A NOVA FRONTEIRA DA BIOÉTICA: DESAFIOS E IMPERATIVOS NA ERA DA INTELIGÊNCIA ARTIFICIAL GENERATIVA

Paulo Henrique Matayoshi Calixto[1]

[1] Instituto Federal Goiano

Resumo

A ascensão da Inteligência Artificial (IA) generativa representa uma das mais significativas transformações tecnológicas do século XXI, com um impacto profundo e iminente na medicina, na pesquisa biomédica e na saúde pública. Modelos de linguagem avançados, redes generativas adversariais e outras arquiteturas de IA são capazes de criar conteúdo novo e complexo, desde textos diagnósticos e planos de tratamento personalizados até o design de novas moléculas e a análise preditiva de dados genômicos. Contudo, essa capacidade disruptiva introduz um espectro de dilemas éticos sem precedentes, que desafiam os pilares tradicionais da bioética. Esta análise explora a intersecção entre a bioética e a IA generativa, examinando as implicações desta tecnologia através dos princípios da autonomia, beneficência, não maleficência e justiça. Discute-se a complexidade do consentimento informado quando os dados do paciente são utilizados para treinar algoritmos, a questão da responsabilidade em caso de erro diagnóstico gerado por IA, o risco de exacerbação de desigualdades em saúde devido a vieses algorítmicos e o acesso desigual à tecnologia. Além disso, o artigo aborda questões emergentes como a privacidade de dados de saúde em larga escala, a integridade da pesquisa científica e a própria natureza da tomada de decisão clínica em uma parceria homem-máquina. Conclui-se que a integração ética da IA generativa na saúde exige o desenvolvimento de novos quadros regulatórios, uma governança de dados robusta e uma formação contínua dos profissionais de saúde, a fim de assegurar que a inovação tecnológica sirva ao bem-estar humano, respeitando a dignidade e os direitos fundamentais dos pacientes.

Palavras-chave: Inteligência Artificial Generativa; Bioética; Ética Médica; Governança de Dados; Vieses Algorítmicos; Saúde Digital.

THE NEW FRONTIER OF BIOETHICS: CHALLENGES AND IMPERATIVES IN THE ERA OF GENERATIVE ARTIFICIAL INTELLIGENCE

Abstract

The rise of Generative Artificial Intelligence (AI) represents one of the most significant technological transformations of the 21st century, with a profound and imminent impact on medicine, biomedical research, and public health. Advanced language models, generative adversarial networks, and other AI architectures are capable of creating novel and complex content, ranging from diagnostic reports and personalized treatment plans to the design of new molecules and the predictive analysis of genomic data. However, this disruptive capability introduces an unprecedented spectrum of ethical dilemmas that challenge the traditional pillars of bioethics. This analysis explores the intersection between bioethics and generative AI,

examining the implications of this technology through the principles of autonomy, beneficence, non-maleficence, and justice. The complexity of informed consent is discussed, particularly when patient data are used to train algorithms, as well as the issue of accountability in cases of AI-generated diagnostic errors, the risk of exacerbating health disparities due to algorithmic bias, and unequal access to technology. In addition, the article addresses emerging concerns such as large-scale health data privacy, the integrity of scientific research, and the very nature of clinical decision-making within a human–machine partnership. It concludes that the ethical integration of generative AI in healthcare requires the development of new regulatory frameworks, robust data governance, and the continuous education of healthcare professionals, in order to ensure that technological innovation serves human well-being while respecting patients' dignity and fundamental rights.

Keywords: Generative Artificial Intelligence; Bioethics; Medical Ethics; Data Governance; Algorithmic Bias; Digital Health.

1. Introdução

A Inteligência Artificial (IA) generativa, definida como o subcampo da IA focado na criação de conteúdo original a partir de dados existentes, emergiu como uma força transformadora em múltiplos setores da sociedade (Goodfellow *et al.,* 2016). Na área da saúde, o seu potencial é particularmente vasto, prometendo revolucionar desde o diagnóstico por imagem e a descoberta de fármacos até a personalização do tratamento e a gestão de sistemas de saúde (Sallam, 2023). Ferramentas capazes de redigir relatórios médicos, gerar hipóteses de pesquisa e até mesmo desenhar novas estruturas proteicas estão deixando de ser ficção científica para se tornarem realidade nos laboratórios e hospitais. A velocidade acelerada desta evolução tecnológica, no entanto, ultrapassa a capacidade dos quadros éticos e regulatórios existentes de acompanhar e dar resposta aos desafios que ela suscita (Cohen; Price, 2023). A aplicação da IA generativa em contextos biomédicos levanta questões fundamentais que tocam no cerne da bioética: a dignidade do paciente, a relação médico-paciente, a equidade no acesso aos cuidados de saúde e a própria definição de responsabilidade profissional (Floridi, 2019). A promessa de avanços sem precedentes caminha lado a lado com riscos significativos, que exigem uma análise criteriosa e proativa.

Este artigo propõe-se a realizar uma análise aprofundada dos principais dilemas bioéticos impostos pela IA generativa. Para tal, a discussão será estruturada em torno dos quatro princípios fundamentais da bioética — autonomia, beneficência, não maleficência e justiça —, conforme estabelecido por Beauchamp e Childress (2019). Será demonstrado como cada um destes princípios é reinterpretado e desafiado pela introdução de agentes não humanos com capacidades generativas no ecossistema da saúde. O objetivo não é oferecer respostas definitivas, mas sim mapear o terreno ético, identificar os pontos de maior tensão e propor caminhos para o desenvolvimento de uma governança responsável (WHO, 2021). A integração da IA generativa na medicina não é apenas uma questão de implementação técnica; é, fundamentalmente, uma questão de valores. A forma como a sociedade navegar nestas

águas determinará se esta poderosa tecnologia será uma força para o bem comum ou um catalisador de novas formas de desigualdade e desumanização no cuidado (Char; Shah; Magnus, 2018).

2. A Natureza da Inteligência Artificial Generativa e as Suas Aplicações na Saúde

Para compreender os desafios bioéticos, é imperativo primeiro delinear o que distingue a IA generativa de outras formas de inteligência artificial. Ao contrário da IA analítica, que se concentra em classificar e prever com base em padrões de dados, a IA generativa cria novos artefatos digitais que mimetizam as características dos dados com os quais foi treinada (Gao *et al.,* 2023). Modelos como as Redes Generativas Adversariais (GANs) e os Transformadores Generativos Pré-treinados (GPTs) são exemplos proeminentes desta tecnologia, capazes de produzir imagens, textos e dados sintéticos com um nível de realismo surpreendente (Vaswani *et al.,* 2017).

As aplicações biomédicas desta tecnologia são vastas e em rápida expansão. Na radiologia, por exemplo, as GANs podem ser usadas para gerar imagens médicas sintéticas para aumentar os conjuntos de dados de treinamento, melhorando a precisão dos algoritmos de diagnóstico de doenças raras (Yi; Walia; Babyn, 2019). Na descoberta de fármacos, modelos generativos podem projetar novas moléculas com propriedades farmacológicas específicas, acelerando drasticamente o processo de pesquisa e desenvolvimento de novos medicamentos, um processo tradicionalmente longo e dispendioso (Zhavoronkov *et al.,* 2019). No campo da medicina personalizada, os grandes modelos de linguagem (LLMs - Large Language Models) podem analisar o histórico de saúde de um paciente, dados genômicos e a literatura médica mais recente para gerar recomendações de tratamento altamente individualizadas (Kortz; Lee, 2023). Além disso, podem funcionar como assistentes virtuais para médicos, automatizando a documentação clínica, resumindo informações do paciente e redigindo comunicações, liberando tempo valioso para a interação direta com o paciente (Davenport; Kalakota, 2019).

Esta capacidade de síntese e criação de informação nova representa uma mudança de paradigma. A IA deixa de ser apenas uma ferramenta de análise para se tornar um parceiro "criativo" na prática clínica e na pesquisa (Wang *et al.,* 2023). É precisamente esta agência generativa que introduz uma nova camada de complexidade ética, pois as decisões e os resultados já não podem ser atribuídos apenas a um operador humano ou a um algoritmo determinístico (Nyholm, 2018).

3. Os Pilares da Bioética Revisitados pela IA Generativa

A estrutura principialista da bioética, popularizada por Beauchamp e Childress (2019), oferece um roteiro robusto para analisar dilemas morais na medicina. Os seus quatro princípios, autonomia, beneficência, não maleficência e justiça, não são regras absolutas, mas sim guias para a deliberação moral que devem ser ponderados e equilibrados. A introdução da IA generativa no cenário clínico e de pesquisa obriga a uma reavaliação profunda do significado e da aplicação de cada um destes pilares. A tecnologia não opera num vácuo ético; ela reconfigura as relações de poder, as fontes

de conhecimento e as modalidades de cuidado (Sieck *et al.,* 2021). A IA generativa, com a sua capacidade de operar com um grau de autonomia e opacidade (a chamada "caixa-preta"), desafia diretamente as noções tradicionais de agência e responsabilidade que são centrais para a ética médica (Wachter, 2017). A seguir, cada princípio será dissecado em detalhe, explorando as novas questões que emergem com a implementação desta tecnologia.

4. O Princípio da Autonomia na Era Algorítmica

A autonomia, ou o direito do paciente ao autogoverno e à tomada de decisões informadas sobre os seus próprios cuidados, é uma pedra angular da ética médica moderna (Faden; Beauchamp, 1986). A IA generativa interfere neste princípio de, pelo menos, duas formas cruciais: através do uso de dados do paciente para o treinamento de modelos e através da sua influência na tomada de decisão clínica partilhada (Price; Cohen, 2019).

O consentimento informado para o uso de dados de saúde assume uma nova dimensão. Tradicionalmente, o consentimento é obtido para um propósito específico, como um procedimento ou a participação numa pesquisa (Nunes, 2020). No entanto, os modelos de IA generativa são treinados em vastos conjuntos de dados, muitas vezes agregados de milhões de registros de pacientes, tornando o consentimento específico e granular praticamente impossível de obter (Shay; Cohen, 2023). Surge a questão de saber se um consentimento amplo, dado no momento da admissão hospitalar, é eticamente suficiente para permitir que dados altamente sensíveis sejam usados para treinar um sistema de IA que servirá a propósitos futuros e imprevisíveis.

A privacidade é um componente essencial da autonomia. A capacidade dos modelos generativos de criar dados sintéticos de pacientes levanta um novo tipo de risco: a reidentificação (El Emam *et al.,* 2020). Embora os dados sintéticos sejam criados para preservar o anonimato, estudos têm demonstrado que, com técnicas sofisticadas, é possível correlacionar esses dados com indivíduos reais, especialmente se contiverem informações genômicas ou demográficas raras (Rocher *et al.,* 2019). A proteção da confidencialidade do paciente exige, portanto, salvaguardas técnicas e legais muito mais robustas do que as atualmente existentes.

Por outro lado, a autonomia do paciente no processo de decisão clínica pode ser tanto ampliada quanto diminuída pela IA generativa. Um sistema de IA pode processar uma quantidade de informação muito superior à de um médico humano, apresentando ao paciente opções de tratamento, probabilidades de sucesso e riscos de uma forma mais completa e personalizada (Topol, 2019). Esta riqueza de informação pode, teoricamente, capacitar o paciente a tomar uma decisão mais informada, fortalecendo a sua autonomia (Gille *et al.,* 2020). Contudo, existe o risco de uma "automação da deferência", onde tanto o médico quanto o paciente podem sentir-se compelidos a seguir a recomendação da IA, percebida como objetiva e infalível, sem uma deliberação crítica adequada (Parikh; Obermeyer, 2021). A complexidade e a natureza de "caixa-preta" de muitos algoritmos tornam difícil explicar o "porquê" de uma determinada

recomendação, minando a base para um diálogo verdadeiramente informado entre médico e paciente. A autonomia, neste cenário, corre o risco de se tornar uma mera formalidade de assentimento a uma decisão algorítmica (Zou; Schiebinger, 2018).

5. O Princípio da Beneficência e a Promessa da Inovação

O princípio da beneficência obriga os profissionais de saúde a agir no melhor interesse do paciente, promovendo o seu bem-estar e procurando ativamente benefícios para a sua saúde (Beauchamp; Childress, 2019). A IA generativa oferece caminhos promissores para cumprir este imperativo ético, através da aceleração da inovação e da melhoria da qualidade dos cuidados (Yu; Beam; Kohane, 2018). Como mencionado, a capacidade de projetar novas moléculas candidatas a fármacos pode reduzir drasticamente o tempo e o custo do desenvolvimento de novos tratamentos para doenças como o câncer ou doenças raras (Fleming, 2018). Da mesma forma, a utilização de IA para analisar dados genômicos e identificar novos alvos terapêuticos representa um benefício potencial incalculável para a humanidade (Zhavoronkov *et al.,* 2019). A beneficência, neste contexto, implica um dever de explorar e desenvolver responsavelmente estas tecnologias.

Na prática clínica diária, a beneficência pode ser promovida através da redução de erros médicos. Sistemas de IA generativa podem atuar como uma "segunda opinião" inteligente, analisando exames de imagem ou dados laboratoriais para detectar anomalias que poderiam passar despercebidas ao olho humano (Singhal *et al.,* 2023). Ao automatizar tarefas administrativas, a IA também pode reduzir o esgotamento (*burnout*) dos médicos, permitindo-lhes dedicar mais tempo e atenção ao cuidado direto do paciente, o que é um benefício claro para a qualidade da relação terapêutica (Meskó; Dhole, 2021).

Contudo, o dever de beneficência não é absoluto e deve ser ponderado com os riscos associados. A implementação apressada de tecnologias de IA não validadas adequadamente pode causar mais mal do que bem (Keane; Topol, 2018). Existe uma obrigação ética de garantir que os sistemas de IA generativa passem por testes clínicos rigorosos, semelhantes aos exigidos para novos fármacos ou dispositivos médicos, antes de serem implementados em larga escala. A validação deve ocorrer em populações diversas para garantir que os benefícios se apliquam a todos os grupos de pacientes (Nagendran *et al.,* 2020).

6. O Princípio da Não Maleficência: Vieses, Erros e o Risco de Dano

O princípio da não maleficência, resumido no aforismo hipocrático *primum non nocere* ("primeiro, não causar dano"), exige que se evite ativamente causar dano a um paciente (Beauchamp; Childress, 2019). Com a IA generativa, os potenciais para dano são múltiplos e, por vezes, subtis, sendo o viés algorítmico um dos mais preocupantes (Obermeyer *et al.,* 2019). Os modelos de IA são treinados com dados do mundo real, e se esses dados refletem preconceitos e desigualdades sociais existentes, o algoritmo irá aprender e perpetuar esses vieses (Cirillo *et al.,* 2020). Por exemplo, se um algoritmo de diagnóstico é treinado predominantemente com dados de pacientes de uma determinada etnia, a sua precisão pode ser

significativamente menor para pacientes de outras etnias, levando a diagnósticos tardios ou incorretos e, consequentemente, a danos (Rajkomar *et al.,* 2018). O dano, neste caso, não é intencional, mas é uma consequência sistémica de práticas de recolha de dados não equitativas.

Outra fonte de dano potencial reside nos erros gerados pela IA, conhecidos como "alucinações". Os modelos de linguagem, por vezes, podem gerar informações que são factualmente incorretas, mas apresentadas de forma eloquente e convincente (Ji *et al.,* 2023). Se um médico confiar num resumo clínico gerado por IA que contém uma alucinação — como a inclusão de um medicamento ao qual o paciente é alérgico — as consequências podem ser fatais (Rao *et al.,* 2023). A verificação humana de todos os resultados gerados por IA é essencial, mas isto levanta questões sobre a eficiência e a carga de trabalho.

A questão da responsabilidade em caso de erro é um dos maiores desafios legais e éticos. Se um paciente é prejudicado por uma recomendação de uma IA, quem é o responsável: o médico que a utilizou, o hospital que a implementou, a empresa que a desenvolveu, ou o próprio algoritmo? (Price; Cohen, 2019). A falta de clareza sobre a responsabilidade pode deixar os pacientes sem recurso em caso de dano e pode criar uma cultura de "passagem de culpa" que mina a confiança no sistema de saúde (Matthias, 2004).

A não maleficência também se estende à segurança dos dados. A centralização de enormes volumes de dados de saúde para treinar modelos de IA cria um alvo atrativo para ataques cibernéticos (Finlayson *et al.,* 2019). Uma violação de dados em larga escala pode levar não apenas à perda de privacidade, mas também à manipulação de informações de saúde, com potencial para danos físicos e psicológicos generalizados.

7. O Princípio da Justiça e a Ameaça da Brecha Digital

O princípio da justiça, no contexto da bioética, refere-se à distribuição equitativa de benefícios, riscos e custos (Rawls, 1971). A IA generativa tem o potencial tanto para reduzir quanto para exacerbar as desigualdades existentes na saúde (Reddy *et al.,* 2020). O desafio ético é garantir que os benefícios desta tecnologia sejam partilhados por toda a sociedade e não apenas por uma minoria privilegiada.

Por um lado, a IA poderia promover a justiça ao democratizar o acesso a conhecimentos especializados. Um sistema de IA de diagnóstico poderia fornecer apoio a médicos em áreas rurais ou em países com poucos recursos, onde o acesso a especialistas é limitado (Wahl *et al.,* 2018). A automação de tarefas também poderia reduzir os custos operacionais dos sistemas de saúde, tornando os cuidados mais acessíveis (Sallam, 2023).

No entanto, os riscos de aprofundar as desigualdades são imensos. O desenvolvimento e a implementação de sistemas de IA sofisticados são extremamente caros, o que pode levar a uma "brecha da IA" na saúde, onde hospitais e sistemas de saúde ricos têm acesso a ferramentas de ponta, enquanto os mais pobres ficam para trás (Ghassemi *et al.,* 2021). Isso poderia criar um sistema de saúde de dois níveis, onde a qualidade do cuidado depende da capacidade de pagar pela tecnologia mais avançada.

O problema do viés algorítmico é também uma questão central de justiça. Se

os algoritmos são menos precisos para grupos populacionais sub-representados, eles irão sistematicamente fornecer um padrão de cuidado inferior a esses grupos, reforçando as disparidades de saúde já existentes (Obermeyer *et al.,* 2019). A justiça exige que sejam feitos esforços proativos para garantir que os conjuntos de dados de treinamento sejam diversificados e representativos da população que a IA se destina a servir (Zou; Schiebinger, 2018).

Além disso, a justiça distributiva questiona quem lucra com a IA na saúde. Os dados dos pacientes são o recurso fundamental que alimenta estes sistemas. Se empresas de tecnologia privadas utilizam dados de saúde pública para desenvolver algoritmos lucrativos, como devem esses benefícios ser partilhados com os pacientes e com a sociedade que forneceram os dados? (Vaidhyanathan, 2018). Modelos de governança de dados que tratam os dados de saúde como um bem comum e garantem uma partilha justa de benefícios são urgentemente necessários (Blasimme; Vayena, 2019).

8. Desafios Éticos Específicos e Questões Emergentes

Para além da aplicação dos quatro princípios clássicos, a IA generativa levanta um conjunto de questões éticas novas e específicas que merecem uma análise dedicada. Estas incluem a natureza da autoria na pesquisa científica, o impacto na relação médico-paciente e o estatuto moral dos próprios sistemas de IA.

8.1. Integridade e Autoria na Pesquisa Científica

A capacidade dos modelos de linguagem de escrever textos científicos coerentes e plausíveis representa uma ameaça à integridade da pesquisa (Thorp, 2023). A utilização não declarada de IA para escrever artigos pode levar à disseminação de informações falsas ou "alucinadas" na literatura científica, com consequências graves se essas informações forem usadas para fundamentar decisões clínicas (Bender *et al.,* 2021). As revistas científicas e as instituições de pesquisa estão a debater-se com a criação de políticas claras sobre o uso aceitável de IA na redação e na análise de dados. A questão da autoria também se torna complexa. Se uma IA gera uma hipótese inovadora que leva a uma descoberta significativa, deve a IA ser creditada como coautora? (Salvagno; Taccone, 2023). Embora o consenso atual seja que a autoria requer responsabilidade, o que exclui as máquinas, este debate realça como a IA está a desafiar conceitos fundamentais da prática científica (Birhane, 2021). A transparência sobre o uso de ferramentas de IA na pesquisa torna-se um imperativo ético para manter a confiança no empreendimento científico.

8.2. O Futuro da Relação Médico-Paciente

A relação médico-paciente é construída sobre a confiança, a empatia e a comunicação humana (Peabody, 1927). A introdução de um "terceiro" algorítmico nesta relação pode alterá-la de formas profundas (Verghese; Shah; Harrington, 2018). Se o médico passa mais tempo a interagir com uma tela do que com o paciente, ou se a tomada de decisão é percebida como sendo ditada por uma máquina, a dimensão humana e compassiva do cuidado pode ser erodida (Verghese, 2018). Por outro lado, alguns argumentam que a IA pode, na verdade, "re-humanizar" a medicina. Ao assumir tarefas administrativas e analíticas, a IA pode

libertar os médicos para se concentrarem no que os humanos fazem melhor: ouvir, confortar e construir uma aliança terapêutica com o paciente (Topol, 2019). A realização deste potencial, no entanto, depende de uma implementação cuidadosa que priorize a interação humana e utilize a IA como uma ferramenta de apoio, e não como um substituto para o julgamento clínico e a empatia (Meskó; Dhole, 2021).

8.3. A "Mentalidade" da Máquina e o Estatuto Moral

À medida que os sistemas de IA generativa se tornam mais sofisticados e capazes de conversas que mimetizam a compreensão e a emoção humanas, surgem questões filosóficas sobre o seu estatuto (Boden, 2016). Embora a IA atual não possua consciência, senciência ou intencionalidade, a sua capacidade de gerar respostas empáticas pode levar os pacientes a formar laços emocionais com assistentes de saúde virtuais (Turkle, 2011). Isto levanta dilemas éticos sobre a autenticidade dessas interações e o potencial para manipulação emocional. É eticamente permissível usar uma IA para simular empatia a fim de melhorar a adesão do paciente ao tratamento? (Gardiner; Bickmore, 2021). Embora a questão do estatuto moral da IA pertença, por agora, à ficção científica, os debates sobre a ética da simulação de qualidades humanas na interação com pessoas vulneráveis são extremamente relevantes e atuais (Coeckelbergh, 2020).

9. A Necessidade de Governança e Quadros Regulatórios

A complexidade e a escala dos desafios éticos apresentados pela IA generativa na saúde tornam evidente que a autorregulação pela indústria tecnológica ou a simples adaptação de diretrizes éticas existentes serão insuficientes (Price; Cohen, 2019). É necessária uma abordagem multifacetada à governança, que envolva legislação, regulação, desenvolvimento de padrões técnicos e educação (Viljoen, 2021).

Uma componente crucial da governança é a exigência de transparência e explicabilidade (o chamado XAI - Explainable AI). Os reguladores devem exigir que os sistemas de IA utilizados em contextos de alto risco, como o diagnóstico médico, sejam auditáveis e que as suas decisões possam ser, de alguma forma, explicadas e contestadas (Amann *et al.,* 2020). Embora a explicabilidade total possa ser tecnicamente impossível para alguns modelos, graus de interpretabilidade são essenciais para a responsabilidade e a confiança (Rudin, 2019).

A criação de agências reguladoras, ou a expansão do mandato das existentes (como as agências de medicamentos e dispositivos médicos), para supervisionar os "algoritmos como dispositivos médicos" é outro passo fundamental (Price; Cohen, 2019). Estas agências seriam responsáveis por certificar a segurança, a eficácia e a equidade dos sistemas de IA antes da sua comercialização e por monitorizar o seu desempenho no mundo real, exigindo atualizações ou a sua retirada do mercado se surgirem problemas (FDA, 2021).

A governança de dados é igualmente crítica. Legislações como o Regulamento Geral sobre a Proteção de Dados (RGPD) na Europa fornecem um ponto de partida, mas podem ser necessárias leis mais específicas para os dados de saúde no contexto da IA (European Commission, 2021). Modelos de

"fiduciários de dados" ou "cooperativas de dados", onde os dados dos pacientes são geridos por entidades independentes com o mandato de agir no interesse dos pacientes, são propostas inovadoras que merecem consideração (Kariya; Chopra, 2023).

Finalmente, a educação é uma ferramenta de governança indispensável. Os profissionais de saúde precisam de ser formados para compreender as capacidades e as limitações da IA, para interpretar criticamente os seus resultados e para comunicar eficazmente com os pacientes sobre o seu uso (Masters, 2019). Os pacientes e o público em geral também precisam de uma maior literacia em IA para poderem participar de forma significativa no debate social sobre o futuro da saúde digital (Long; Magerko, 2020).

10. Considerações finais

A inteligência artificial generativa não é apenas mais uma ferramenta no arsenal da área da saúde; é uma tecnologia com o potencial de redefinir fundamentalmente a prática clínica, a pesquisa biomédica e a própria natureza da relação de cuidado. A sua capacidade de criar, sintetizar e inovar oferece promessas imensas para o avanço do bem-estar humano, alinhando-se diretamente com o princípio da beneficência. No entanto, esta mesma capacidade introduz riscos profundos e complexos que desafiam todos os pilares da bioética. A autonomia do paciente é posta em causa pelas novas dinâmicas do consentimento de dados e pela opacidade da tomada de decisão algorítmica. O princípio da não maleficência exige uma vigilância constante contra os danos causados por vieses, erros e falhas de segurança. O princípio da justiça compele-nos a lutar contra a criação de uma nova brecha digital na saúde e a garantir que os benefícios da IA sejam distribuídos de forma equitativa por toda a sociedade.

Navegar nesta nova fronteira exige mais do que soluções técnicas. Requer um diálogo social amplo e inclusivo, envolvendo pacientes, médicos, tecnólogos, filósofos, advogados e decisores políticos. Exige a construção de quadros de governança robustos e adaptáveis, que promovam a inovação responsável e protejam os direitos e a dignidade dos indivíduos. A tarefa não é rejeitar a tecnologia, mas sim moldá-la de acordo com os valores humanísticos que sustentam a ética médica.

O futuro dos serviços de saúde será, inevitavelmente, de colaboração entre a inteligência humana e a artificial. O grande desafio bioético do nosso tempo é garantir que, nesta nova parceria, a tecnologia permaneça ao serviço da humanidade, ampliando a nossa capacidade de cuidar, curar e confortar, sem nunca perder de vista a pessoa única e vulnerável que é o centro de toda a prática de saúde. A jornada está apenas a começar, e as escolhas que fizermos agora irão moldar o panorama da saúde para as gerações vindouras.

Referências

AMANN, J. *et al.* Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. **BMC Medical Informatics and Decision Making**, v. 20, n. 1, p. 310, 2020.

BEAUCHAMP, T. L.; CHILDRESS, J. F. **Principles of Biomedical Ethics**. 8. ed. New York: Oxford University Press, 2019.

BENDER, E. M. *et al.* On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. *In*:** PROCEEDINGS OF THE 2021 ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY. [S. l.]: ACM, 2021. p. 610-623.

BIRHANE, A. Algorithmic injustice: a relational ethics approach. **Patterns**, v. 2, n. 2, 100205, 2021.

BLASIMME, A.; VAYENA, E. Governing health data as a common good. **Journal of Medical Ethics**, v. 45, n. 6, p. 359-360, 2019.

BODEN, M. A. **AI: Its Nature and Future**. Oxford: Oxford University Press, 2016.

CHAR, D. S.; SHAH, N. H.; MAGNUS, D. Implementing machine learning in health care—addressing ethical challenges. **The New England Journal of Medicine**, v. 378, n. 11, p. 981-983, 2018.

CIRILLO, D. *et al.* Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. **npj Digital Medicine**, v. 3, n. 81, 2020.

COECKELBERGH, M. **AI Ethics**. Cambridge, MA: The MIT Press, 2020.

COHEN, I. G.; PRICE, W. N. Governing Generative AI in Health Care. **JAMA**, v. 330, n. 10, p. 899-900, 2023.

DAVENPORT, T.; KALAKOTA, R. The potential for artificial intelligence in healthcare. **Future Healthcare Journal**, v. 6, n. 2, p. 94-98, 2019.

EL EMAM, K. *et al.* Evaluating the privacy risks of synthetic health data. **Journal of Medical Internet Research**, v. 22, n. 4, e18158, 2020.

EUROPEAN COMMISSION. **Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)**. Brussels: European Commission, 2021.

FADEN, R. R.; BEAUCHAMP, T. L. **A History and Theory of Informed Consent**. New York: Oxford University Press, 1986.

FINLAYSON, S. G. *et al.* Adversarial attacks against medical machine learning. **Science**, v. 363, n. 6433, p. 1287-1289, 2019.

FLEMING, N. How artificial intelligence is changing drug discovery. **Nature**, v. 557, p. S55-S57, 2018.

FLORIDI, L. Translating principles into practices: a framework for the ethics of AI in healthcare. **Philosophy & Technology**, v. 32, n. 4, p. 571-574, 2019.

GAO, C. *et al.* A comprehensive survey on generative ai. **arXiv preprint arXiv:2303.11717**, 2023.

GARDINER, P. M.; BICKMORE, T. W. The Role of Empathetic Avatars in Health Coaching. **American Journal of Health Promotion**, v. 35, n. 6, p. 858-861, 2021.

GHASSEMI, M. *et al.* The AI Divide in Healthcare. **The Lancet Digital Health**, v. 3, n. 9, e542-e543, 2021.

GILLE, F. *et al.* The role of artificial intelligence in shared decision making: a

scoping review. **Journal of Medical Internet Research**, v. 22, n. 9, e17928, 2020.

GOODFELLOW, I. *et al.* **Deep Learning**. Cambridge, MA: The MIT Press, 2016.

JI, Z. *et al.* Survey of Hallucination in Natural Language Generation. **ACM Computing Surveys**, v. 55, n. 12, p. 1-38, 2023.

KARIYA, T.; CHOPRA, V. A Data Fiduciary Model for Health Data. **JAMA**, v. 330, n. 1, p. 15-16, 2023.

KEANE, P. A.; TOPOL, E. J. With an eye to AI and accuracy. **npj Digital Medicine**, v. 1, n. 25, 2018.

KORTZ, T.; LEE, H. Large Language Models in Oncology: Many Opportunities and One Major Risk. **JCO Clinical Cancer Informatics**, v. 7, e2300021, 2023.

LONG, D.; MAGERKO, B. What is AI Literacy? Competencies and Design Considerations. *In*:** PROCEEDINGS OF THE 2020 CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS. [S. l.: s. n.], 2020. p. 1–16.

MASTERS, K. Artificial intelligence in medical education. **Medical Teacher**, v. 41, n. 9, p. 976-980, 2019.

MATTHIAS, A. The responsibility gap: Ascribing responsibility for the actions of learning automata. **Ethics and Information Technology**, v. 6, n. 3, p. 175-183, 2004.

MESKÓ, B.; DHOLE, C. The clinician's guide to the role of artificial intelligence in preventing physician burnout. **Journal of Medical Internet Research**, v. 23, n. 9, e29532, 2021.

NAGENDRAN, M. *et al.* Artificial intelligence versus clinicians in disease diagnosis: a systematic review and meta-analysis. **BMJ**, v. 368, m689, 2020.

NUNES, R. **Bioética**. 4. ed. Lisboa: Almedina, 2020.

NYHOLM, S. Attributing agency to autonomous systems: Reflections on the moral responsibility of manufacturers. **Science and Engineering Ethics**, v. 24, n. 4, p. 1201-1219, 2018.

OBERMEYER, Z. *et al.* Dissecting racial bias in an algorithm used to manage the health of populations. **Science**, v. 366, n. 6464, p. 447-453, 2019.

PARIKH, R. B.; OBERMEYER, Z. Overcoming Automation Bias and Deference—The Case for Human-Centered AI in Medicine. **The American Journal of Bioethics**, v. 21, n. 5, p. 14-16, 2021.

PEABODY, F. W. The Care of the Patient. **JAMA**, v. 88, n. 12, p. 877-882, 1927.

PRICE, W. N.; COHEN, I. G. Privacy in the age of medical big data. **Nature Medicine**, v. 25, p. 37-43, 2019.

RAJKOMAR, A. *et al.* Ensuring fairness in machine learning to advance health equity. **Annals of Internal Medicine**, v. 169, n. 12, p. 866-872, 2018.

RAO, A. *et al.* Assessing the potential of GPT-4 to Ccot clinical notes. **NEJM Catalyst Innovations in Care Delivery**, 2023.

RAWLS, J. **A Theory of Justice**. Cambridge, MA: Harvard University Press, 1971.

REDDY, S. *et al.* Artificial intelligence and health equity: a scoping review. **Journal of the National Medical Association**, v. 112, n. 5, p. 508-522, 2020.

ROCHER, L. *et al.* Estimating the success of re-identifications in incomplete datasets using generative models. **Nature Communications**, v. 10, n. 3069, 2019.

RUDIN, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. **Nature Machine Intelligence**, v. 1, p. 206-215, 2019.

SALLAM, M. Generative AI in public health: applications, challenges and ethical considerations. **Frontiers in Public Health**, v. 11, 1238670, 2023.

SALVAGNO, M.; TACCONE, F. S. When the author is a Large Language Model: can we trust it? A case report for the scientific community. **Critical Care**, v. 27, n. 150, 2023.

SHAY, G.; COHEN, I. G. Rethinking data sharing and consent for the AI era. **The New England Journal of Medicine**, v. 389, n. 1, p. 1-4, 2023.

SIECK, C. J. *et al.* Technology as a social determinant of health: The role of technology in health inequities and a call for participatory design. **JMIR**, v. 23, n. 7, e27434, 2021.

SINGHAL, K. *et al.* Large language models encode clinical knowledge. **Nature**, v. 620, p. 172-180, 2023.

THORP, H. H. ChatGPT is fun, but not an author. **Science**, v. 379, n. 6629, p. 313, 2023.

TOPOL, E. J. **Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again**. New York: Basic Books, 2019.

TURKLE, S. **Alone Together: Why We Expect More from Technology and Less from Each Other**. New York: Basic Books, 2011.

U.S. FOOD AND DRUG ADMINISTRATION. **Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan**. Silver Spring, MD: FDA, 2021.

VAIDHYANATHAN, S. **Antisocial Media: How Facebook Disconnects Us and Undermines Democracy**. New York: Oxford University Press, 2018.

VASWANI, A. *et al.* Attention is All You Need. *In*:** ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 30. [S. l.: s. n.], 2017.

VERGHESE, A. The Pen and the Stethoscope. **The New England Journal of Medicine**, v. 379, p. 2381-2383, 2018.

VERGHESE, A.; SHAH, N. H.; HARRINGTON, R. A. What this computer needs is a physician: humanism and artificial intelligence. **JAMA**, v. 319, n. 1, p. 19-20, 2018.

VILJOEN, S. A relational theory of data governance. **Yale Law Journal**, v. 131, 2021.

WACHTER, R. M. **The Digital Doctor: Hope, Hype, and Harm at the Dawn of Medicine's Computer Age**. New York: McGraw-Hill, 2017.

WAHL, B. *et al.* Artificial intelligence and global health: opportunities and challenges. **The Lancet**, v. 392, n. 10164, p. 2489-2491, 2018.

WANG, D. *et al.* Scientific discovery in the age of artificial intelligence. **Nature**, v. 620, p. 47-60, 2023.

WORLD HEALTH ORGANIZATION. **Ethics and governance of artificial intelligence for health: WHO guidance**. Geneva: World Health Organization, 2021.

YI, X.; WALIA, E.; BABYN, P. Generative adversarial network in medical imaging: A review. **Medical Image Analysis**, v. 58, 101552, 2019.

YU, K. H.; BEAM, A. L.; KOHANE, I. S. Artificial intelligence in healthcare. **Nature Biomedical Engineering**, v. 2, n. 10, p. 719-731, 2018.

ZHAVORONKOV, A. *et al.* Deep learning enables rapid identification of potent DDR1 kinase inhibitors. **Nature Biotechnology**, v. 37, n. 9, p. 1038-1040, 2019.

ZOU, J.; SCHIEBINGER, L. AI can be sexist and racist — it's time to make it fair. **Nature**, v. 559, p. 324-326, 2018.