

Entre o algoritmo e o Juramento de Hipócrates: bioética na era da inteligência artificial

Juracy Barbosa dos Santos¹, Guilhermina Rego¹, Rui Nunes¹

1. Faculdade de Medicina da Universidade do Porto, Porto, Portugal.

Resumo

O avanço da inteligência artificial tem transformado profundamente a prática médica. De sistemas de apoio à decisão clínica a algoritmos de triagem e diagnóstico, a inteligência artificial tem demonstrado potencial para diagnósticos precoces, terapias personalizadas, otimização de recursos, redução de erros e ampliação do acesso a cuidados especializados. Contudo, essa revolução tecnológica impõe desafios éticos significativos aos princípios clássicos da bioética, como beneficência, não maleficência, confidencialidade e respeito à autonomia do doente, consagrados desde o Juramento de Hipócrates. A incorporação da inteligência artificial na medicina suscita tensões entre eficiência automatizada e valores humanos. Este artigo revisa a literatura sobre os principais dilemas éticos da aplicação da inteligência artificial na prática médica e avalia em que medida os princípios hipocráticos podem ser preservados ou adaptados diante das transformações tecnológicas, a fim de contribuir para o debate sobre os rumos éticos da medicina contemporânea em um cenário de crescente automação.

Palavras-chave: Bioética. Inteligência artificial. Ética médica. Autonomia pessoal.

Resumen

Entre el algoritmo y el Juramento Hipocrático: bioética en la era de la inteligencia artificial

El avance de la inteligencia artificial ha transformado profundamente la práctica médica. Desde sistemas de apoyo a la decisión clínica hasta algoritmos de triaje y diagnóstico, la inteligencia artificial ha demostrado potencial para diagnósticos precoces, terapias personalizadas, optimización de recursos, reducción de errores y ampliación del acceso a cuidados especializados. No obstante, esta revolución tecnológica impone desafíos éticos significativos a los principios clásicos de la bioética, tales como la beneficencia, la no maleficencia, la confidencialidad y el respeto a la autonomía del paciente, consagrados desde el Juramento Hipocrático. La incorporación de la inteligencia artificial en la medicina suscita tensiones entre la eficiencia automatizada y los valores humanos. Este artículo revisa la literatura sobre los principales dilemas éticos de la aplicación de la inteligencia artificial en la práctica médica y evalúa en qué medida los principios hipocráticos pueden preservarse o adaptarse ante las transformaciones tecnológicas, con el fin de contribuir al debate sobre los rumbos éticos de la medicina contemporánea en un escenario de creciente automatización.

Palabras clave: Bioética. Inteligencia artificial. Ética médica. Autonomía personal.

Abstract

Between the algorithm and the Hippocratic Oath: bioethics in the age of artificial intelligence

The advancement of artificial intelligence has profoundly transformed medical practice. From clinical decision-support systems to screening and diagnostic algorithms, artificial intelligence has demonstrated potential for early diagnosis, personalized therapies, resource optimization, error reduction, and expanded access to specialized care. However, this technological revolution imposes significant ethical challenges on the classic principles of bioethics, such as beneficence, non-maleficence, confidentiality, and respect for patient autonomy, enshrined since the Hippocratic Oath. The incorporation of artificial intelligence into medicine raises tensions between automated efficiency and human values. This article reviews the literature on the primary ethical dilemmas of applying artificial intelligence in medical practice and evaluates to what extent Hippocratic principles can be preserved or adapted in the face of technological transformations, aiming to contribute to the debate on the ethical directions of contemporary medicine in a scenario of increasing automation.

Keywords: Bioethics. Artificial intelligence. Ethics, medical. Personal autonomy.

Declararam não haver conflito de interesse.

A introdução da inteligência artificial (IA) na medicina representa um dos maiores saltos tecnológicos da história recente da saúde. De sistemas de apoio à decisão clínica a algoritmos de triagem e diagnóstico, a IA tem demonstrado potencial para diagnósticos precoces, terapias personalizadas, otimização de recursos, redução de erros médicos e ampliação do acesso a cuidados especializados, especialmente em regiões remotas ou com escassez de profissionais de saúde qualificados^{1,2}. Em áreas como dermatologia, radiologia e cardiologia, algoritmos baseados em *deep learning* (aprendizado profundo) já demonstram desempenho comparável, ou mesmo superior, ao de especialistas humanos em tarefas específicas³.

No entanto, a rápida expansão e incorporação da IA à prática clínica não ocorre sem desafios. A adoção acelerada dessas tecnologias suscita preocupações éticas significativas, especialmente no que diz respeito aos princípios fundamentais da bioética. Desde seus primórdios, a medicina tem sido regida por princípios éticos que visam à proteção e ao bem-estar do paciente. O Juramento de Hipócrates, datado do século V a.C., consagrou valores fundamentais como beneficência, não maleficência, confidencialidade e respeito à autonomia do doente – pilares da bioética ainda hoje⁴. Tais valores foram reforçados no século XX por documentos como o *Código de Nuremberg* e a *Declaração de Helsinki*, que consolidaram as bases para a ética em pesquisa e prática médica³.

Contudo, o advento da IA na medicina introduz novos e complexos desafios éticos que tensionam esses princípios. Entre os dilemas que emergem estão a opacidade e a não explicabilidade dos algoritmos – os chamados *black boxes* –, que dificultam a compreensão e a validação das decisões automatizadas por parte de médicos e pacientes, e essa falta de transparência algorítmica compromete o princípio da autonomia, pois impede o consentimento plenamente informado. Além disso, surgem questões sobre a responsabilização por erros clínicos decorrentes de decisões baseadas em IA, com dúvidas sobre se a culpa recai sobre o médico, o programador, a instituição de saúde ou o próprio algoritmo^{2,5,6}.

A privacidade e a proteção de dados sensíveis dos pacientes são outros pontos críticos. Como a IA depende de grandes volumes de dados para treinamento e operação, é essencial garantir que

o tratamento das informações esteja em conformidade com legislações e regulamentações, como a Lei 14.510/2022, que dispõe sobre a prestação de serviços de telessaúde⁷, a Lei Geral de Proteção de Dados (LGPD, Lei 13.709/2018)⁸, e a Resolução do Conselho Federal de Medicina (CFM) 2.314/2022, que define e regulamenta a telemedicina como forma de prestação de serviços médicos⁹, no Brasil; o General Data Protection Regulation (GDPR)¹⁰ e a Network and Information Systems (NIS Directive)¹¹, na União Europeia; o Health Insurance Portability and Accountability Act (HIPAA)¹² e o Health Information Technology for Economic and Clinical Health Act (HITECH)¹³, nos Estados Unidos. A violação da privacidade ou o uso indevido de dados médicos podem ter consequências devastadoras para os pacientes, incluindo discriminação, estigmatização e danos morais e financeiros.

Adicionalmente, há o risco de ampliar desigualdades no acesso à saúde. A implantação de soluções baseadas em IA, muitas vezes desenvolvidas em contextos de alta renda, pode não considerar adequadamente as necessidades, idiomas e realidades socioculturais de populações vulneráveis ou de países em desenvolvimento. Isso pode acentuar disparidades já existentes no sistema de saúde e comprometer o princípio da justiça distributiva, fundamental para a equidade no cuidado médico¹⁴.

Também merece preocupação a possível desumanização na medicina. A relação médico-paciente, historicamente centrada na escuta, no cuidado e na empatia, pode ser afetada negativamente se a confiança for deslocada para sistemas automatizados. O excesso de dependência tecnológica pode tornar o atendimento mais impessoal, comprometendo a integralidade do cuidado e a dimensão subjetiva da experiência do adoecimento¹⁵. A formação médica, nesse cenário, também deve ser repensada, uma vez que habilidades como pensamento crítico, julgamento clínico e sensibilidade ética precisam ser preservadas e fortalecidas, mesmo em ambiente cada vez mais orientado por dados¹⁶.

O descompasso entre a velocidade da inovação tecnológica e a capacidade de regulação ética e jurídica impõe, portanto, a necessidade urgente de revisar os fundamentos bioéticos sob nova perspectiva. Tais questões exigem não

apenas análise técnica, mas abordagem bioética rigorosa, que considere o impacto da IA nos valores fundamentais que regem a prática médica. A literatura recente reforça que princípios éticos isolados não são suficientes para garantir a integridade moral das aplicações de IA: é necessário um sistema de governança robusto, transparente, participativo e que envolva múltiplos stakeholders, incluindo médicos, engenheiros, pacientes, juristas e representantes da sociedade civil^{17,18}.

Este artigo tem como objetivo revisar a literatura existente sobre os principais dilemas éticos relacionados à aplicação da inteligência artificial na prática médica, resgatar os princípios hipocráticos e analisar como eles podem (ou não) ser preservados diante das transformações tecnológicas em curso. Por meio de análise crítica e multidisciplinar, pretende-se contribuir para o debate sobre os rumos éticos da medicina contemporânea em um cenário de crescente automação.

Método

Este artigo caracteriza-se como pesquisa de natureza qualitativa, do tipo exploratória, com abordagem metodológica baseada em revisão bibliográfica narrativa. O objetivo é reunir, analisar e interpretar criticamente a literatura científica acerca dos dilemas éticos emergentes na aplicação da IA artificial na prática médica, à luz dos princípios clássicos da bioética, conforme estabelecidos no Juramento de Hipócrates.

Segundo Gil¹⁹, a pesquisa bibliográfica permite o exame de contribuições teóricas já consolidadas e é especialmente útil em estudos que visam compreender e discutir temas ainda em consolidação no meio acadêmico, como é o caso da interface entre bioética e IA. Abordagem exploratória, por sua vez, é adequada para o aprofundamento inicial de um problema ainda pouco delimitado, na medida em que permite identificar categorias e lacunas relevantes para futuras investigações empíricas²⁰.

O percurso metodológico seguiu as seguintes etapas:

1) Definição do tema e do problema de pesquisa mediante delimitação do escopo bioético da atuação da IA em contextos clínicos e hospitalares.

2) Levantamento bibliográfico sistemático, realizado nas bases de dados PubMed, Scopus,

SciELO e Web of Science, com os seguintes descritores e combinações: “bioethics”, “artificial intelligence”, “machine learning”, “medical ethics”, “Hippocratic Oath”, “clinical decision support”, “algorithmic bias”, entre outros, com o auxílio do operador booleano “and”. Foram incluídos artigos publicados entre 2015 e 2024, nos idiomas português e inglês, que abordassem diretamente os aspectos éticos da IA em saúde.

3) Seleção e triagem das fontes com base na leitura dos títulos, resumos e, posteriormente, textos completos. Os critérios de inclusão foram: relevância para o tema, publicação em periódicos revisados por pares e contribuição teórica ou prática para o debate ético. Foram excluídas publicações duplicadas, resumos de conferências sem texto completo e artigos com foco estritamente técnico ou computacional da IA, sem discussão ética.

4) Análise e síntese dos resultados por meio de leitura crítica e categorização temática, com ênfase em tópicos recorrentes, como autonomia médica, responsabilidade por decisões automatizadas, explicabilidade algorítmica e equidade no acesso a tecnologias baseadas em IA.

A metodologia adotada visa não apenas mapear o estado da arte dos dilemas bioéticos da IA médica, mas também contextualizá-los à luz dos princípios hipocráticos clássicos – como beneficência, não maleficência e confidencialidade – e, com isso, identificar pontos de convergência e tensão com nossos paradigmas tecnológicos. Tal abordagem contribui para o amadurecimento crítico do debate e fornece subsídios conceituais para formulação de políticas públicas e diretrizes clínicas e desenvolvimento de tecnologias mais sensíveis aos princípios éticos da medicina.

Discussão

A ética médica entre Hipócrates e a era digital

Desde a Antiguidade, a medicina tem se norteadado por princípios éticos que refletem compromisso moral com o bem-estar do paciente. O Juramento de Hipócrates, datado do século V a.C., sintetiza esse legado e constitui-se em marco simbólico e normativo da ética médica ocidental, na medida em que enuncia compromissos

fundamentais como promoção do bem, abstenção de causar danos, respeito à confidencialidade e conduta moral do profissional de saúde³. Esses valores, que deram origem aos pilares bioéticos da beneficência, não maleficência, autonomia e justiça, permaneceram como bússola moral ao longo de séculos, mesmo diante das transformações científicas e socioculturais. Contudo, a ascensão da inteligência artificial representa novo ponto de inflexão, pois desafia a aplicabilidade e a suficiência dessas diretrizes tradicionais para orientar práticas médicas mediadas por tecnologias que tomam decisões, aprendem autonomamente e operam em escala².

Tal transformação cria desafios éticos, pois a lógica algorítmica que rege os sistemas de IA nem sempre é transparente ou compatível com os valores humanísticos da medicina tradicional⁴. A medicina digital, ao incorporar algoritmos de *machine learning*, sistemas de apoio a decisão clínica e tecnologias preditivas baseadas em *big data*, expande as possibilidades diagnósticas e terapêuticas, mas também desloca a centralidade do juízo clínico humano. Ou seja, a IA pode ampliar exponencialmente a capacidade diagnóstica e preditiva da prática médica, mas também deslocar a autoridade decisória do profissional para modelos estatísticos complexos, frequentemente opacos. Nesse contexto, surge a tensão entre a tradição ética – ancorada na relação interpessoal e no julgamento prudente – e a nova racionalidade algorítmica, cuja lógica frequentemente se torna incompreensível, não auditável ou incontrolável².

Essa transição tecnológica impõe à bioética tradicional o desafio de reconfigurar seus fundamentos para que esta seja capaz de orientar práticas em ambientes tecnologizados, ou seja, nos quais as decisões não são mais tomadas exclusivamente por sujeitos humanos, mas por sistemas híbridos em que o médico delega ou compartilha responsabilidades com máquinas^{5,21}. A medicina do futuro será, assim, além de prática humana apoiada pela técnica, uma coprodução entre humanos e algoritmos⁶. Tal cenário exige profunda reconfiguração ética, capaz de lidar com questões de responsabilidade compartilhada, justiça algorítmica, transparência e governança de dados. Ao mesmo tempo, a ética hipocrática continua a oferecer referencial fundamental, desde que reinterpretada à luz das novas mediações tecnológicas. Como destacam

Floridi e colaboradores¹⁷, o desenvolvimento ético da IA deve integrar princípios humanísticos desde sua concepção, visando à construção de uma sociedade digital justa, transparente e digna.

Autonomia e opacidade algorítmica

O princípio da autonomia é central na ética médica moderna e representa o direito do paciente de tomar decisões informadas sobre seu próprio cuidado. A realização plena desse princípio, fundamentado na transparência da relação médico-paciente, depende de comunicação clara, compreensão dos riscos e participação ativa do paciente no processo terapêutico. Contudo, a ascensão da IA, especialmente em formas de *deep learning*, ameaça minar essa transparência essencial. Algoritmos complexos, treinados com grandes volumes de dados, frequentemente operam como “caixas-pretas”, ou seja, são incapazes de explicar, de forma compreensível, como chegaram a determinada recomendação^{1,2,4,22}.

Tal opacidade compromete a base do consentimento informado, já que nem mesmo profissionais de saúde compreendem integralmente os critérios utilizados pelos algoritmos. Como destacam Morley e colaboradores²¹, quando o médico não consegue explicar ao paciente como uma decisão foi tomada – seja um diagnóstico, classificação de risco ou proposta terapêutica –, o próprio pacto ético que fundamenta a autonomia acaba sendo abalado. O paciente, em vez de decidir livremente, torna-se dependente de um sistema cuja lógica lhe é inacessível.

Além disso, a opacidade algorítmica introduz novas camadas de assimetria informacional. Pacientes com menor escolaridade, acesso digital limitado ou baixo letramento em saúde são ainda mais vulneráveis a perda de autonomia, uma vez que enfrentam barreiras adicionais à compreensão dos processos decisórios. Como defendido por Gerke, Minssen e Cohen, a autonomia não é apenas direito individual, mas fenômeno relacional que exige contextos adequados para ser exercida – entre eles, o acesso a informação clara e inteligível⁶.

A questão se agrava quando consideramos a ambiguidade em torno da responsabilidade ética e jurídica das decisões tomadas com auxílio de IA. Quando uma decisão clínica é tomada com base em recomendação algorítmica, quem deve ser

responsabilizado por seus resultados? O médico? O desenvolvedor do algoritmo? A instituição^{5?} A ausência de clareza quanto à responsabilização moral – o que é chamado por alguns autores de “vácuo ético”²³ – compromete não apenas a autonomia, mas a própria legitimidade do cuidado.

Nesse cenário, torna-se urgente avançar na direção de sistemas mais explicáveis e auditáveis²⁴. A transparência algorítmica, aliada ao desenvolvimento de mecanismos participativos de consentimento e mediação, é condição essencial para preservar a autonomia do paciente em uma era dominada por decisões automatizadas.

Beneficência e risco de viés algorítmico

A beneficência, enquanto obrigação moral de agir no melhor interesse do paciente, adquire novos contornos em uma medicina mediada por inteligência artificial. Ferramentas de IA têm demonstrado, em diversos estudos, acurácia comparável ou superior à humana em tarefas como interpretação de imagens radiológicas para classificação de câncer de pele¹ e predição de eventos cardiovasculares²². No entanto, essa superioridade técnica não implica neutralidade ética. Como advertem alguns autores, o desempenho dos algoritmos depende fortemente da qualidade, diversidade e representatividade dos dados utilizados em seu treinamento²¹.

Diversos estudos apontam que sistemas treinados com dados historicamente enviesados tendem a replicar e até amplificar desigualdades preexistentes. Um exemplo emblemático é o algoritmo de gerenciamento populacional de saúde que subestimava a gravidade da condição de pacientes negros, pois utilizava como *proxy* o custo com cuidados de saúde¹⁴ – um indicador influenciado por desigualdade estrutural. A aplicação acrítica desses sistemas pode, portanto, violar o princípio da beneficência por produzir resultados discriminatórios e ineficazes para determinados grupos.

Além disso, o risco de viés algorítmico não se limita a questões raciais ou socioeconômicas. Até mesmo sistemas desenvolvidos para recrutamento automatizado apresentaram viés de gênero, penalizando candidatas mulheres²⁵, e, ainda, demonstraram desempenho inferior em populações negras, idosos ou pacientes com

comorbidades múltiplas, refletindo vieses de representatividade dos dados de treinamento. Quando esse tipo de lógica é transferido para o campo da saúde, o potencial de dano ético aumenta, pois decisões clínicas podem determinar desfechos de vida ou morte. Em vez de promover o bem, a IA pode, inadvertidamente, reforçar padrões discriminatórios e dificultar o acesso equitativo a cuidados de qualidade – o que infringe o princípio da beneficência e fere o ideal hipocrático de tratar igualmente todos os pacientes.

Nesse contexto, a beneficência exige não apenas eficácia técnica, mas também vigilância constante sobre a equidade dos sistemas de IA. É necessário desenvolver metodologias de auditoria algorítmica capazes de identificar padrões discriminatórios, corrigir vieses e garantir que os benefícios da IA sejam distribuídos de forma justa⁶. A avaliação ética deve fazer parte do próprio ciclo de vida do sistema, desde o *design* até a implementação clínica²⁵.

Não maleficência e danos invisíveis

A não maleficência, princípio ético expresso pela máxima “*primum non nocere*” – ou seja, “primeiro, não causar dano” –, exige do profissional o compromisso de evitar ou minimizar quaisquer prejuízos à saúde e ao bem-estar do paciente durante o cuidado médico. No contexto da IA, no entanto, esse entendimento deve ser ampliado para englobar riscos não intencionais, já que os danos nem sempre são visíveis, imediatos ou diretos, o que não minimiza o potencial de prejuízos significativos à saúde, à integridade ou à dignidade dos pacientes. O uso de sistemas automatizados pode acarretar consequências clínicas, psicológicas e sociais não intencionais, muitas vezes imperceptíveis no início, mas com efeitos profundos em médio e longo prazos^{21,23}.

Entre os riscos mais discutidos estão os erros diagnósticos causados por mau funcionamento do sistema, interpretação inadequada ou uso fora de escopo. Mesmo que a IA possa oferecer precisão superior em determinados contextos, ela é tão falível quanto seus dados e parâmetros. Decisões clínicas automatizadas podem ignorar singularidades do paciente e levar a condutas inadequadas em situações complexas ou atípicas²⁶. Isso pode ser ainda mais problemático quando há

excesso de confiança dos profissionais na recomendação algorítmica, fenômeno conhecido como *automation bias* (viés de automação).

Outro risco ético relevante é a desumanização do cuidado, uma vez que, com o avanço da tecnologia, tende-se a valorizar o que é quantificável e previsível em detrimento da subjetividade²⁷. A substituição da escuta clínica pela interação com interfaces digitais pode reduzir o espaço para a expressão do sofrimento, do medo e da vulnerabilidade do paciente e resultar em desumanização do cuidado e negligência de aspectos subjetivos e afetivos essenciais à experiência do adoecer²⁷. Esse esvaziamento afetivo, embora difícil de medir, constitui um tipo de dano invisível – emocional e existencial – que contraria a integralidade do cuidado hipocrático.

Além disso, o uso de IA em saúde pode gerar efeitos coletivos adversos. Falhas de segurança, vazamento de dados ou decisões sistêmicas enviesadas minam a confiança da sociedade nos sistemas de saúde. A não maleficência, aqui, deve ir além da dimensão individual e considerar os impactos éticos sistêmicos e institucionais da automação.

A fim de mitigar esses riscos, é essencial incorporar medidas de robustez, rastreabilidade e auditoria contínua²³. Transparência, capacidade de revisão e a possibilidade de “*override humano*” (capacidade de intervir ou substituir a decisão tomada pela IA por um ser humano) são componentes éticos indispensáveis para que a IA médica opere sob o princípio da não maleficência²⁴. A tecnologia deve ser testada não apenas quanto a sua eficiência, mas também quanto a sua resiliência ética.

Justiça, equidade e o acesso digital à IA em saúde

O princípio da justiça, entendido como a obrigação de tratar todos com equidade e de promover acesso justo aos recursos em saúde, adquire novas camadas de complexidade na era da inteligência artificial. Se, por um lado, a IA promete ampliar o alcance do cuidado médico por meio de telemedicina, triagem remota e diagnósticos automatizados, por outro, esses avanços tecnológicos frequentemente esbarram em barreiras estruturais que aprofundam desigualdades já existentes^{21,28,29}.

A literatura evidencia que países de baixa e média renda, assim como regiões periféricas e comunidades vulneráveis, frequentemente carecem da infraestrutura mínima necessária para implementar soluções baseadas em IA: conectividade, interoperabilidade de sistemas, qualificação profissional, equipamentos e políticas públicas orientadas para inclusão digital^{29,30}. A ausência dessas condições compromete não apenas a eficácia das tecnologias, mas também a equidade na distribuição de seus benefícios.

A justiça social deve ser princípio norteador no desenvolvimento de tecnologias de IA, de modo a evitar a criação de uma medicina “em duas velocidades”: uma de alta *performance* e precisão, para quem pode pagar; outra, limitada e convencional, para quem depende do sistema público¹⁷. A mercantilização da IA na saúde pode ampliar o abismo entre pacientes de diferentes classes sociais e consolidar o que foi chamado de “injustiça automatizada”³¹.

Esse risco exige abordagem regulatória centrada em justiça distributiva. As diretrizes éticas globais sobre IA convergem para a necessidade de políticas que promovam acesso inclusivo, dados representativos e participação de populações sub-representadas no design e na validação dos sistemas⁵. No Brasil, por exemplo, a implementação ética da IA depende de investimentos em saúde digital pública, incentivo à interoperabilidade e capacitação profissional²⁸.

Portanto, a justiça na era da IA não é apenas questão de distribuição de recursos, mas também de inclusão epistemológica. Grupos historicamente marginalizados devem ser incorporados ao processo de desenvolvimento tecnológico, desde a produção de dados até a formulação de critérios de validação. Só assim será possível construir uma inteligência artificial verdadeiramente plural e comprometida com os valores hipocráticos da medicina universal.

Confidencialidade, LGPD e a nova ética da privacidade

Sigilo médico é um dos pilares mais antigos da ética profissional em saúde. Desde o Juramento de Hipócrates, preservar o segredo das informações obtidas durante o cuidado é considerado dever moral e jurídico. No entanto, a IA, por operar com

grandes volumes de dados sensíveis, impõe desafios inéditos a esse princípio^{6,21}.

Sistemas de IA dependem de dados clínicos extensivos para treinamento, validação e operação em tempo real. Muitas vezes, há informações altamente identificáveis, como imagens médicas, genomas, prontuários e metadados pessoais. Mesmo técnicas de anonimização podem ser ineficazes, já que algoritmos conseguem reidentificar indivíduos por cruzamento de dados^{15,30}. A coleta massiva desses dados, muitas vezes em segundo plano (via aplicativos, *wearables*, portais on-line), desafia a capacidade de obtenção de consentimento informado e a governança ética desses bancos de informação.

No Brasil, a LGPD⁸, que estabelece diretrizes para o tratamento de dados pessoais, exige base legal clara, minimização de dados, mecanismos de revogação, direito a portabilidade, eliminação e consentimento informado, especialmente em relação aos chamados “dados sensíveis”, como os de saúde.

A regulamentação da telessaúde no Brasil, por meio da Lei 14.510/2022⁷, reforça a obrigatoriedade de preservar a integridade, confidencialidade e disponibilidade das informações digitais em atendimentos remotos. Já a Resolução CFM 2.314/2022⁹ estabelece que o médico deve garantir, por meios tecnológicos adequados, a proteção dos dados do paciente e informá-lo sobre os riscos do atendimento virtual. Assim, essas normativas convergem para construir um ecossistema digital seguro, ético e centrado no paciente quando o serviço médico é prestado em território brasileiro.

No contexto internacional, destacam-se o HIPAA¹² e o HITECH¹³, nos Estados Unidos, que definem padrões de segurança, integridade e auditoria de dados clínicos digitais; e o GDPR¹⁰ e a NIS Directive¹¹, na Europa, que estabelecem o princípio de *privacy by design* (princípio de proteção de dados segundo o qual privacidade deve ser incorporada desde as fases iniciais do desenvolvimento de sistemas, produtos, serviços ou processos) e exigem *accountability* (dever de responsabilização e prestação de contas) dos controladores de dados e segurança cibernética em setores críticos, como a saúde.

Essas legislações indicam que a ética médica contemporânea deve incorporar uma bioética dos

dados: um conjunto de responsabilidades éticas e técnicas voltadas à proteção da privacidade e da autodeterminação informacional dos pacientes. Vivemos em uma era de “capitalismo de vigilância”, na qual dados pessoais são continuamente coletados e comercializados, de modo que a proteção da confidencialidade não é apenas questão técnica, mas também política e cultural³².

Responsabilidade moral e accountability em sistemas híbridos

Com a introdução da inteligência artificial na prática clínica, a noção tradicional de responsabilidade moral – centrada na figura do médico como decisor autônomo – passa por profunda transformação. Em ambientes onde decisões clínicas são compartilhadas entre humanos e algoritmos, a linha entre causalidade, responsabilidade e culpa torna-se nebulosa^{4,6}.

Essa ambiguidade é conhecida na literatura como “vácuo de responsabilidade”²³ e caracteriza-se pela dificuldade de atribuir responsabilidade em casos de erro médico assistido por IA. Por exemplo, se um algoritmo recomenda uma conduta baseada em evidência estatística e o médico, ao segui-la, causa dano ao paciente, quem responde?

A literatura propõe o conceito de “responsabilidade dividida”, que reconhece a multiplicidade de agentes envolvidos, desde os programadores e cientistas de dados até os fatores hospitalares e médicos¹⁷. Entretanto, o simples reconhecimento dessa complexidade não resolve o dilema ético, pois pode gerar efeito de diluição de culpa, enfraquecendo os mecanismos de *accountability* institucional³³.

Como resposta a esse desafio, autores defendem o desenvolvimento de normas éticas e legais que responsabilizem todas as partes envolvidas, com ênfase na auditabilidade dos sistemas e na rastreabilidade das decisões. Isso implica, por exemplo, que todo sistema de IA utilizado em saúde deve manter registros acessíveis dos dados que foram utilizados, dos modelos que foram aplicados e das recomendações que foram emitidas, dos contextos em que o foram e por quem⁵.

Na prática clínica, é essencial que o médico mantenha autonomia crítica diante dos sistemas automatizados e os utilize como ferramentas de apoio, e não como substitutos do julgamento clínico.

A IA deve ser uma “inteligência aumentada”², e não autoridade autônoma. Retomar a centralidade do médico como guardião do julgamento clínico, mesmo em contextos tecnológicos avançados, pode ser um caminho para reequilibrar o princípio hipocrático da prudência com a inovação. Nesse sentido, preservar a centralidade ética do médico no processo decisório é medida de proteção tanto à segurança do paciente quanto à integridade da profissão médica.

Crise da relação médico-paciente na era algorítmica

A medicina tradicional sempre foi construída sobre uma relação interpessoal complexa, marcada por escuta ativa, empatia, confiança e julgamento clínico humanizado. No entanto, a crescente mediação da prática médica por sistemas digitais – desde triagens automatizadas por *chatbots* (programas desenvolvidos para simular conversa com seres humanos, por meio de linguagem natural) até algoritmos de decisão clínica – ameaça reduzir a dimensão subjetiva e ética da interação entre médico e paciente^{15,21}.

Alguns autores apontam que, ao priorizar eficiência, padronização e objetividade, sistemas baseados em IA tendem a marginalizar aspectos humanísticos do cuidado, como compreensão do sofrimento, da dor ou das expectativas do paciente²⁷. O modelo hipocrático pressupõe um médico que conhece o paciente em sua integralidade – biológica, emocional e social – e que, portanto, é capaz de adaptar condutas de maneira ética e prudente. Todavia, o uso excessivo de tecnologias pode levar à chamada “despersonalização algorítmica” do cuidado, na qual a subjetividade do paciente é reduzida a variáveis numéricas ou probabilísticas, sem espaço para narrativas pessoais ou singularidades²⁸. Além disso, a confiança, que é componente essencial da aliança terapêutica, pode ser transferida para a tecnologia – fenômeno conhecido como *trust displacement* –, o que fragiliza o vínculo com o profissional, prejudica a adesão ao tratamento e cria expectativas infundadas de infalibilidade técnica³⁴.

Adicionalmente, o uso de IA em etapas iniciais do cuidado, como triagem automatizada por *chatbots*, pode reduzir o contato humano em momentos decisivos para o paciente, o que gera

insegurança, solidão ou sensação de abandono. A tecnologia tende a desvalorizar a escuta e a intuição clínica – capacidades que, embora não quantificáveis, são essenciais para reconhecer o sofrimento humano em sua complexidade²⁷.

Esse cenário desafia a própria identidade do médico. Se antes ele era visto como cuidador e conselheiro, hoje corre o risco de ser reduzido a operador técnico de sistemas, perdendo espaço para o raciocínio clínico e para a aética do encontro. É importante resgatar a noção de que a medicina deve ser, antes de tudo, compromisso com o bem do outro³ – algo que não pode ser delegado a máquinas, por mais eficientes que sejam.

Para enfrentar essa crise, torna-se urgente promover uma formação médica voltada à ética digital, ao pensamento crítico sobre IA e ao desenvolvimento de competências comunicacionais para interações mediadas por tecnologia¹⁶. Restaurar a centralidade da relação humana no cuidado é imperativo ético que nenhuma inovação pode substituir.

Reinterpretação dos princípios hipocráticos no século XXI

A análise dos estudos revisados revela que os princípios éticos enunciados por Hipócrates continuam sendo uma base moral valiosa para a medicina contemporânea, mas necessitam de reinterpretação diante dos desafios da inteligência artificial. Essa reinterpretação pode ser sintetizada em quatro movimentos:

1) Autonomia 2.0: a autonomia do paciente, no contexto da IA, não se limita à escolha de tratamentos e inclui o direito a explicação, não discriminação algorítmica e proteção de dados sensíveis. Surge a necessidade de consentimento informado ampliado, capaz de abordar riscos tecnológicos e o funcionamento de sistemas automatizados²⁴.

2) Beneficência inclusiva: fazer o bem requer assegurar que os benefícios das tecnologias sejam distribuídos de forma equitativa. Isso implica monitorar vieses de dados, priorizar desenvolvimento ético de algoritmos e evitar que inovações aumentem desigualdades sociais¹⁴.

3) Não maleficência sistêmica: prevenir danos não significa apenas evitar erros clínicos, mas também antecipar consequências sistêmicas da automação, como exclusão digital, perda de

competências médicas e descaracterização do cuidado baseado nos princípios hipocráticos²³.

4) Justiça digital: a justiça na era da IA exige que todos os envolvidos tenham acesso a tecnologias seguras, eficazes e eticamente projetadas. A governança algorítmica deve considerar o pluralismo cultural, a diversidade de contextos e o respeito aos direitos fundamentais dos pacientes¹⁷.

Assim, o legado hipocrático não é descartado, mas atualizado, funcionando como matriz ética viva que pode orientar decisões morais em ambientes tecnocientíficos complexos.

Propostas bioéticas para uma IA centrada no paciente

Com base na literatura revisada, é possível delinear algumas diretrizes bioéticas para o desenvolvimento e a aplicação da IA na medicina:

1) Transparência e explicabilidade algorítmica: os sistemas devem ser auditáveis e compreensíveis, para permitir ao médico e ao paciente entenderem os critérios das decisões sugeridas^{4,24}.

2) Consentimento informado dinâmico: mecanismos de consentimento contínuo e contextualizado devem ser desenvolvidos de modo que se adaptem à natureza interativa das tecnologias digitais³⁵.

3) Participação inclusiva: o desenvolvimento de IA deve incluir múltiplos atores, com especial atenção às vozes dos usuários finais – médicos e pacientes –, para assegurar alinhamento com valores clínicos^{25,35}.

4) Governança ética: a implementação da IA deve ser acompanhada por comitês éticos multidisciplinares que revisem seus impactos em tempo real, a fim de promover ajustes de responsabilização^{18,36}.

5) Educação médica para a era digital: os currículos de medicina devem incluir formação em ética da tecnologia, literacia de dados e pensamento crítico sobre IA, para evitar dependência acrítica ou delegação irresponsável¹⁶.

Essas propostas, embora ainda em construção, apontam caminhos para uma medicina algorítmica mais ética, transparente e centrada na dignidade humana.

Limitações da literatura atual e lacunas para pesquisa futura

Apesar da crescente produção científica sobre ética da IA na medicina, a literatura apresenta

lacunas importantes. Em primeiro lugar, há predomínio de estudos teóricos e normativos, com escassez de dados empíricos sobre a experiência de médicos e pacientes em contextos reais de uso de IA. A ausência de estudos qualitativos limita a compreensão de como os dilemas éticos se manifestam na prática cotidiana²¹.

Em segundo lugar, muitas pesquisas têm enfoque euro-americano e, portanto, desconsideram desigualdades digitais e especificidades culturais, legais e sociais de outros contextos, como de países da América Latina e da África. Estudos sobre justiça algorítmica precisam incorporar perspectivas pós-coloniais e de saúde global, sob risco de repetir padrões tecnocêntricos e excludentes³⁷ e reproduzir exclusões históricas com roupagem tecnológica.

Além disso, observa-se falta de integração entre marcos legais, clínicos e éticos. Poucos modelos oferecem orientações práticas que articulem em um só *framework* as exigências da LGPD, da ética médica e da eficiência clínica. Nesse sentido, há demanda urgente por pesquisas interdisciplinares que dialoguem com direito, sociologia, ciência da computação e fisiologia moral, e o engajamento com pacientes e a ampliação dos fóruns de deliberação ética podem ser caminhos promissores, sem perder a inovação.

O avanço ético da IA na medicina exige não apenas inovação técnica, mas também comprometimento público e institucional com dignidade humana. Preservar o espírito do Juramento de Hipócrates ao adaptá-lo aos tempos digitais é o maior desafio – e a maior oportunidade – da bioética contemporânea.

Considerações finais

A inteligência artificial representa uma das transformações mais significativas da prática médica no século XXI. Ao mesmo tempo que oferece recursos poderosos para diagnósticos, prognósticos e terapias personalizadas, sua adoção em larga escala impõe desafios éticos profundos que reconfiguram os alicerces tradicionais da medicina. Tais transformações demandam não apenas atualização técnica dos profissionais de saúde, mas também reflexão crítica sobre os valores que devem nortear o uso dessas tecnologias.

A análise crítica da literatura revela que os princípios hipocráticos da beneficência, não maleficência, autonomia e justiça continuam sendo fundamentais, mas precisam ser reinterpretados à luz dos riscos e potencialidades da era digital. Autonomia, por exemplo, já não pode ser garantida apenas pelo consentimento informado tradicional: ela requer explicabilidade algorítmica e transparência das decisões automatizadas. Beneficência demanda vigilância constante contra vieses nos dados e nos sistemas de aprendizagem da máquina, que podem comprometer a qualidade e a equidade do cuidado. Não maleficência exige abordagem sistêmica e interdisciplinar, que leve em conta os efeitos não intencionais dos algoritmos na prática clínica. E justiça implica compromisso com a democratização do acesso a tecnologias, a fim de evitar a concentração de recursos em contextos privilegiados.

Adicionalmente, os marcos legais nacionais e internacionais oferecem importante base normativa, mas ainda insuficiente para dar conta da complexidade ética que os sistemas de IA em saúde introduzem. Embora forneçam diretrizes relevantes para proteção de dados e segurança da informação, muitos desses regulamentos ainda não abrangem integralmente os dilemas ligados à autonomia algorítmica, à responsabilidade em

sistemas híbridos (homem-máquina) e à transparência em modelos de aprendizado profundo, muitas vezes considerados “caixas-pretas”.

Nesse contexto, a bioética contemporânea precisa evoluir para incorporar, de forma verdadeiramente interdisciplinar, os conceitos de governança algorítmica, proteção de dados sensíveis, justiça distributiva e responsabilização moral. Isso envolve a participação ativa de profissionais da saúde, juristas, engenheiros, filósofos e representantes da sociedade civil na formulação de políticas públicas e diretrizes clínicas. O desenvolvimento ético da IA em saúde não pode ser responsabilidade exclusiva dos programadores ou das empresas de tecnologia: deve ser um esforço conjunto, orientado pelo bem-estar humano.

A preservação do espírito do Juramento de Hipócrates não significa resistir ao progresso tecnológico, mas assegurar que a inovação seja colocada a serviço da dignidade humana, da equidade e do cuidado compassivo. Cabe à medicina, à ética e ao direito construir juntos um modelo de inteligência artificial que, mais do que aprender com dados, respeite valores fundamentais da vida e do cuidado. Afinal, o verdadeiro avanço da medicina digital dependerá da capacidade não apenas de prever doenças, mas também de promover humanidade em cada decisão automatizada.


Referências

1. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* [Internet]. 2017 [acesso 5 jun 2025];542(7639):115-8. DOI: 10.1038/nature21056
2. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* [Internet]. 2019 [acesso 5 jun 2025];25(1):44-56. DOI: 10.1038/s41591-018-0300-7
3. Pellegrino ED, Thomasma DC. *For the patient's good: the restoration in health care*. New York: Oxford University Press; 1998.
4. London AJ. Artificial intelligence and black-box medical decisions: accuracy versus explainability. *Hastings Cent Rep* [Internet]. 2019 [acesso 5 jun 2025];49(1):15-21. DOI: 10.1002/hast.973
5. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. *Nat Mach Intell* [Internet]. 2019 [acesso 5 jun 2025];1(9):389-99. DOI: 10.1038/s42256-019-0088-2
6. Gerke S, Minssen T, Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. *Artificial Intelligence in Healthcare*. 2020:295-336. DOI: 10.1016/B978-0-12-818438-7.00012-5
7. Brasil. Lei nº 14.510, de 27 de dezembro de 2022. Dispõe sobre a prestação de serviços de telessaúde. *Diário Oficial da União* [Internet]. Brasília, 28 dez 2022 [acesso 6 jun 2025]. Disponível: <https://tinyurl.com/mr39srpb>

8. Brasil. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Diário Oficial da União [Internet]. Brasília, 15 ago 2018 [acesso 6 jun 2025]. Disponível: <https://bit.ly/4b4cssd>
9. Conselho Federal de Medicina. Resolução nº 2.314, de 20 de abril de 2022. Define e regulamenta a telemedicina como forma de prestação de serviços médicos. Diário Oficial da União [Internet]. Brasília, p. 227, 5 maio 2022 [acesso 6 jun 2025]. Seção 1. Disponível: <https://tinyurl.com/hyrzwr3f>
10. European Parliament and Council. Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Official Journal of the European Union [Internet]. Geneva, 2016 [acesso 5 jun 2025]. Disponível: <https://bit.ly/409B7qj>
11. European Commission. Directive (EU) 2016/1148 on Security of Network and INformation Systems (NIS Directive). Official Journal of the European Union [Internet]. 2016 [acesso 6 jun 2025]. Disponível: <https://bit.ly/4aQwtUa>
12. United States. Public Law 104-191. Health Insurance Portability and Accountability Act of 1996 (HIPAA). Federal Register [Internet]. Washington, 20 ago 1996 [acesso 6 jun 2025]. Disponível: <https://tinyurl.com/vvkr9vm8>
13. United States. Health Information Technology for Economic and Clinical Health Act (HITECH). Federal Register [Internet]. Washington, nº 209, Rules and Regulations, 30 out 2009 [acesso 10 set 2024]. Disponível: <https://bit.ly/4d4yRYR>
14. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* [Internet]. 2019 [acesso 6 jun 2025];366(6464):447-53. DOI: 10.1126/science.aax2342
15. Lupton D. The digitally engaged patient: self-monitoring and self-care in the digital health era. *Soc Theory Health* [Internet]. 2013 [acesso 8 jun 2025];11(3):256-70. DOI: 10.1057/sth.2013.10
16. Masters K. Artificial intelligence in medical education. *Med Teach* [Internet]. 2019 [acesso 14 jun 2025];41(9):976-80. DOI: 10.1080/0142159X.2019.1595557
17. Floridi L, Cowls J, Beltrametti M, Chatila R, Chazerand P, Dignum V *et al*. AI4People – an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. *Minds Mach* [Internet]. 2018 [acesso 6 jun 2025];28(4):689-707. DOI: 10.1007/s11023-018-9482-5
18. Mittelstadt BD. Principles alone cannot guarantee ethical AI. *Nat Mach Intell* [Internet]. 2019 [acesso 6 jun 2025];1(11):501-7. DOI: 10.1038/s42256-019-0114-4
19. Gil AC. Métodos e técnicas de pesquisa social. 6ª ed. São Paulo: Atlas; 2008.
20. Severino AJ. Metodologia do trabalho científico. 24ª ed. São Paulo: Cortez; 2016.
21. Morley J, Machado C, Burr C, Cowls J, Taddeo M, Floridi L. The ethics of AI in health care: a mapping review. *Soc Sci Med* [Internet]. 2020 [acesso 5 jun 2025];260:113172. DOI: 10.1016/j.socscimed.2020.113172
22. Krittanawong C, Johnson KW, Rosenson RS, Wang Z, Aydar M, Narayan SM. Deep learning for cardiovascular medicine: a practical primer. *Eur Heart J* [Internet]. 2019 [acesso 6 jun 2025];40(25):2058-73. DOI: 10.1093/eurheartj/ehz056
23. Mittelstadt BD, Floridi L. The ethics of big data: current and foreseeable issues in biomedical contexts. *Sci Eng Ethics* [Internet]. 2016 [acesso 6 jun 2025];22(2):303-41. DOI: 10.1007/s11948-015-9652-2
24. Wachter S, Mittelstadt B, Floridi L. Transparent, explainable, and accountable AI for robotics. *Sci Robot* [Internet]. 2017 [acesso 8 jun 2025];2(6):eaan6080. DOI: 10.1126/scirobotics.aan6080
25. Dastin J. Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters* [Internet]. 10 out 2018 [acesso 2 jul 2025]; Disponível: <https://bit.ly/4rhewTv>
26. Shortliffe EH, Sepúlveda MJ. Clinical decision support in the era of artificial intelligence. *JAMA* [Internet]. 2018 [acesso 6 jun 2025];320(21):2199-200. DOI: 10.1001/jama.2018.17163
27. Harari YN. *Homo Deus: a brief history of tomorrow*. New York: Harper; 2017.
28. Silva AG, Oliveira TC, Lima MADS. Artificial intelligence in Brazilian health care: a viewpoint. *JMIR Med Inform* [Internet]. 2020 [acesso 7 jun 2025];8(10):e23557. DOI: 10.2196/23557

29. Whitelaw S, Mamas MA, Topol E, Van Spall HGC. Applications of digital technology in COVID-19 pandemic planning and response. *Lancet Digit Health* [Internet]. 2020 [acesso 7 jun 2025];2(8):e435-40. DOI: 10.1016/S2589-7500(20)30142-4
30. Goodman B. The algorithmic governance of health: a case study of the United States. *Health Care Anal* [Internet]. 2021 [acesso 8 jun 2025];29(1):39-57. DOI: 10.1007/s10728-019-00385-0
31. Eubanks V. Automating inequality: how high-tech tools profile, police, and punish the poor. New York: Picador; 2019.
32. Zuboff S. The age of surveillance capitalism. New York: PublicAffairs; 2019.
33. Nemitz P. Constitutional democracy and technology in the age of artificial intelligence. *Philos Trans A Math Phys Eng Sci* [Internet]. 2018 [acesso 7 jun 2025];376(2133):20180089. DOI: 10.1098/rsta.2018.0089
34. Zhang B, Dafoe A. Artificial intelligence: American attitudes and trends [Internet]. Oxford: University of Oxford; 2019 [acesso 8 jun 2025]. DOI: 10.2139/ssrn.3312874
35. Sadan T. Toward participatory design of ethical AI. *Philos Technol* [Internet]. 2020 [acesso 8 jun 2025];33(2):167-88. DOI: 10.1007/s13347-019-00371-1
36. Dove ES, Townend D, Meslin EM, Bobrow M, Littler K, Nicol D *et al.* Ethics review for international data-intensive research. *Science* [Internet]. 2016 [acesso 8 jun 2025];351(6280):1399-400. DOI: 10.1126/science.aad5269
37. Benjamin R. Race after technology: abolitionist tools for the New Jim Code [Internet]. Cambridge: Polity Press; 2019 [acesso 8 jun 2025]. Disponível: <https://tinyurl.com/272s7b8z>

Juracy Barbosa dos Santos – Doutorando – drjuracybarbosa@gmail.com

 0009-0009-1729-9778

Guilhermina Rego – Doutora – guilherminarego@med.up.pt

 0000-0002-8590-9832

Rui Nunes – Doutor – ruinunes@med.up.pt

 0000-0002-1377-9899

Correspondência

Juracy Barbosa dos Santos – Faculdade de Medicina/Universidade do Porto. Al. Professor Hernani Monteiro, CP 4200-319. Porto, Portugal.

Contribuição dos autores

Juracy Barbosa dos Santos foi responsável pela concepção do estudo, análise de dados e redação do manuscrito. Guilhermina Rego revisou o manuscrito. Rui Nunes realizou a análise crítica e a revisão final do manuscrito.

Disponibilidade de dados: Todos os dados utilizados ou gerados na pesquisa estão integralmente descritos e apresentados no corpo do artigo.

Editora responsável: Dilza Teresinha Ambrós Ribeiro

Recebido: 24.8.2025

Revisado: 18.12.2025

Aprovado: 27.1.2026