y entender la IA y, a continuación, decidir de forma autónoma, aunque basándose en las recomendaciones de la IA²¹.

Al igual que la imprevisibilidad, la predicción es otro punto relevante que se discute en la IA. Una herramienta capaz de predecir riesgos y priorizar pacientes para el tratamiento de la enfermedad renal crónica ayudó en la rápida identificación de individuos que necesitaban reemplazo renal¹⁷. Otro estudio evaluó la reacción y decisión de enfermeros sobre la práctica de la eutanasia en pacientes terminales utilizando tecnología de rastreo ocular (*eye-tracking*), trayendo una división entre decisiones racionales y emocionales, demostrando la efectividad del modelo de predicción¹¹. Incluso cuando la IA recomienda tratamientos, los profesionales de la salud y los pacientes pueden emitir juicios morales que el programa no es capaz de producir¹³.

Pese a las recomendaciones de la IA, la decisión final con relación a la salud debe ser tomada por un profesional¹³. Un estudio realizado en una unidad básica de salud constató que los pacientes atendidos creen que sus médicos deben protegerlos de los daños derivados de errores de IA, por lo que la decisión final debe ser del médico y de la aseguradora de salud⁷. En ese sentido y teniendo en cuenta las malas prácticas, las aseguradoras deben ser claras en cuanto a la cobertura cuando las decisiones son tomadas por sistemas de IA, aunque sea parcialmente⁶.

Es probable que la IA en el contexto de la salud dé a los médicos más tiempo para otras tareas, como establecer contacto directo con sus pacientes⁶. Sin embargo, la IA puede causar conflictos sobre el papel de los médicos, que deben brindar cuidados mientras alimentan el sistema con datos de pacientes con fines de apoyo e investigación clínica¹³. A medida que la IA se implemente para tareas clínicas específicas, las funciones de los profesionales de la salud seguirán evolucionando a la vez que se incorporen varios módulos de IA a los cuidados⁶.

La integridad profesional es uno de los tres principales valores éticos del *big data* en la toma de decisiones, ya que engloba la responsabilidad hacia el paciente, haciendo central la formación médica¹³. Los médicos tendrán que adaptarse a sus nuevas funciones como integradores e intérpretes de información y como apoyadores de los pacientes, y el sistema de educación médica tendrá

que proporcionarles las herramientas y los métodos necesarios⁶. Las herramientas de IA son útiles, pero los médicos no han sido capacitados para entenderlas y, a la vez, su participación activa en su desarrollo es vital¹⁸.

Para enfrentar los desafíos, los investigadores de IA y los médicos deben trabajar juntos para priorizar y desarrollar aplicaciones que satisfagan necesidades clínicas cruciales⁶. Ante el riesgo potencial de causar daño, es necesaria la participación de bioeticistas en el diseño de estas tecnologías³. Serán necesarias colaboraciones multidisciplinarias y multisectoriales para facilitar el desarrollo y la implementación de aplicaciones de IA en el ámbito médico⁶.

Se deben manipular varios conjuntos de datos para desarrollar herramientas de IA, lo que significa que esta tecnología solo puede avanzar mediante el uso de *big data*⁵. Para desarrollar y entrenar algoritmos, es necesario acceder a los datos de los registros médicos, y los pacientes no siempre son conscientes de que dicha información se comparte¹⁶. Este dilema requiere un estudio en profundidad de estas tecnologías emergentes y la evolución de los principios bioéticos de privacidad y confidencialidad del paciente³.

Las discusiones sobre la ética de los registros médicos electrónicos y de la IA se han centrado principalmente en la privacidad, la confidencialidad, la seguridad de los datos, el consentimiento informado y la propiedad de los datos. Sin embargo, la relevancia de cada uno varía según las diferencias culturales, la alfabetización, la relación paciente-profesional, la infraestructura disponible y la regulación de cada país⁴.

A pesar de muchas discusiones sobre la ética del *big data* en una variedad de contextos, hay poca orientación disponible sobre qué valores están en juego y cómo se deben tomar decisiones en un entorno de salud e investigación cada vez más complejo⁵. El anonimato y la privacidad del paciente son esenciales para generar confianza, además de la necesidad de enumerar los niveles de permiso en las conexiones entre la institución de salud, el titular de los datos (paciente) y las instituciones con ánimo de lucro relacionadas (planes de salud, proveedores y socios de tratamiento/examen)¹⁶.

Al considerar el uso de *big data* surgen varias cuestiones éticas complejas y, por lo tanto, es importante contar con un marco ético para abordarlas y

guiar las acciones⁵. Asimismo, esta planificación ética debería aplicarse a otros proyectos que impliquen IA en el ámbito de la salud, como marco que utilizar en todas las etapas de la obtención de drones en la salud pública que incluyan principios éticos, valores humanos, estándares y requisitos de salud¹⁴.

Se cuestiona el papel de las empresas de *big data* y IA, un segmento centrado en la ganancia, dadas las implicaciones éticas en la organización de la salud¹⁸. La sostenibilidad de los sistemas de salud se ve actualmente perjudicada por la forma en que se conciben y se llevan al mercado las innovaciones en salud¹². Los aspectos financieros y el poder concentrado en los tenedores de datos y en las tecnologías parecen ser más relevantes que los relacionados con la equidad, el prejuicio y las desigualdades.¹⁸

Aún se debe determinar quién controlará u obtendrá ganancias de la aplicación de la IA, por lo que la prioridad debe ser el equilibrio entre las salvaguardias regulatorias y las fuerzas del mercado para garantizar que los pacientes se beneficien⁶. La creación de normativas es un punto fundamental en la evolución de este tema¹⁸.

A pesar de tener un inmenso potencial para corregir errores humanos y mejorar la prestación de cuidados, las aplicaciones de IA basadas en ML pueden exacerbar los sesgos³. A los pacientes les preocupa que las herramientas de IA puedan reforzar los sesgos existentes, lo que puede ocurrir si un conjunto de datos de aprendizaje es tendencioso o si los desarrolladores incorporan involuntariamente su propio sesgo en un algoritmo⁷, lo que ya es una posibilidad con los algoritmos CDSS¹³.

La IA puede ser transformadora de aspectos sociales para la salud pública en los países de escasos recursos, a medida que estos lugares estén más conectados y creen datos de mayor calidad, es probable que aumente la capacidad de las herramientas de IA para enfrentar los desafíos de salud⁴.

La portabilidad es otro aspecto importante cuando se habla de la disponibilidad de la IA en ubicaciones remotas; por ejemplo, los escáneres de resonancia magnética más asequibles y portátiles ofrecen oportunidades para abordar las necesidades de salud insatisfechas y las iniquidades en entornos remotos y con recursos limitados¹⁹. La infraestructura requerida en estos entornos requiere importantes inversiones para su implementación, no siempre teniendo el acceso necesario para transmitir grandes conjuntos de datos a sistemas en la nube⁴. En este caso,

cobran relevancia las regulaciones de privacidad y seguridad de datos (locales e internacionales) y la forma de brindar apoyo y derivación clínica en estas comunidades ubicadas en lugares remotos¹⁹.

En un estudio sobre la IA en el diagnóstico asistido por computadora, se aceleraron los pedidos de indemnización para los trabajadores mineros con cáncer de pulmón ocupacional²⁰. Incluso con una baja demanda de inversiones públicas y privadas y un alto riesgo de descalificación de la herramienta en el ámbito legal, este enfoque proporciona un beneficio social comprobado cuando los protocolos éticos y el seguimiento son parte del proceso desde el principio.

Dos tecnologías predictivas de IA fueron mejor evaluadas entre ancianos e individuos con menos alfabetización digital, lo que demuestra que el alivio de enfermedades graves, como infarto y cáncer, casos que implican urgencia, vulnerabilidad y riesgo, puede aceptarse bajo una premisa de confiabilidad¹⁵.

Si bien los pacientes suelen mostrarse entusiasmados con la posibilidad de que la IA mejore sus cuidados, también les preocupan las posibilidades de seguridad y supervisión^{7,15}. También se plantean preocupaciones de que las herramientas de IA puedan aumentar los costos de salud y de que dichos costos puedan trasladarse a los pacientes⁷. Sin embargo, como las decisiones médicas se toman mediante sistemas automatizados, se pueden priorizar la rentabilidad y el racionamiento de gastos³.

El uso de la IA en la toma de decisiones clínicas probablemente sea inevitable dados los crecientes costos de la asistencia a la salud¹³. Los estudios prospectivos podrán identificar mejor las debilidades de los modelos de IA en entornos clínicos heterogéneos e indicar formas de integrarlos en los flujos de trabajo clínicos actuales⁶. Sin embargo, los esfuerzos para superar los desafíos técnicos en la implementación de la IA deben seguir aspectos éticos desde el principio²⁰.

Corroborando la literatura, se evidencian los beneficios y riesgos en el uso de la IA en la salud. Con el fin de maximizar los beneficios para el interés público y limitar las amenazas, la Organización Mundial de la Salud propone seis principios para el uso de la IA que se centran en los temas

cubiertos en esta discusión, a saber: 1) proteger la autonomía humana; 2) promover el bienestar, la seguridad y el interés público; 3) asegurar la transparencia, la explicabilidad y la inteligibilidad; 4) promover la responsabilidad y la rendición de cuentas; 5) garantizar la inclusión y la equidad; y 6) promover tecnologías responsivas y sostenibles⁸.

Consideraciones finales

Entender la IA en el ámbito de la salud requiere comprender sus pasos. Alimentación de datos, desarrollo de algoritmos y salida de decisiones. En este proceso, los cuatro principios bioéticos, además de los de explicabilidad y de legislación de derechos humanos, son los constructos que garantizan la ética.

La garantía de anonimato y del consentimiento para el uso de datos, sin fines comerciales de mayor relevancia en la implementación de la IA, siguen siendo diversificados y legalmente regionalizados en los procedimientos de investigación, diagnóstico y tratamiento. Por lo tanto, la transparencia es la base principal de la IA en salud, y las preguntas son más amplias que las respuestas.

Entender cómo se programaron las decisiones, sacándolas de la caja negra de la tecnología incomprensible, es el aspecto central. Cuando la máquina aprende por sí sola, ¿se puede responsabilizar al médico de sus decisiones? Las regulaciones deben evolucionar, tanto en la definición de responsabilidades como en el entendimiento y en la confianza en la autonomía de estas decisiones generadas por la IA. Permitir que las máquinas decidan por sí solas, basándose en su propio aprendizaje, es un concepto en evolución, ya que ni siquiera los expertos en este desarrollo pueden predecir sus consecuencias, tanto en términos de asertividad y confiabilidad, como en el mantenimiento del uso inicialmente definido.

Lo que podemos afirmar es que estamos en un camino sin retorno donde todas las cuestiones planteadas en este artículo deberán regularse considerando el acceso, la privacidad, cuestiones sociales y de justicia, sin dejar de estar tocadas por intereses de control, poder y rentabilidad.

Esta investigación no recibió subvenciones específicas de agencias de fomento de los sectores público, comercial o ánimo de lucro.

Referencias

1. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. *Wiley Interdiscip Rev Data Min Knowl Discov* [Internet]. 2019 [acesso 5 nov 2021];9(4):e1312. DOI: 10.1002/widm.1312
2. Tai MCT. The impact of artificial intelligence on human society and bioethics. *Tzu Chi Med J* [Internet]. 2020 [acesso 5 nov 2021];32(4):339-43. DOI: 10.4103/tcmj.tcmj_71_20
3. Nabi J. How bioethics can shape artificial intelligence and machine learning. *Hastings Cent Rep* [Internet]. 2018 [acesso 5 nov 2021];48(5):10-3. DOI: 10.1002/hast.895
4. Wahl B, Cossy-Gantner A, Germann S, Schwalbe NR. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? *BMJ Glob Heal* [Internet]. 2018 [acesso 5 out 2021];3(4):e000798. DOI: 10.1136/bmjgh-2018-000798
5. Xafis V, Schaefer GO, Labude MK, Brassington I, Ballantyne A, Lim HY *et al.* An ethics framework for big data in health and research. *Asian Bioeth Rev* [Internet]. 2019 [acesso 5 nov 2021];11(3):227-254. DOI: 10.1007/s41649-019-00099-x
6. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng* [Internet]. 2018 [acesso 5 nov 2021];2(10):719-731. DOI: 10.1038/s41551-018-0305-z
7. Richardson JP, Smith C, Curtis S, *et al.* Patient apprehensions about the use of artificial intelligence in healthcare. *NPJ Digit Med* [Internet]. 2021 [acesso 5 nov 2021];4(1):140. DOI: 10.1038/s41746-021-00509-1
8. World Health Organization. Ethics and governance of artificial intelligence for health [Internet]. Geneva: WHO Guidance; 2021 [acesso 4 nov 2021]. Disponível: <https://bit.ly/3Ry4M8P>
9. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* [Internet]. 2009 [acesso 5 set 2021];339(7716):332-6. DOI: 10.1136/bmj.b2535
10. Santos CMDC, Pimenta CADM, Nobre MRC. A estratégia PICO para a construção da pergunta de pesquisa e busca de evidências. *Rev Latinam Enferm* [Internet]. 2007 [acesso 5 set 2021];15(3):508-11. DOI: 10.1590/S0104-11692007000300023
11. Fernandes DL, Siqueira-Batista R, Gomes AP, Souza CR, Costa IT, Cardoso FSL *et al.* Investigation of the visual attention role in clinical bioethics decision-making using machine learning algorithms. *Procedia Comput Sci* [Internet]. 2017 [acesso 11 out 2022];108:1165-74. DOI: 10.1016/j.procs.2017.05.032
12. Silva HP, Lehoux P, Hagemester N. Developing a tool to assess responsibility in health innovation: results from an international delphi study. *Heal Policy Technol* [Internet]. 2018 [acesso 11 out 2022];7(4):388-96. DOI: 10.1016/j.hlpt.2018.10.007
13. Lysaght T, Lim HY, Xafis V, Ngiam KY. AI-assisted decision-making in healthcare. *Asian Bioeth Rev* [Internet]. 2019 [acesso 11 out 2022];11(3):299-314. DOI: 10.1007/s41649-019-00096-0
14. Cawthorne D, Robbins-van Wynsberghe A. An ethical framework for the design, development, implementation, and assessment of drones used in public healthcare. *Sci Eng Ethics* [Internet]. 2020 [acesso 11 out 2022];26(5):2867-91. DOI: 10.1007/s11948-020-00233-1
15. Antes AL, Burrous S, Sisk BA, Schuelke MJ, Keune JD, DuBois JM. Exploring perceptions of healthcare technologies enabled by artificial intelligence: an online, scenario-based survey. *BMC Med Inform Decis Mak* [Internet]. 2021 [acesso 11 out 2022];21(1):221. DOI: 10.1186/s12911-021-01586-8
16. Batlle JC, Dreyer K, Allen B, Cook T, Roth CJ, Kitts AB *et al.* Data sharing of imaging in an evolving health care world: report of the ACR Data Sharing Workgroup, part 1: data ethics of privacy, consent, and anonymization. *J Am Coll Radiol* [Internet]. 2021 [acesso 11 out 2022];18(12):1646-54. DOI: 10.1016/j.jacr.2021.07.014
17. Green JA, Ephraim PL, Hill-Briggs F, Browne T, Strigo TS, Hauer CL *et al.* Integrated digital health system tools to support decision making and treatment preparation in CKD: the PREPARE NOW Study. *Kidney Med* [Internet]. 2021 [acesso 11 out 2022];3(4):565-75.e1. DOI: 10.1016/j.xkme.2021.03.009

18. Martinho A, Kroesen M, Chorus C. A healthy debate: exploring the views of medical doctors on the ethics of artificial intelligence. *Artif Intell Med* [Internet]. 2021 [acceso 11 out 2022];121(October):102190. DOI: 10.1016/j.artmed.2021.102190
19. Shen FX, Wolf SM, Bhavnani S, Deoni S, Elison JT, Fair D *et al.* Emerging ethical issues raised by highly portable MRI research in remote and resource-limited international settings. *Neuroimage* [Internet]. 2021 [acceso 11 out 2022];238:118210. DOI: 10.1016/j.neuroimage.2021.118210
20. Spiegel JM, Ehrlich R, Yassi A, Riera F, Wilkinson J, Lockhart K *et al.* Using artificial intelligence for high-volume identification of silicosis and tuberculosis a bio-ethics approach. *Ann Glob Heal* [Internet]. 2021 [acceso 11 out 2022];87(1):58. DOI: 10.5334/aogh.3206
21. Stahl BC, Andreou A, Brey P, Hatzakis T, Kirichenko A, Macnish K *et al.* Artificial intelligence for human flourishing: beyond principles for machine learning. *J Bus Res* [Internet]. 2021 [acceso 11 out 2022];124:374-88. DOI: 10.1016/j.jbusres.2020.11.030

Mariele Abadia Elias – Graduada – marieleabadia@uniplaclages.edu.br

 0000-0001-5724-8400

Luciana Arruda Faversoni – Graduada – luarrudaf@uniplaclages.edu.br

 0000-0002-5495-1122

Josiane Aparecida Vieira Moreira – Graduada – josiane@uniplaclages.edu.br

 0000-0003-2683-5172

Anelise Viapiana Masiero – Doctora – anemasiero@uniplaclages.edu.br

 0000-0003-3203-5876

Natalia Veronez da Cunha – Doctora – nat_cunha@uniplaclages.edu.br

 0000-0002-8522-5836

Correspondencia

Mariele Abadia Elias – Av. Marechal Castelo Branco, 170, bloco 2, Universitário CEP 88509-900. Lages/SC, Brasil.

Participación de las autoras

Mariele Abadia Elias, Luciana Arruda Faversoni y Josiane Aparecida Vieira Moreira fueron responsables de la conceptualización, método, investigación, redacción y elaboración del artículo original. Anelise Viapiana Masiero y Natalia Veronez da Cunha Bellinati fueron responsables de la preparación, revisión y edición del artículo.

Recibido: 7.4.2023

Revisado: 6.7.2023

Aprobado: 13.10.2023