

ENTRE ALGORITMOS E DECISÕES CLÍNICAS: O PAPEL DA INTELIGÊNCIA ARTIFICIAL NA CONSTRUÇÃO DO DIAGNÓSTICO EM SAÚDE

BETWEEN ALGORITHMS AND CLINICAL DECISIONS: THE ROLE OF ARTIFICIAL INTELLIGENCE IN BUILDING A HEALTH DIAGNOSIS

ENTRE ALGORITMOS Y DECISIONES CLÍNICAS: EL PAPEL DE LA INTELIGENCIA ARTIFICIAL EN LA ELABORACIÓN DE DIAGNÓSTICOS DE SALUD



10.56238/sevened2026.016-015

Ana Karine Brito do Couto

Instituição: Faculdade de Odontologia de Bauru – Universidade de São Paulo (FOB-USP)

E-mail: anacouto26.1@usp.br

Orcid: <https://orcid.org/0009-0000-6174-989X>

Lattes: <http://lattes.cnpq.br/4294872216599907>

George Monteiro Filho

Instituição: Faculdade de Odontologia de Bauru – Universidade de São Paulo (FOB-USP)

E-mail: georgemonteirofilho@usp.br

Orcid: <https://orcid.org/0009-0002-6930-8604>

Lattes: <http://lattes.cnpq.br/9029229132614734>

Kettelyn Macêdo da Cruz

Instituição: Faculdade de Odontologia de Bauru – Universidade de São Paulo (FOB-USP)

E-mail: kettelynmacedo@usp.br

Orcid: <https://orcid.org/0009-0005-0745-4990>

Lattes: <http://lattes.cnpq.br/3399865174340480>

Raíssa Estefane Vaz Damião

Instituição: Faculdade de Odontologia de Bauru – Universidade de São Paulo (FOB-USP)

E-mail: raissadamiao@usp.br

Orcid: <https://orcid.org/0009-0007-8728-4788>

Lattes: <https://lattes.cnpq.br/5130051174613911>

RESUMO

A incorporação da Inteligência Artificial (IA) no diagnóstico em saúde representa uma transformação qualitativa na prática clínica, distinta de todas as inovações tecnológicas anteriores pela capacidade de analisar, interpretar e gerar inferências a partir de grandes volumes de dados. O capítulo examina essa transformação a partir de suas raízes históricas, percorrendo os paradigmas que moldaram o diagnóstico desde a medicina hipocrática, com seu saber empírico e centrado na narrativa do paciente, até a era digital, passando pela medicina anatomo-clínica, pelo advento da tecnologia diagnóstica e pela consolidação da Medicina Baseada em Evidências. Em seguida, apresenta os fundamentos técnicos das principais abordagens de IA, incluindo aprendizado de máquina, aprendizado profundo,

processamento de linguagem natural e integração de Big Data, discutindo como cada uma dessas abordagens opera e quais são suas capacidades e limitações específicas no contexto clínico. O capítulo analisa as implicações dessa incorporação para a própria natureza do diagnóstico, que passa de um ato interpretativo pontual para um processo probabilístico e contínuo, explicitando a incerteza inerente à prática clínica e reconfigurando a dinâmica de autoridade entre profissional e sistema. São examinados fenômenos como a opacidade algorítmica, os vieses sistemáticos decorrentes de dados históricos não representativos, a tensão entre a singularidade do paciente e a lógica generalizante dos algoritmos, e o risco de viés de automação na tomada de decisão. As aplicações contemporâneas da IA são discutidas em seus contextos mais relevantes, incluindo o diagnóstico por imagem, a prática odontológica e os sistemas de apoio à decisão clínica, com atenção tanto ao desempenho demonstrado na literatura quanto às limitações que impedem a transposição direta dos resultados de pesquisa para o mundo real. O capítulo aborda ainda as dimensões éticas da IA no diagnóstico, tratando de princípios como autonomia, equidade e responsabilização, bem como os desafios regulatórios impostos pela natureza adaptativa desses sistemas, e discute as implicações para a formação profissional em saúde, argumentando que preparar o profissional contemporâneo exige integrar fundamentos tecnológicos ao raciocínio clínico, sem substituir as competências humanas que os algoritmos ainda não conseguem reproduzir. Conclui-se que o futuro do diagnóstico depende da construção deliberada de uma parceria entre humano e algoritmo que preserve a capacidade analítica da IA e o julgamento contextual, ético e humano do profissional de saúde, com responsabilidades compartilhadas entre quem desenvolve, regula, forma e cuida.

Palavras-chave: Inteligência Artificial. Saúde. Diagnóstico. Aprendizado Profundo. Aprendizado de Máquina.

ABSTRACT

The incorporation of Artificial Intelligence (AI) in healthcare diagnosis represents a qualitative transformation in clinical practice, distinct from all previous technological innovations due to its ability to analyze, interpret, and generate inferences from large volumes of data. This chapter examines this transformation from its historical roots, tracing the paradigms that have shaped diagnosis from Hippocratic medicine, with its empirical knowledge centered on the patient's narrative, to the digital age, passing through anatomical-clinical medicine, the advent of diagnostic technology, and the consolidation of Evidence-Based Medicine. It then presents the technical foundations of the main AI approaches, including machine learning, deep learning, natural language processing, and Big Data integration, discussing how each of these approaches operates and what their specific capabilities and limitations are in the clinical context. The chapter analyzes the implications of this incorporation for the very nature of diagnosis, which shifts from a punctual interpretive act to a probabilistic and continuous process, highlighting the uncertainty inherent in clinical practice and reconfiguring the dynamics of authority between professional and system. This chapter examines phenomena such as algorithmic opacity, systematic biases arising from unrepresentative historical data, the tension between patient uniqueness and the generalizing logic of algorithms, and the risk of automation bias in decision-making. Contemporary AI applications are discussed in their most relevant contexts, including diagnostic imaging, dental practice, and clinical decision support systems, paying attention both to the performance demonstrated in the literature and to the limitations that prevent the direct transposition of research results to the real world. The chapter also addresses the ethical dimensions of AI in diagnosis, dealing with principles such as autonomy, equity, and accountability, as well as the regulatory challenges imposed by the adaptive nature of these systems, and discusses the implications for professional health training, arguing that preparing the contemporary professional requires integrating technological fundamentals with clinical reasoning, without replacing the human skills that algorithms are not yet able to reproduce. It is concluded that the future of diagnosis depends on the deliberate construction of a partnership between humans and algorithms that preserves the analytical

capacity of AI and the contextual, ethical, and human judgment of the healthcare professional, with shared responsibilities among those who develop, regulate, train, and provide care.

Keywords: Artificial Intelligence. Health. Diagnosis. Deep Learning. Machine Learning.

RESUMEN

La incorporación de la Inteligencia Artificial (IA) al diagnóstico sanitario representa una transformación cualitativa en la práctica clínica, distinta de todas las innovaciones tecnológicas anteriores debido a su capacidad para analizar, interpretar y generar inferencias a partir de grandes volúmenes de datos. Este capítulo examina esta transformación desde sus raíces históricas, rastreando los paradigmas que han configurado el diagnóstico desde la medicina hipocrática, con su conocimiento empírico centrado en la narrativa del paciente, hasta la era digital, pasando por la medicina anatómica-clínica, el advenimiento de la tecnología diagnóstica y la consolidación de la Medicina Basada en la Evidencia. A continuación, presenta los fundamentos técnicos de los principales enfoques de IA, incluyendo el aprendizaje automático, el aprendizaje profundo, el procesamiento del lenguaje natural y la integración de macrodatos, analizando cómo opera cada uno de estos enfoques y cuáles son sus capacidades y limitaciones específicas en el contexto clínico. El capítulo analiza las implicaciones de esta incorporación para la naturaleza misma del diagnóstico, que pasa de un acto interpretativo puntual a un proceso probabilístico y continuo, destacando la incertidumbre inherente a la práctica clínica y reconfigurando la dinámica de autoridad entre el profesional y el sistema. Este capítulo examina fenómenos como la opacidad algorítmica, los sesgos sistemáticos derivados de datos históricos no representativos, la tensión entre la singularidad del paciente y la lógica generalizadora de los algoritmos, y el riesgo de sesgo de automatización en la toma de decisiones. Se analizan las aplicaciones contemporáneas de la IA en sus contextos más relevantes, incluyendo la imagenología diagnóstica, la práctica odontológica y los sistemas de apoyo a la decisión clínica, prestando atención tanto al rendimiento demostrado en la literatura como a las limitaciones que impiden la transposición directa de los resultados de la investigación al mundo real. El capítulo también aborda las dimensiones éticas de la IA en el diagnóstico, tratando principios como la autonomía, la equidad y la responsabilidad, así como los desafíos regulatorios que impone la naturaleza adaptativa de estos sistemas, y analiza las implicaciones para la formación profesional en salud, argumentando que la preparación del profesional contemporáneo requiere integrar los fundamentos tecnológicos con el razonamiento clínico, sin reemplazar las habilidades humanas que los algoritmos aún no pueden reproducir. Se concluye que el futuro del diagnóstico depende de la construcción deliberada de una colaboración entre humanos y algoritmos que preserve la capacidad analítica de la IA y el juicio contextual, ético y humano del profesional de la salud, con responsabilidades compartidas entre quienes desarrollan, regulan, capacitan y brindan atención.

Palabras clave: Inteligencia Artificial. Salud. Diagnóstico. Aprendizaje Profundo. Aprendizaje Automático.

1 INTRODUÇÃO

A incorporação de tecnologias digitais na saúde não constitui um fenômeno recente, mas a ascensão da Inteligência Artificial (IA) representa algo qualitativamente distinto das inovações anteriores. Enquanto tecnologias tradicionais, como o estetoscópio ampliaram a percepção sensorial do clínico, e os métodos de imagem possibilitaram a visualização de estruturas antes inacessíveis, a Inteligência Artificial (IA) introduz um novo nível de atuação no processo diagnóstico: a capacidade de analisar, interpretar e gerar inferências a partir de grandes volumes de dados.

Essa evolução tem implicações relevantes para o processo diagnóstico. O raciocínio clínico constitui o alicerce da prática em saúde, exigindo que os profissionais sintetizem informações complexas, reconheçam padrões e apliquem causalidade para alcançar diagnósticos precisos, entretanto a cognição humana é limitada. A partir disso, o diagnóstico clínico passa a ser mediado por algoritmos treinados em populações, operando por lógica probabilística e, frequentemente, não oferece explicações legíveis ao profissional que os utiliza. Sendo assim, o que muda não é apenas o processo: muda a própria natureza do saber clínico, a maneira como o conhecimento é produzido, validado e aplicado à beira do leito ou na cadeira do consultório.

A transformação não ocorre de forma isolada. Ela se insere em um contexto mais amplo de digitalização acelerada dos sistemas de saúde, de crescente disponibilidade de dados clínicos em formato eletrônico e de pressão por maior eficiência nos processos assistenciais. A IA emerge, nesse contexto, como uma promessa de solução para problemas antigos: variabilidade diagnóstica, sobrecarga dos profissionais, dificuldade de acesso a especialistas em regiões remotas, demora na detecção de doenças em estágio inicial.

Ao mesmo tempo, as promessas associadas à IA são sustentadas por evidências consistentes. Em atividades específicas e sob condições controladas, sistemas baseados em aprendizado profundo têm demonstrado desempenho comparável ou superior ao de especialistas humanos na interpretação de imagens. De forma semelhante, ferramentas de apoio à decisão tem mostrado potencial para reduzir erros e apoiar decisões em cenários clínicos de maior complexidade. A IA detecta padrões em grandes volumes de dados que nenhum clínico conseguiria processar manualmente. Ignorar esses avanços seria tão equivocado quanto adotá-los sem senso crítico.

É precisamente nesse tensionamento que este capítulo se situa. Seu objetivo é examinar o papel da IA na construção do diagnóstico em saúde, não como celebração tecnológica nem como alarmismo, mas como análise crítica de um fenômeno que reconfigura práticas, responsabilidades e fundamentos epistemológicos da clínica. Para isso, percorre as raízes históricas do diagnóstico, examina os mecanismos técnicos da IA, discute suas aplicações contemporâneas e seus limites, e propõe uma reflexão sobre o que significa, afinal, diagnosticar na era dos dados.

A estrutura do capítulo segue uma lógica deliberada: começa pelo passado, porque compreender o impacto da IA exige saber de onde veio o diagnóstico clínico; avança para os fundamentos técnicos, porque avaliar criticamente uma tecnologia requer conhecê-la minimamente; e chega às implicações práticas, éticas e formativas, porque, a partir desse momento, as transformações se tornam concretas para os profissionais de saúde. O fio condutor, ao longo de todo o percurso, é a pergunta sobre como integrar o que a máquina oferece ao que o clínico não pode deixar de ser.

2 EVOLUÇÃO DO DIAGNÓSTICO EM SAÚDE: BASES HISTÓRICAS E EPISTEMOLÓGICAS

Compreender o impacto da IA no diagnóstico em saúde exige recuar no tempo, pois cada inovação tecnológica na medicina reflete uma visão específica sobre o conhecimento, a doença e o papel do clínico. O diagnóstico não é um procedimento neutro: é uma construção histórica, moldada por paradigmas que se sucedem sem nunca se apagar completamente. As marcas de cada período persistem na prática contemporânea, muitas vezes sem que os próprios profissionais as percebam.

O diagnóstico médico, historicamente fundamentado na observação clínica e na experiência do profissional desde a medicina hipocrática, consolidou-se como um processo essencial que integra análise, interpretação e tomada de decisão. Ao longo do tempo, evoluiu de uma prática empírica para um modelo cada vez mais complexo, incorporando tecnologias avançadas e, mais recentemente, a Inteligência Artificial, que não rompe com essa trajetória, mas amplia na construção diagnóstica.

2.1 O DIAGNÓSTICO NA MEDICINA CLÁSSICA: OBSERVAÇÃO, NARRATIVA E SUBJETIVIDADE

Na tradição hipocrática, o diagnóstico era inseparável da narrativa. O médico ouvia o paciente, observava sua aparência, sua urina, seu modo de respirar e de mover-se, e integrava tudo isso em uma história coerente sobre o desequilíbrio que acometia aquele corpo. A doença não era considerada isoladamente, mas uma condição singular, contextual, inseparável da pessoa que a vivia e do ambiente em que ela existia.

Esse modelo tinha o clínico como intérprete central. Sua experiência, sua memória de casos anteriores e sua sensibilidade eram os instrumentos diagnósticos mais sofisticados disponíveis. Não havia exame laboratorial, não havia imagem. O que havia era observação acurada, escuta atenta e raciocínio ancorado na tradição transmitida de mestre a discípulo. A subjetividade não era um defeito do método: era sua condição de possibilidade.

Do ponto de vista epistemológico, esse período se caracteriza por um saber empírico e contextual, no qual a singularidade do paciente constitui simultaneamente o ponto de partida e de chegada do processo diagnóstico. O conhecimento não é transferível de forma abstrata: ele reside na

experiência acumulada do clínico e na sua capacidade de interpretar cada caso como único. Essa característica, frequentemente vista como limitação pelos paradigmas posteriores, é hoje reconhecida como um elemento que a IA ainda não consegue reproduzir adequadamente.

2.2 A VIRADA ANATOMO-CLÍNICA: O CORPO COMO OBJETO DE INVESTIGAÇÃO

A ruptura começa nos séculos XVIII e XIX, com a medicina anatomo-clínica, uma mudança estrutural na forma de investigar e classificar a doença. A autópsia permitiu correlacionar sintomas com alterações morfológicas internas; o estetoscópio, tornou acessíveis fenômenos fisiológicos antes imperceptíveis. A doença passou a ser localizada em estruturas anatômicas específicas (órgãos, tecidos, lesões identificáveis). Esse deslocamento reorganizou os fundamentos do diagnóstico: a queixa do paciente deixou de ser o dado central do processo clínico, cedendo lugar à lesão como critério de validação diagnóstica.

Trata-se de uma transição epistemológica decisiva. O diagnóstico se torna progressivamente objetivo, verificável e potencialmente reprodutível. O conhecimento médico passa a se organizar em categorias compartilhadas e socialmente validadas, reduzindo sua dependência exclusiva da experiência individual do clínico. A doença deixa de ser uma experiência subjetiva e passa a ser definida como entidade biológica localizada, descrita em termos anatômicos precisos e identificável por profissionais devidamente treinados.

Nesse período, consolida-se uma abordagem de base positivista do conhecimento clínico, na qual o real é associado ao que pode ser observado, mensurado e verificado. Essa lógica abrirá caminho para todas as transformações tecnológicas que virão a seguir, pois estabelece que o corpo pode e deve ser investigado de forma cada vez mais objetiva, por meio de instrumentos cada vez mais precisos.

2.3 A EMERGÊNCIA DA TECNOLOGIA DIAGNÓSTICA: AMPLIAÇÃO DA PERCEPÇÃO CLÍNICA

O século XX radicaliza a tendência inaugurada pela medicina anatomo-clínica. A radiografia, a tomografia, a ressonância magnética, os analisadores laboratoriais automatizados: cada tecnologia expande o alcance dos sentidos do clínico, mas também introduz uma mediação progressiva entre o profissional e o paciente. O diagnóstico passa a depender, cada vez mais, de imagens produzidas por máquinas e de valores gerados por equipamentos. O clínico interpreta resultados, não apenas sinais.

Essa mediação não é neutra. Ela implica novos saberes técnicos, novos vocabulários e, sobretudo, uma nova distribuição de autoridade. Parte do conhecimento diagnóstico migra do clínico para o instrumento. O especialista em medicina de laboratório, o radiologista, o patologista tornam-se figuras centrais em um processo diagnóstico que se fragmenta em múltiplas especialidades. O clínico

geral deixa de ser o único intérprete e passa a ser, em muitos casos, o articulador de saberes produzidos em outros lugares.

É nesse contexto que se instala a tensão, ainda presente, entre a medicina como arte interpretativa e a medicina como ciência aplicada. O avanço tecnológico aumentou a precisão e a objetividade do diagnóstico, mas também criou novas formas de incerteza: a incerteza sobre como integrar múltiplos resultados de exames, sobre como lidar com achados incidentais, sobre como comunicar ao paciente probabilidades e riscos que os próprios profissionais têm dificuldade de compreender.

2.4 MEDICINA BASEADA EM EVIDÊNCIAS: PADRONIZAÇÃO E RACIONALIZAÇÃO DA DECISÃO CLÍNICA

Na segunda metade do século XX, consolidou-se a Medicina Baseada em Evidências (MBE), que propõe hierarquizar o conhecimento clínico a partir de critérios metodológicos rigorosos. Ensaios clínicos randomizados, revisões sistemáticas e metanálises passam a orientar condutas. O objetivo declarado é reduzir a variabilidade arbitrária e fundamentar decisões em dados populacionais robustos, diminuindo a dependência do julgamento individual do clínico.

A MBE representa um avanço significativo na prática médica. Ao introduzir rigor metodológico onde havia tradição não examinada, substituiu a autoridade de opinião por hierarquia de evidências e criou instrumentos para avaliar a qualidade do conhecimento disponível. No entanto, apresenta limitações relevantes para a compreensão do papel posterior da IA. Sua lógica é particularmente eficaz no nível populacional, mas encontra desafios ao lidar com a singularidade de cada paciente. Nesse sentido, a crítica que se fará à IA data-driven, de que generaliza demais e individualiza de menos, é, em certa medida, uma crítica que já se fazia à própria MBE.

Além disso, a MBE pressupõe que o profissional seja capaz de acessar, interpretar e aplicar a evidência disponível. Diante do crescimento exponencial da produção científica, essa pressuposição tornou-se cada vez mais difícil de sustentar. A distância entre o conhecimento produzido pela pesquisa e o conhecimento efetivamente utilizado na prática clínica cotidiana permanece um problema não resolvido, e é precisamente aí que os sistemas de apoio à decisão baseados em IA prometem atuar.

2.5 A TRANSIÇÃO PARA O PARADIGMA DIGITAL: DADOS, CONECTIVIDADE E COMPLEXIDADE

O final do século XX e o início do século XXI introduzem uma nova dimensão no campo da saúde: a digitalização dos processos clínicos. Prontuários eletrônicos, sistemas de imagem integrados, bases de dados populacionais e dispositivos de monitoramento remoto contribuem para a constituição de um ecossistema clínico, que se torna progressivamente informacional. Nesse contexto, o conceito

de Big Data em saúde não designa apenas um aumento de volume de dados, mas uma mudança qualitativa na forma de produzir e organizar o conhecimento clínico.

Esse cenário constitui o substrato no qual a IA se desenvolve. Sem dados digitais em larga escala, o aprendizado de máquina não existe. A IA, nesse sentido, é produto do paradigma digital e amplifica suas características essenciais: a quantificação, a padronização, a análise em escala. O que muda com a IA não é apenas a quantidade de dados processados, mas a forma como os padrões nesses dados são identificados e utilizados para orientar decisões.

A digitalização também criou novas vulnerabilidades. Dados de saúde em formato digital são ativos valiosos, sujeitos a violações de privacidade, uso indevido por terceiros e exploração comercial. A tensão entre o potencial analítico dos dados e a proteção dos direitos dos pacientes é um dos temas centrais da governança da saúde digital, e não encontra soluções simples.

2.6 A EMERGÊNCIA DA INTELIGÊNCIA ARTIFICIAL: DO CONHECIMENTO BASEADO EM HIPÓTESES AO CONHECIMENTO ORIENTADO POR DADOS

A ruptura introduzida pela IA é mais profunda do que parece à primeira vista. Na ciência clínica tradicional, o conhecimento se organiza em torno de hipóteses: formula-se uma explicação sobre um fenômeno, testa-se empiricamente, valida-se ou rejeita-se. Há uma lógica causal que sustenta o processo: queremos saber não apenas o que acontece, mas por quê acontece. Entretanto, a IA data-driven opera de outra forma, não parte de hipóteses, mas de padrões: correlações estatísticas identificadas em grandes conjuntos de dados, frequentemente sem que se saiba por que aquelas variáveis se relacionam.

Isso levanta questões epistemológicas que não têm resposta simples. Em que medida é possível confiar em modelos que explicam o quê, mas não o porquê? Como integrar inferência algorítmica com raciocínio causal? O que acontece com a clínica quando o conhecimento deixa de ser explicado e passa a ser apenas calculado? Essas perguntas não têm respostas definitivas no estado atual da discussão científica e filosófica, mas precisam ser formuladas com clareza para que a adoção da IA em saúde seja responsável.

É relevante destacar que a tensão entre conhecimento correlacional e conhecimento causal não é inédita na medicina. Áreas como a epidemiologia trabalham com correlações muito antes da IA existir. O que a IA muda é a escala e a opacidade: nunca antes foi possível identificar correlações em tantas variáveis simultaneamente, e nunca antes foi tão difícil compreender como essas correlações foram identificadas.

2.7 PARADIGMAS EM TENSÃO: INTERPRETAÇÃO CLÍNICA VERSUS INFERÊNCIA ALGORÍTMICA

A história do diagnóstico é, em síntese, uma história de tensões entre modelos de conhecimento: entre subjetividade e objetividade, entre experiência e evidência, entre arte clínica e protocolo científico. A IA não resolve essas tensões, na verdade, ela as intensifica, ao propor uma forma de conhecimento que é simultaneamente mais poderosa e menos compreensível do que qualquer paradigma anterior.

O desafio contemporâneo não reside em escolher entre o clínico e o algoritmo como alternativas excludentes, mas em desenvolver práticas capazes de articular criticamente as contribuições de ambos. Isso implica reconhecer as potencialidades e as limitações da IA no processo diagnóstico. Essa compreensão é fundamental para orientar seu uso de forma responsável e será aprofundado nas seções seguintes.

2.8 IMPLICAÇÕES PARA A FORMAÇÃO E PRÁTICA PROFISSIONAL

As transformações históricas e epistemológicas do diagnóstico têm consequências diretas para quem forma e para quem é formado nas ciências da saúde. Cada novo paradigma exigiu a atualização das competências profissionais: a medicina anatomo-clínica exigiu anatomia sistemática; o advento dos exames laboratoriais exigiu interpretação de valores de referência; a MBE exigiu estatística e metodologia de pesquisa. A IA exige mais uma camada de competências, que precisa ser integrada à formação sem substituir as anteriores.

O profissional contemporâneo precisa transitar entre paradigmas. Precisa saber interpretar uma curva de aprendizado de máquina com a mesma naturalidade com que interpreta um hemograma. Isso não significa formar cientistas de dados, mas garantir que a alfabetização digital seja parte constitutiva da competência clínica, e não um complemento opcional para os interessados em tecnologia.

3 FUNDAMENTOS DA INTELIGÊNCIA ARTIFICIAL APLICADA AO DIAGNÓSTICO EM SAÚDE

A Inteligência Artificial (IA) não constitui uma tecnologia monolítica. O termo engloba um conjunto de abordagens distintas, com pressupostos, capacidades e limitações específicas, que se aplicam a diferentes tipos de problemas. Discutir IA em saúde sem distinguir essas abordagens pode levar a interpretações inadequadas que comprometam sua compreensão crítica. Para avaliar adequadamente seu papel no diagnóstico, é necessário compreender minimamente como cada um desses diferentes modelos funcionam.

3.1 APRENDIZADO DE MÁQUINA (MACHINE LEARNING)

O aprendizado de máquina (Machine Learning, ML) é o núcleo técnico da IA contemporânea. Em vez de ser programado com regras explícitas, um algoritmo de ML aprende padrões a partir de dados, ajustando seus parâmetros internos até que suas previsões se tornem progressivamente mais precisas. A ideia central é simples: dado um conjunto suficientemente grande de exemplos rotulados, o algoritmo aprende a generalizar, aplicando o que aprendeu em casos novos que nunca viu antes.

Os modelos supervisionados, amplamente utilizados no diagnóstico clínico, são treinados a partir de dados rotulados, como exames associados a diagnósticos previamente confirmados. Nesse contexto, o algoritmo aprende padrões discriminativos e relações entre variáveis clínicas e desfechos, e, assim, passa a aplicar essa aprendizagem a casos novos. Modelos não supervisionados, por sua vez, identificam estruturas ocultas em dados não rotulados, sendo úteis para descobrir subgrupos clínicos que ainda não foram descritos na literatura. Já os modelos por reforço aprendem por meio de interação com o ambiente, otimizando sequências de decisões ao longo do tempo, com aplicações potenciais em protocolos de tratamento.

O desempenho desses modelos depende criticamente da qualidade, quantidade e da representatividade dos dados utilizados no treinamento. Um modelo treinado com dados insuficientes, enviesados ou pouco representativos da população de interesse terá desempenho comprometido precisamente nos casos em que seria mais necessário. Essa dependência dos dados constitui uma das principais limitações dos sistemas de ML em contextos clínicos reais.

Na prática clínica, modelos de aprendizado de máquina têm sido amplamente aplicados à análise de imagens médicas, sendo capazes de realizar tarefas como detecção, classificação e segmentação de achados relevantes para o diagnóstico. Esses sistemas podem identificar padrões em dados de imagem e atuar como ferramentas de apoio ao diagnóstico, contribuindo para a interpretação de exames e para a tomada de decisão clínica.

3.2 APRENDIZADO PROFUNDO (DEEP LEARNING)

O aprendizado profundo (*deep learning*) é uma subárea do ML que utiliza redes neurais artificiais com múltiplas camadas de processamento para lidar com dados de alta complexidade. Sua principal vantagem em relação aos modelos tradicionais de ML é a capacidade de extrair características relevantes diretamente dos dados brutos, sem a necessidade de definição prévia de atributos por um especialista. Sendo assim, em vez de programar explicitamente quais características devem ser identificadas em uma imagem, os modelos de *deep learning* aprendem, de forma autônoma, quais padrões são mais informativos para a tarefa em questão.

Na prática clínica, isso se traduz em algoritmos capazes de identificar padrões em radiografias, tomografias, imagens histopatológicas e exames de fundo de olho com precisão comparável e, em

contextos específicos, potencialmente superior ao de especialistas humanos. As redes neurais convolucionais, especialmente desenvolvidas para análise de imagens, têm se consolidado como a principal técnica no diagnóstico por imagem assistida por IA. Na odontologia, esses modelos vêm sendo aplicados na detecção de cáries, lesões periapicais e perda óssea periodontal em exames radiográficos.

Uma das principais limitações do *deep learning* é, ao mesmo tempo, sua principal característica: os modelos aprendem representações internas dos dados que não são diretamente interpretáveis por humanos. O que exatamente o algoritmo aprendeu a identificar em uma imagem de radiografia, e por que essa identificação funciona, frequentemente não é possível de descrever em termos clínicos compreensíveis. Isso é o que origina o problema da explicabilidade, discutido em seção posterior.

3.3 PROCESSAMENTO DE LINGUAGEM NATURAL (PLN)

Uma parcela significativa do conhecimento clínico está registrado em linguagem natural: evoluções de prontuário, laudos médicos, relatos de anamnese, discussões de caso. Esse conhecimento, produzido em volume enorme ao longo de décadas, permaneceu durante muito tempo inacessível para análise computacional sistemática. O processamento de linguagem natural (PLN) muda esse cenário, permitindo que sistemas computacionais extraiam informações relevantes de textos não estruturados com precisão crescente.

No contexto diagnóstico, o PLN viabiliza a extração automática de informações a partir de prontuários eletrônicos, a identificação de padrões linguísticos associados a diferentes condições clínicas, a síntese de evidências presentes na literatura científica e o desenvolvimento de sistema de apoio à decisão integrados ao fluxo clínico. A capacidade de processar a narrativa clínica, e não apenas dados estruturados, representa um avanço significativo na direção de um sistema de IA verdadeiramente integrado à complexidade da prática em saúde.

3.4 BIG DATA E INTEGRAÇÃO DE DADOS EM SAÚDE

A eficácia de qualquer sistema de IA depende, antes de qualquer outra consideração técnica, da qualidade e quantidade dos dados disponíveis. O conceito de Big Data em saúde designa a coleta e análise integrada de dados provenientes de múltiplas fontes: prontuários eletrônicos, resultados de exames laboratoriais, imagens médicas, dados genômicos, registros de dispositivos vestíveis e informações provenientes de sistemas de saúde pública.

A integração dessas fontes cria a base empírica sobre a qual modelos mais robustos e clinicamente relevantes podem ser construídos. Um modelo que combina dados clínicos, laboratoriais e de imagem tem, em princípio, muito mais informação disponível para orientar uma predição do que

qualquer uma dessas fontes isoladamente. Mas essa integração impõe desafios técnicos consideráveis, incluindo a padronização de formatos, a interoperabilidade entre sistemas e a gestão da privacidade de dados de naturezas distintas. Esses desafios são tão relevantes quanto os algoritmos em si para determinar o sucesso de aplicações de IA em contextos clínicos reais.

4 A NATUREZA DO DIAGNÓSTICO NA ERA DOS DADOS

A incorporação da IA no processo diagnóstico não é uma simples adição de ferramentas ao arsenal clínico. É uma reconfiguração da lógica que organiza o próprio ato de diagnosticar. Para compreender essa transformação, é necessário examinar o que o diagnóstico é, epistemologicamente, antes de investigar o que ele se torna com a introdução da IA.

O diagnóstico clínico sempre foi incerto. Essa incerteza não é acidental, mas constitutiva, decorrente da complexidade biológica, da variabilidade individual, da incompletude das informações disponíveis e das limitações do conhecimento médico em determinado momento histórico. O que varia entre diferentes paradigmas não é a existência da incerteza, mas as formas de sua gestão.

Na prática clínica tradicional, essa incerteza era gerenciada de maneira implícita, mediada pelo julgamento e pela experiência do profissional. Com a introdução da IA, a incerteza passa a ser explícita, quantificada e expressa em termos probabilísticos. Essa transformação não apenas altera a forma de tomada de decisão, mas também reconfigura a maneira como o conhecimento clínico é produzido, interpretado e aplicado, trazendo implicações que precisam ser examinadas com cautela.

4.1 DO DIAGNÓSTICO INTERPRETATIVO AO DIAGNÓSTICO PROBABILÍSTICO

No modelo tradicional, o diagnóstico tende a ser apresentado como uma conclusão: o paciente tem ou não determinada condição. Essa aparência de certeza, contudo, encobre um processo essencialmente probabilístico. O clínico trabalha com hipóteses, pesos relativos entre alternativas e exclusões progressivas, ainda que raramente formalize esse raciocínio em termos estatísticos explícitos. A certeza apresentada ao paciente constitui, na maior parte das vezes, uma simplificação comunicativa de um processo internamente mais incerto.

A IA torna esse processo visível. Ao indicar que há 87% de probabilidade de uma determinada lesão estar presente em uma radiografia, o algoritmo não inventa uma linguagem nova. Ele formaliza o que o diagnóstico sempre foi, embora raramente tenha sido explicitado. Essa explicitação tem efeitos positivos: permite decisões mais informadas, facilita a comunicação do grau de certeza e torna o raciocínio diagnóstico mais auditável. Contudo, também exige do profissional uma competência que nem sempre é desenvolvida nas graduações em saúde: a capacidade de interpretar e comunicar incerteza de forma rigorosa, sem transformar probabilidades em certezas por conveniência comunicativa.

A formação de profissionais aptos a lidar com o diagnóstico probabilístico torna-se uma demanda central para os currículos em saúde. Isso envolve não apenas o desenvolvimento de competências estatísticas, mas também habilidades de comunicar ao paciente o grau de incerteza de forma compreensível, sem gerar interpretações equivocadas ou ansiedade desnecessária.

4.2 A TENSÃO ENTRE INDIVIDUALIDADE CLÍNICA E GENERALIZAÇÃO ALGORÍTMICA

Os algoritmos aprendem com populações. Sua lógica é estatística, orientada por padrões majoritários nos dados de treinamento. O paciente individual, entretanto, pode não se encaixar nesses padrões, seja por condições raras, por características demográficas sub-representadas nas bases de dados ou por contextos clínicos incomuns. Essa tensão entre a singularidade do paciente e a lógica generalizante do algoritmo é estrutural, não acidental.

Essa tensão manifesta-se de forma concreta na ocorrência de vieses algorítmicos, que podem resultar em erros diagnósticos sistemáticos que afetam desproporcionalmente grupos específicos. Evidências têm demonstrado que modelos de IA podem apresentar menor acurácia em populações historicamente sub-representadas nos dados de treinamento, incluindo grupos étnicos minoritários, mulheres, pacientes idosos com múltiplas comorbidades, e indivíduos em contexto socioeconômicos vulneráveis. Esses padrões refletem não apenas limitações técnicas, mas também desigualdade estruturais presentes na produção e disponibilidade de dados de saúde.

A resposta a esse problema não é simples. Ampliar as bases de dados de treinamento para incluir maior diversidade é necessária, mas não suficiente: é preciso também desenvolver métricas de desempenho que avaliem explicitamente a equidade dos modelos entre diferentes subgrupos, bem como estabelecer requisitos regulatórios que exijam tal avaliação antes da implementação clínica.

4.3 DIAGNÓSTICO COMO PROCESSO DINÂMICO E CONTÍNUO

Historicamente, o diagnóstico se caracteriza como um ato pontual: ocorre em uma consulta, a partir dos dados disponíveis naquele momento, e resulta em uma conclusão que orienta a conduta subsequente. Essa temporalidade discreta da interpretação clínica reflete as limitações tecnológicas de cada época. Com dispositivos de monitoramento contínuo, prontuários que se atualizam em tempo real e algoritmos capazes de processar fluxos de dados em tempo real, o diagnóstico tende a se tornar um processo: uma hipótese que se refina continuamente à medida que novas informações chegam.

Esse modelo apresenta grande potencial para a detecção precoce de deterioração clínica, especialmente em contextos de cuidado intensivo e monitoramento de doenças crônicas. Um algoritmo que analisa continuamente os sinais vitais de um paciente hospitalizado pode identificar padrões de deterioração horas antes de qualquer sinal clínico manifesto, permitindo intervenção precoce. De forma semelhante, no manejo de condições como diabetes e insuficiência cardíaca, o monitoramento contínuo

assistido por IA pode contribuir para a identificação precoce de descompensações, reduzindo a ocorrência de eventos agudos.

Mas esse modelo também impõe novos desafios. Como definir o momento de agir quando o diagnóstico é uma probabilidade em constante atualização? Quem monitora os alertas gerados por sistemas automatizados, e como evitar que o volume de alertas produza fadiga e inatenção? Como distribuir a responsabilidade por um diagnóstico que emerge de um processo contínuo e distribuído ao longo do tempo, e não de um ato deliberado de um profissional em um momento específico? Essas questões permanecem em aberto e constituem um campo central para futuras investigações.

4.4 A OPACIDADE ALGORÍTMICA E O DESAFIO DA EXPLICABILIDADE

O diagnóstico clínico sempre pressupõe a capacidade de justificar: o profissional deve ser capaz de explicar ao paciente, ao colega e, se necessário, ao comitê de auditoria, por que chegou a determinada conclusão. Essa capacidade de justificação é o que permite o controle de qualidade, a comunicação com o paciente, o aprendizado a partir de erros e a responsabilização em casos de dano. Quando parte da conclusão diagnóstica deriva de um modelo de *deep learning* com dezenas de camadas e milhões de parâmetros, essa justificativa se torna difícil ou mesmo impossível de formular em termos clínicos compreensíveis.

O problema da opacidade algorítmica, frequentemente referido como o problema da 'caixa-preta', não é apenas técnico: é ético. Ele toca em princípios fundamentais da relação terapêutica, como a transparência e o respeito pela autonomia do paciente. Um paciente que recebe um diagnóstico parcialmente baseado em um algoritmo tem o direito de saber disso, e de receber uma explicação sobre como o algoritmo chegou a essa conclusão. Se essa explicação não é possível, algo importante sobre a ética do processo diagnóstico foi comprometido.

O campo da IA explicável (Explainable AI, denominada pela sigla XAI) busca abordar esse problema, desenvolvendo ferramentas que permitam identificar quais variáveis ou regiões de uma imagem influenciaram determinada predição. Essas ferramentas aumentam a transparência dos modelos, mas ainda têm limitações: fornecem apenas uma representação simplificada do funcionamento interno do algoritmo, e não uma explicação completa. Assim, a explicabilidade ainda não é plenamente alcançada e continua sendo um desafio no uso da IA na prática clínica.

4.5 RECONFIGURAÇÃO DA AUTORIDADE NO PROCESSO DIAGNÓSTICO

A presença da IA altera, inevitavelmente, a dinâmica de autoridade no processo diagnóstico. Quando um sistema algorítmico sugere um diagnóstico diferente do que o clínico formulou, surge um conflito que não tem resolução automaticamente clara. Seguir a máquina é uma possibilidade; ignorá-la também. Nenhuma das duas alternativas é isenta de risco.

Na prática, a introdução de recomendações algorítmicas tende a influenciar o processo decisório mesmo quando o profissional acredita estar exercendo um julgamento independente. Evidências da psicologia da decisão indicam que sugestões externas, ainda que explicitamente não vinculadas, podem influenciar significativamente as escolhas humanas. Nesse contexto, o clínico que utiliza sistema de apoio à decisão não está, portanto, tomando decisões completamente independentes: está tomando decisões em interação com o sistema, o que tem implicações tanto para a qualidade das decisões quanto para a responsabilidade sobre elas.

4.6 IMPLICAÇÕES PARA A PRÁTICA ODONTOLÓGICA E ÁREAS AFINS

A odontologia oferece um contexto particularmente rico para examinar essas transformações. Trata-se de uma área intensamente visual, com ampla disponibilidade de dados radiográficos padronizados e alta regularidade nos protocolos de aquisição de imagens. Isso a torna terreno fértil para algoritmos de análise de imagem, e os resultados publicados na literatura são expressivos: modelos de deep learning alcançam acurácia elevada na detecção de cáries, lesões periapicais e alterações periodontais, frequentemente identificando alterações em estágios iniciais que passariam despercebidas em avaliação convencional.

Por outro lado, a tomada de decisão terapêutica em odontologia envolve dimensões que os algoritmos atuais não capturam adequadamente. As preferências estéticas do paciente, as condições sistêmicas que influenciam a indicação de determinados procedimentos, o contexto econômico que define o que é viável, o grau de ansiedade e cooperação do paciente: todos esses fatores são essenciais para a tomada de decisão clínica, e nenhum deles está capturado nas imagens radiográficas que os algoritmos analisam. O diagnóstico assistido por IA na odontologia é mais sensível, mas não necessariamente mais abrangente: exige, tanto quanto o diagnóstico convencional, a integração pelo julgamento clínico do profissional, que permanece insubstituível na conduta do cuidado.

4.7 O DIAGNÓSTICO COMO CONSTRUÇÃO COLABORATIVA HUMANO-MÁQUINA

A imagem mais adequada para descrever o diagnóstico na era da IA não é a de uma substituição, o algoritmo no lugar do clínico, mas a de uma colaboração assimétrica, na qual cada agente contribui com o que faz melhor. A máquina processa volume, detecta padrões em alta velocidade e quantifica incerteza de forma explícita. O clínico contextualiza, integra, pondera aspectos que os dados não capturam e, em última instância, decide.

Essa colaboração, contudo, não é espontânea. Ela precisa ser construída deliberadamente, por meio de interfaces bem projetadas, de fluxos de trabalho cuidadosamente redesenhados e, sobretudo, de profissionais adequadamente formados para interagir com sistemas de IA de forma crítica e informada. Sem esse investimento, a colaboração tende a degenerar em dois extremos igualmente

problemáticos: a submissão acrítica às recomendações do algoritmo ou a rejeição sistemática motivada por desconfiança ou falta de familiaridade com a tecnologia.

5 RACIOCÍNIO CLÍNICO E INTELIGÊNCIA ARTIFICIAL: CONVERGÊNCIAS E TENSÕES

O raciocínio clínico é caracterizado na literatura de cognição médica, como um processo dual. De um lado, opera o raciocínio analítico, lento, deliberado, baseado na elaboração de hipóteses e na avaliação sistemática de evidências. De outro, opera o reconhecimento de padrões, rápido, intuitivo, fundamentado na experiência acumulada ao longo de anos de prática. Profissionais experientes transitam entre esses dois modos de forma fluida, frequentemente sem consciência explícita de qual predomina em cada situação.

Essa estrutura dual tem implicações importantes para compreender como a IA se relaciona com o raciocínio clínico. Os sistemas de IA baseados em deep learning operam predominantemente na lógica do reconhecimento de padrões, identificando estruturas nos dados que se associam a determinados desfechos. Funcionalmente, o que o algoritmo realiza é comparável ao que o clínico experiente faz ao deparar-se com um caso e reconhecer de imediato um padrão já consolidado em sua experiência. A diferença é de escala: o algoritmo foi treinado com muito mais exemplos do que qualquer clínico humano poderia acumular em uma carreira inteira.

O clínico experiente que reconhece um padrão o faz dentro de um contexto rico: ele sabe quem é o paciente, conhece sua história, compreende as implicações clínicas do diagnóstico e é capaz de articular por que aquele padrão é relevante. O algoritmo identifica o padrão, mas não incorpora esse contexto ao seu processamento. Essa diferença, tem consequências clínicas diretas, especialmente nos casos em que o contexto é o que diferencia uma decisão adequada de uma decisão inadequada.

A convergência entre raciocínio clínico e raciocínio algorítmico não é, portanto, automática. Ela requer a construção de modelos de interação que preservem o contexto, que garantam que o profissional compreenda o que o algoritmo oferece e que evitem que a velocidade da resposta algorítmica comprometa a reflexão analítica que muitas situações clínicas exigem.

6 HEURÍSTICAS, VIESES COGNITIVOS E O PAPEL DA IA

O raciocínio clínico humano é eficiente, mas não é infalível. As heurísticas são atalhos mentais que permitem decidir rapidamente em contextos de alta complexidade e pressão de tempo. Elas são funcionais na maior parte das situações, pois permitem ao clínico experiente chegar a conclusões razoáveis sem processar exaustivamente toda a informação disponível. Mas produzem erros sistemáticos quando aplicadas em contextos para os quais não foram desenvolvidas, ou quando as características do caso diferem das regularidades que sustentam seu funcionamento.

Entre os vieses mais documentados no raciocínio diagnóstico, destacam-se a ancoragem, que leva o clínico a supervalorizar a primeira hipótese formulada e resistir à revisão mesmo diante de evidências contrárias; o viés de disponibilidade, que leva a superestimar a probabilidade de diagnósticos recentemente encontrados ou particularmente memoráveis; o fechamento prematuro, que interrompe o processo diagnóstico antes de todas as hipóteses relevantes serem consideradas; e o viés de confirmação, que leva o clínico a buscar informações que confirmem sua hipótese inicial e a desvalorizar as que a contradizem.

A IA tem sido apresentada, com frequência, como uma solução para esses vieses. Algoritmos não estão sujeitos à fadiga, à pressão temporal, à superestimação de diagnósticos recentes ou à ancoragem em hipóteses iniciais da mesma forma que o raciocínio humano. Essa vantagem é real em certas condições. Mas a visão de que a IA simplesmente elimina os vieses do diagnóstico é uma simplificação que não encontra suporte na evidência disponível.

Os algoritmos de aprendizado de máquina carregam vieses próprios, que emergem diretamente das características dos dados utilizados em seu treinamento. Se os dados de treinamento refletem práticas clínicas historicamente enviesadas, como o subdiagnóstico de determinadas condições em grupos específicos, o modelo aprenderá e reproduzirá esses vieses em escala. Se determinada população está sub-representada nos dados de treinamento, o modelo terá desempenho inferior para essa população. Se os rótulos utilizados no treinamento foram produzidos por especialistas com vieses próprios, esses vieses serão incorporados ao modelo.

Além disso, a interação entre humanos e sistemas de IA introduz formas específicas de viés que não existiam antes. O viés de automação, a tendência de aceitar acriticamente as recomendações de um sistema percebido como mais confiável ou mais objetivo do que o julgamento humano, é particularmente problemático em contextos clínicos. Um profissional que aceita a sugestão algorítmica sem avaliação crítica pode cometer erros sistemáticos de forma muito mais eficiente do que cometeria agindo sozinho. Nesse sentido, a IA amplifica tanto o acerto quanto o erro.

7 MODELOS DE TOMADA DE DECISÃO HÍBRIDOS

A resposta mais promissora às limitações de ambos os agentes, humano e algorítmico, é o modelo híbrido de tomada de decisão. Nesse modelo, a IA atua como sistema de apoio, fornecendo análises, probabilidades e recomendações que são subsequentemente avaliadas pelo profissional de saúde, que mantém a autoridade e a responsabilidade pela decisão final. A lógica direta é: combinar a capacidade analítica dos algoritmos com a compreensão contextual e o julgamento ético do clínico.

Na prática, entretanto, a implementação de modelos híbridos eficientes é mais complexa do que a descrição conceitual sugere. Para que a colaboração funcione, o profissional precisa compreender o que o algoritmo oferece e o que ele não oferece. Isso inclui a capacidade de interpretar probabilidades

e intervalos de confiança, identificar as condições em que o modelo tende a errar e reconhecer quando o caso em questão está fora do domínio para o qual o modelo foi desenvolvido. Sem essa compreensão, a colaboração degenera em submissão acrítica ou em rejeição sistemática, ambas clinicamente inadequadas.

A construção de modelos híbridos eficientes não é apenas uma questão tecnológica. Envolve a reestruturação cuidadosa dos fluxos de trabalho clínicos, para que a interação com o sistema de IA seja fluida e não adicione carga cognitiva desnecessária. Envolve o desenvolvimento de interfaces que apresentem as recomendações algorítmicas de forma compreensível e contextualizada, sem sobrecarregar o profissional com informações irrelevantes. E envolve, acima de tudo, a formação de profissionais que compreendam o que significa trabalhar em conjunto com sistemas de IA.

8 EXPLICABILIDADE, TRANSPARÊNCIA E CONFIANÇA NO DIAGNÓSTICO ASSISTIDO POR IA

A questão da explicabilidade merece atenção específica, porque toca em um princípio fundamental da prática clínica: a responsabilidade de justificar. Em medicina, não basta chegar ao diagnóstico correto. É necessário ser capaz de articular o raciocínio que levou até ele. Essa capacidade de justificação é o que permite o controle de qualidade, a comunicação com o paciente, a formação de novos profissionais e o aprendizado coletivo a partir de erros.

Quando decisões diagnósticas são influenciadas por modelos de deep learning com milhões de parâmetros internos, essa justificação se torna tecnicamente difícil. O campo da IA explicável desenvolveu um conjunto de ferramentas destinadas a tornar os processos decisórios dos algoritmos mais compreensíveis. Entre elas, destacam-se os métodos que identificam quais regiões de uma imagem foram mais relevantes para uma determinada classificação, e os métodos que identificam quais variáveis clínicas mais contribuíram para uma predição. Essas ferramentas não revelam o processamento interno completo do modelo, mas oferecem uma aproximação com utilidade clínica demonstrável.

A confiança no diagnóstico assistido por IA não pode ser construída apenas sobre o desempenho dos modelos. Ela depende de transparência sobre como foram desenvolvidos, em quais dados foram treinados, para quais populações foram validados e quais são seus limites conhecidos. Depende também da informação clara sobre as situações em que o modelo provavelmente funciona bem e aquelas em que provavelmente não funciona. Sem essa informação, a adoção da IA na clínica equivale a uma decisão baseada em confiança irrefletida, fundamentada apenas na reputação do desenvolvedor ou no desempenho observado em condições ideais que podem não se reproduzir na prática.

9 CONFLITOS DECISÓRIOS: QUANDO O ALGORITMO DIVERGE DO CLÍNICO

O cenário mais revelador e mais desconfortável da clínica assistida por IA é aquele em que o algoritmo sugere algo diferente do que o profissional formulou. Nesses momentos, a pergunta abstrata sobre quem decide se torna urgente e concreta. O profissional está diante de duas fontes de informação em conflito: seu próprio raciocínio clínico, construído sobre experiência, contexto e julgamento, e a recomendação do sistema, construída sobre padrões estatísticos em grandes bases de dados.

Em termos jurídicos, a responsabilidade permanece com o profissional. Mas em termos práticos, a dinâmica é mais complexa. Ignorar a sugestão da IA, quando o desfecho for adverso, pode ser interpretado retrospectivamente como recusa injustificada de uma ferramenta disponível. Segui-la acriticamente, quando ela estiver errada, também não exime o profissional de responsabilidade por uma decisão que deveria ter sido submetida a avaliação crítica. O espaço de manobra existe, mas exige do profissional um nível de preparo analítico que nem sempre é viável no tempo real da consulta.

A resolução responsável desses conflitos exige mais do que boa vontade individual. Exige protocolos institucionais que orientem como proceder quando o julgamento clínico e a recomendação algorítmica divergem. Exige documentação adequada do raciocínio que levou à decisão final. E exige um ambiente institucional que valorize o pensamento crítico em relação à tecnologia, em vez de gerar incentivos implícitos para a aceitação automática das recomendações do sistema.

10 IMPACTOS NA FORMAÇÃO DO PROFISSIONAL DE SAÚDE

A incorporação da IA na prática clínica implica em uma questão direta e urgente para as instituições formadoras: o que significa preparar um profissional de saúde para um ambiente em que algoritmos participam ativamente do diagnóstico? A resposta não pode ser reduzida à adição de uma disciplina de 'introdução à IA' ao final do currículo. Exige uma revisão mais profunda das competências centrais que se espera que o profissional desenvolva ao longo da formação.

O profissional contemporâneo precisa compreender o suficiente sobre aprendizado de máquina para avaliar criticamente um sistema de apoio à decisão: reconhecer seus limites, identificar quando seus resultados devem ser questionados e comunicar ao paciente, de forma honesta e compreensível, o papel que a tecnologia desempenhou no diagnóstico. Isso não exige que todos os clínicos se tornem especialistas em ciência de dados. Exige que nenhum clínico utilize a IA como uma caixa mágica cujos resultados são aceitos por padrão.

Simultaneamente, as competências humanas ganham novo relevo em um contexto de crescente automação da análise técnica. A empatia, a escuta ativa, a capacidade de comunicar incerteza de forma compreensível, o julgamento ético diante de situações em que os dados não oferecem resposta clara: essas competências são precisamente as que os algoritmos ainda não conseguem reproduzir, e são elas que representam o diferencial do profissional em relação aos sistemas automatizados. Formar para

essas competências é, portanto, uma resposta estratégica à transformação tecnológica, não uma fuga dela.

A formação continuada torna-se, nesse contexto, indispensável. A evolução rápida das ferramentas de IA disponíveis na prática clínica significa que a formação inicial, por mais completa que seja, será rapidamente ultrapassada pelos desenvolvimentos tecnológicos. Modelos de educação permanente que integrem a atualização tecnológica ao desenvolvimento das competências clínicas tradicionais são uma necessidade estrutural, não um complemento opcional.

10.1 ESGOTAMENTO PROFISSIONAL E INTELIGÊNCIA ARTIFICIAL: PROMESSA E RISCO

A incorporação da IA na prática clínica é frequentemente apresentada nos discursos institucionais e na literatura de implementação, como uma resposta ao esgotamento profissional. O argumento tem base empírica parcial: ferramentas de transcrição automática, geração assistida de registros clínicos e triagem algorítmica de resultados têm demonstrado redução mensurável no tempo dedicado a tarefas administrativas, que figuram entre os principais fatores associados ao burnout em saúde. A carga documental imposta pelos prontuários eletrônicos, em particular, consome uma parcela significativa da jornada de trabalho dos profissionais, e sistemas baseados em IA têm mostrado potencial real para reduzir essa demanda.

Entretanto, a relação entre IA e esgotamento profissional é mais complexa do que essa narrativa sugere. A introdução de novas tecnologias no ambiente clínico raramente reduz a carga de trabalho de forma simples e direta, ela a redistribui. O tempo liberado pela automação de tarefas rotineiras tende a ser preenchido por novas demandas: monitoramento de alertas gerados pelos sistemas, verificação de saídas algorítmicas, documentação adicional para justificar decisões que divergem das recomendações do sistema. Em contextos em que os sistemas de apoio à decisão geram volumes elevados de notificações, o fenômeno da fadiga de alertas contribui para uma forma específica de sobrecarga cognitiva que pode agravar, e não aliviar, o estado de esgotamento do profissional.

Há ainda uma dimensão menos visível desse problema: o impacto da IA sobre o senso de competência e autonomia do profissional. Quando sistemas algorítmicos assumem tarefas antes consideradas centrais ao exercício clínico, parte dos profissionais experimenta uma redefinição de seu papel que pode ser vivenciada como esvaziamento de identidade profissional. Essa experiência não é universal, mas está documentada, e tem consequências sobre o engajamento com o trabalho e sobre a satisfação profissional a longo prazo. Implementações mal conduzidas, que introduzem a tecnologia sem preparação adequada ou sem envolvimento dos profissionais no processo, tendem a amplificar esse efeito.

O ponto central, portanto, não é se a IA pode ou não contribuir para o bem-estar dos profissionais de saúde, ela pode, em condições adequadas. É que essa contribuição não é automática.

Ela depende de como os sistemas são desenhados, de como são implementados e de como o tempo e a atenção recuperados pela automação são efetivamente reinvestidos. Tratar a IA como solução inerente ao burnout, sem atenção a esses fatores, é substituir um problema por outro.

11 APLICAÇÕES DA INTELIGÊNCIA ARTIFICIAL NO DIAGNÓSTICO CLÍNICO CONTEMPORÂNEO

As aplicações clínicas da IA não são hipotéticas nem futuristas. Elas já existem, já estão em uso em diferentes partes do mundo e acumulam evidências na literatura científica. A pergunta relevante não é mais se a IA funcionará no diagnóstico, mas como integrá-la de forma que seus benefícios se distribuam equitativamente e seus riscos sejam gerenciados com responsabilidade. Isso exige compreender, com algum grau de especificidade, onde e como a IA está sendo aplicada e com que resultados.

As aplicações variam significativamente conforme o tipo de dado analisado, o nível de integração com o fluxo assistencial e o grau de autonomia conferido ao sistema. Em alguns contextos, a IA atua de forma complementar, sugerindo hipóteses diagnósticas que o profissional avalia e aceita ou rejeita. Em outros, assume papel mais ativo, como em sistemas de triagem automatizada que priorizam casos de maior urgência antes de qualquer avaliação humana. Compreender essas variações é fundamental para avaliar as implicações éticas e práticas de cada aplicação.

12 DIAGNÓSTICO POR IMAGEM: PRECISÃO, ESCALA E SENSIBILIDADE

O diagnóstico por imagem é, atualmente, o campo mais maduro de aplicação clínica da IA. A combinação entre ampla disponibilidade de dados visuais rotulados e a capacidade das redes neurais convolucionais para análise de padrões em imagens produziu algoritmos com desempenho notável em tarefas específicas. Detecção de nódulos pulmonares em tomografias, identificação de retinopatia diabética em imagens de fundo de olho, classificação de lesões dermatológicas em fotografias, reconhecimento de fraturas em radiografias: em todas essas tarefas, estudos publicados demonstram desempenho comparável ou superior ao de especialistas humanos em condições controladas.

A vantagem da IA nesses contextos não é apenas a precisão. É também consistência e escala. Um algoritmo analisa cada imagem aplicando os mesmos critérios, sem variação relacionada ao cansaço, à hora do dia ou ao número de casos já avaliados. Isso tem implicações diretas para a qualidade do diagnóstico em sistemas de saúde de alta demanda, onde a fadiga dos especialistas é um fator documentado de variabilidade diagnóstica.

A escala é igualmente relevante. Em programas de rastreamento populacional, como o rastreamento mamográfico para câncer de mama ou o rastreamento de retinopatia diabética, a quantidade de imagens a serem avaliadas é muito superior à capacidade de análise disponível. Sistemas

de IA podem processar esse volume de forma eficiente, identificando casos prioritários para avaliação especializada e permitindo que os especialistas humanos dediquem seu tempo aos casos mais complexos.

Entretanto, o desempenho desses sistemas em condições controladas de pesquisa não se traduz automaticamente para o desempenho no mundo real. Diferenças na qualidade das imagens, nos equipamentos utilizados, nos protocolos de aquisição e nas características populacionais podem degradar significativamente o desempenho dos modelos. Isso reforça a necessidade de validação externa rigorosa antes de qualquer implementação clínica.

13 INTELIGÊNCIA ARTIFICIAL NA ODONTOLOGIA: DIAGNÓSTICO ASSISTIDO E PLANEJAMENTO CLÍNICO

Na odontologia, a IA encontrou terreno especialmente fértil. A prática odontológica é intensamente visual e produz grandes volumes de imagens padronizadas, incluindo radiografias periapicais, panorâmicas e tomografias de feixe cônico, o que facilita a construção de bases de dados adequadas para o treinamento de modelos de aprendizado profundo. A relativa padronização dos protocolos de aquisição de imagens odontológicas também reduz a variabilidade que, em outras especialidades, dificulta a generalização dos modelos.

Os resultados reportados na literatura são consistentes: algoritmos de deep learning alcançam acurácia elevada na identificação de cáries interproximais, lesões periapicais, alterações periodontais com perda óssea mensurável e lesões de bifurcação. Em muitos estudos, esses sistemas detectam alterações em estágios precoces que passariam despercebidas em avaliação convencional, o que representa um avanço significativo para o objetivo de intervenção precoce.

No planejamento terapêutico, especialmente em implantodontia e ortodontia, a análise automatizada de estruturas ósseas e dentárias em tomografias computadorizadas permite simulações mais precisas do posicionamento de implantes e da movimentação dentária. Isso reduz a dependência exclusiva do julgamento visual do profissional em tarefas que envolvem medições precisas e avaliação de estruturas tridimensionais complexas.

Entretanto, como em outras especialidades, o uso da IA na odontologia não elimina a necessidade de julgamento clínico. A decisão de tratar ou monitorar uma lesão cáriosa inicial, por exemplo, depende não apenas da presença da lesão, mas do histórico de cárie do paciente, de seu risco individual, de suas preferências e de sua capacidade de aderir a medidas preventivas. Nenhum desses fatores está capturado na imagem radiográfica. O diagnóstico assistido por IA na odontologia é um componente valioso do processo diagnóstico, não um substituto dele.

14 SISTEMAS DE APOIO À DECISÃO CLÍNICA (CDSS): INTEGRAÇÃO ENTRE DADOS E CONDUTA

Os sistemas de apoio à decisão clínica (Clinical Decision Support Systems, CDSS) representam a modalidade de integração mais abrangente da IA no fluxo assistencial. Diferentemente de ferramentas de análise de imagem, que operam em um domínio específico, os CDSS buscam integrar múltiplas fontes de dados clínicos do paciente, incluindo histórico, resultados laboratoriais, medicações em uso e sinais vitais, para oferecer alertas, sugestões diagnósticas e recomendações terapêuticas em tempo real, integradas ao prontuário eletrônico.

A eficácia dos CDSS depende de dois fatores frequentemente subestimados no desenvolvimento tecnológico: a qualidade dos dados disponíveis e a adequação da interface ao fluxo de trabalho real. Sistemas que geram alertas excessivos ou irrelevantes são sistematicamente ignorados pelos profissionais, fenômeno conhecido como *alert fatigue*, que pode ser ainda mais prejudicial do que a ausência de alertas, pois cria um ambiente em que os alertas genuinamente importantes são também desconsiderados.

A experiência acumulada com CDSS ao longo de décadas demonstra que o sucesso da implementação depende muito mais de fatores organizacionais e de design do que de sofisticação técnica. Sistemas percebidos pelos profissionais como aliados, que reduzem sua carga cognitiva sem criar obstáculos ao fluxo de trabalho, tendem a ser adotados e utilizados de forma eficaz. Sistemas percebidos como interferências ou como instrumentos de controle administrativo tendem a ser contornados ou ignorados, independentemente de sua qualidade técnica.

15 BIOMARCADORES, DADOS MULTIMODAIS E DIAGNÓSTICO INTEGRADO

Uma das perspectivas mais promissoras da IA no diagnóstico em saúde é a capacidade de integrar dados de naturezas profundamente distintas em modelos preditivos unificados. O raciocínio clínico humano processa diferentes tipos de informação de forma relativamente sequencial: examina os resultados laboratoriais, depois analisa as imagens, depois considera o histórico clínico. Os algoritmos de aprendizado de máquina podem, em princípio, processar simultaneamente todas essas dimensões, identificando interações e padrões que seriam invisíveis em análises isoladas.

Modelos multimodais que integram dados de imagem, dados laboratoriais, dados genômicos e informações clínicas estruturadas têm demonstrado, em contextos específicos, desempenho superior aos modelos unimodais. Na oncologia, por exemplo, a integração de dados de imagem radiológica com dados de sequenciamento genômico permite estratificações prognósticas muito mais precisas do que qualquer uma das fontes isoladamente.

No contexto odontológico e sistêmico, essa abordagem multimodal abre perspectivas para uma compreensão mais aprofundada das relações entre saúde bucal e condições sistêmicas como diabetes,

doenças cardiovasculares e condições gestacionais. A integração de dados odontológicos com dados de prontuário sistêmico pode revelar associações que a análise isolada de cada especialidade não seria capaz de identificar. Os desafios metodológicos são consideráveis, especialmente na troca de informações entre sistemas e de interpretabilidade dos modelos, mas a direção é clinicamente promissora.

16 IA GENERATIVA E PRODUÇÃO DE CONHECIMENTO CLÍNICO

A evolução mais recente da IA introduziu uma categoria de sistemas com características qualitativamente distintas dos modelos de classificação e predição: a inteligência artificial generativa. Enquanto os modelos tradicionais analisam dados e produzem classificações ou probabilidades, os sistemas generativos produzem conteúdo novo, como textos, relatórios, sínteses e simulações, a partir de padrões aprendidos, frequentemente condicionados por instruções em linguagem natural.

No contexto clínico, isso se traduz na capacidade de gerar rascunhos de laudos a partir de dados estruturados, sintetizar informações relevantes de prontuários extensos, identificar e sumarizar literatura científica pertinente a um caso específico, ou auxiliar na formulação de hipóteses diagnósticas para casos complexos. Em ambientes de alta demanda, onde a carga de documentação é frequentemente apontada como fator de esgotamento profissional, a IA generativa pode oferecer alívio significativo.

Entretanto, os sistemas generativos têm uma característica que os torna particularmente problemáticos em contextos de saúde: eles produzem texto fluente e coerente mesmo quando o conteúdo é factualmente incorreto. Esse fenômeno, conhecido como alucinação, é inerente à forma como esses modelos funcionam e não é completamente eliminado por quaisquer das técnicas de treinamento disponíveis atualmente. Em saúde, onde a precisão da informação é crítica e erros podem ter consequências graves, a supervisão humana de todo conteúdo produzido por IA generativa não é opcional: é uma exigência ética inegociável.

17 LIMITAÇÕES TÉCNICAS E RISCOS NA APLICAÇÃO CLÍNICA

As limitações da IA no diagnóstico em saúde são tão reais quanto suas capacidades, e merecem atenção proporcional. A primeira e mais fundamental é a dependência de dados de qualidade. Sem dados em volume suficiente, com representatividade adequada e com qualidade técnica satisfatória, os modelos não funcionam. Em muitos contextos clínicos, especialmente em países de média e baixa renda, em especialidades com menor produção de dados digitalizados e em populações historicamente excluídas dos sistemas de saúde, essa dependência representa uma barreira concreta à implementação.

A segunda limitação é a restrição de domínio. Os modelos de IA aprendem a resolver problemas dentro do domínio definido pelos seus dados de treinamento. Quando encontram situações fora desse

domínio, seja uma doença emergente nunca antes vista, uma apresentação clínica atípica ou um paciente com características muito diferentes das da população de treinamento, sua capacidade de generalização é limitada. Ao contrário do clínico humano, que pode raciocinar sobre situações novas a partir de princípios gerais, o algoritmo não tem como saber o que não aprendeu.

A terceira limitação, frequentemente subestimada, é a natureza sistemática dos erros algorítmicos. Quando um clínico comete um erro diagnóstico, esse erro é individual e circunstancial. Quando um algoritmo comete um erro, ele o comete de forma consistente em todos os casos que apresentam as mesmas características. Em um sistema de saúde que processa milhares de exames por dia com auxílio de IA, um erro sistemático pode afetar um número muito grande de pacientes antes de ser detectado e corrigido. Isso coloca exigências muito maiores de monitoramento e de mecanismos de detecção de erros do que as que existem para a prática clínica convencional.

18 INTEGRAÇÃO COM O FLUXO CLÍNICO E DESAFIOS DE IMPLEMENTAÇÃO

A distância entre uma tecnologia que demonstra desempenho impressionante em estudos publicados e uma tecnologia que funciona de forma eficaz no contexto clínico real é frequentemente maior do que os desenvolvedores antecipam. A implementação bem-sucedida da IA em saúde exige atenção a variáveis que vão muito além do desempenho técnico do algoritmo: interoperabilidade com sistemas de informação já existentes, conformidade com regulações de privacidade e segurança de dados, adequação das interfaces às condições reais de trabalho dos profissionais e gestão cuidadosa do processo de mudança organizacional.

A resistência dos profissionais à adoção de novas tecnologias não deve ser interpretada automaticamente como conservadorismo ou falta de abertura à inovação. Em muitos casos, ela reflete experiências prévias com implementações que não atingiram os resultados esperados, bem como preocupações bem fundamentadas sobre como a tecnologia afetará a qualidade do cuidado, a distribuição de responsabilidades e as condições de trabalho. Nesse contexto, o envolvimento dos profissionais no processo de implementação desde as etapas iniciais, aliado a escuta de suas demandas e a adaptação das ferramentas ao contexto clínico real, constituem um elemento para o sucesso dessas iniciativas.

19 DIMENSÕES ÉTICAS DA INTELIGÊNCIA ARTIFICIAL NO DIAGNÓSTICO EM SAÚDE

A ética da Inteligência Artificial (IA) em saúde não constitui um elemento adicional ao desenvolvimento técnico, mas integra o próprio núcleo do problema. Questões éticas estão presentes desde as primeiras decisões sobre que dados coletar, que populações incluir, que objetivos otimizar e que riscos aceitar. Cada escolha no desenvolvimento de sistemas de IA em saúde envolve,

simultaneamente, dimensões técnicas e uma decisão moral, ainda que nem sempre seja reconhecida como tal.

O princípio da autonomia, central na bioética contemporânea, é tensionado em múltiplas direções pela presença da IA no processo diagnóstico. A autonomia do paciente, entendida como direito de participar ativamente das decisões sobre sua saúde com base em informações compreensíveis, pode ser comprometida quando parte do processo diagnóstico é conduzida por um algoritmo cujo funcionamento não é plenamente interpretável. A autonomia do profissional, relacionada a capacidade de exercer julgamento clínico fundamentado, é questionada quando sistemas algorítmicos operam com autoridade implícita que torna difícil ou socialmente custoso discordar de suas recomendações.

Os princípios da beneficência e da não maleficência exigem que os sistemas de IA sejam desenvolvidos e implementados com rigor metodológico suficiente para assegurar que seus benefícios superem seus riscos. Isso inclui a realização de validação externa em populações representativas, o monitoramento contínuo do desempenho após a implementação e a existência de mecanismos que permitam a suspensão ou revisão rápida do sistema quando problemas ou falhas são identificados. Exige também honestidade sobre as limitações dos sistemas, mesmo quando essa transparência contraria interesses institucionais ou comerciais.

A equidade representa um dos desafios éticos mais urgentes e mais sistematicamente negligenciados nesse campo. Embora a IA tenha o potencial de ampliar o acesso a diagnósticos de qualidade em contextos com escassez de especialistas, ela também pode, se desenvolvida sem atenção explícita à diversidade dos dados, reproduzir e amplificar desigualdades presentes nos conjuntos de treinamento. Garantir a distribuição equitativa dos benefícios da IA requer decisões deliberadas em todas as etapas desses sistemas, desde o desenvolvimento até a implementação clínica.

20 RESPONSABILIDADE E ACCOUNTABILITY NO DIAGNÓSTICO ASSISTIDO POR IA

A responsabilidade no diagnóstico assistido por IA é complexa porque envolve mais etapas e mais participantes do que o diagnóstico convencional. Quando um profissional chega a um diagnóstico errado usando apenas seu próprio julgamento, a responsabilidade é relativamente clara: ela recai sobre o profissional, modulada pelas circunstâncias do caso e pelos padrões da prática clínica vigente. Entretanto, quando o diagnóstico é influenciado por um sistema de IA, os responsáveis se multiplicam: o desenvolvedor do algoritmo, a empresa que o comercializou, a instituição que o implementou, o profissional que o utilizou e a empresa que forneceu os dados de treinamento.

Os marcos regulatórios atuais não foram projetados para lidar com esse modelo de responsabilidade distribuída. A tendência predominante, nos sistemas jurídicos atuais, é atribuir ao profissional a responsabilidade final por qualquer decisão clínica, independentemente do papel que a tecnologia desempenhou no processo. Isso cria uma distorção: o profissional é responsável por

decisões que não controlou completamente. A construção de marcos regulatórios mais adequados, que distribuam responsabilidades de forma proporcional ao papel de cada envolvido no processo, é uma necessidade urgente que ainda não foi devidamente atendida na maior parte dos sistemas jurídicos.

21 REGULAÇÃO E GOVERNANÇA DA INTELIGÊNCIA ARTIFICIAL EM SAÚDE

Os sistemas regulatórios da saúde foram estruturados para avaliar medicamentos e dispositivos médicos, produtos com características relativamente estáveis que podem ser testados em condições controladas e certificados em um único momento. Os sistemas baseados em IA apresentam características distintas: modelos de aprendizado de máquina podem ser atualizados continuamente com novos dados, seu desempenho pode variar em função das características dos dados de entrada e seus efeitos podem não se reproduzir de forma consistente em populações diferentes daquelas utilizadas em seu desenvolvimento.

Isso exige regulação adaptativa, capaz de acompanhar a evolução dos sistemas ao longo do tempo, e não apenas de certificar sua segurança em um momento inicial. Agentes reguladores, como a Food and Drug Administration (FDA) nos Estados Unidos, vem desenvolvendo frameworks para regulação contínua de dispositivos baseado em IA. Na Europa, o Artificial Intelligence Act estabelece requisitos específicos para sistemas classificados como de alto risco, incluindo aplicações em saúde. No contexto brasileiro, os avanços regulatórios têm ocorrido de forma mais gradual, o que evidencia a necessidade de consolidação de diretrizes mais robustas para garantir a segurança e a equidade na utilização dessas tecnologias.

A governança institucional desempenha papel igualmente relevante. Hospitais e clínicas que implementam sistemas de IA precisam estabelecer estruturas que assegurem avaliação crítica prévia à implementação, monitoramento contínuo do desempenho, mecanismos de notificações de falhas e definição clara de responsabilização. Essa governança não deve ser restrita às áreas de tecnologia da informação, mas deve envolver profissionais de saúde, gestores, comitês de ética e, sempre que possível, representantes dos pacientes, garantindo uma abordagem multidisciplinar e centrada no cuidado.

22 FORMAÇÃO PROFISSIONAL NA ERA DA INTELIGÊNCIA ARTIFICIAL

Atualizar a formação profissional em saúde para a era da IA não significa substituir o conteúdo clínico por conteúdo tecnológico. Significa integrá-los de forma que o profissional formado seja capaz de exercer julgamento clínico informado pela tecnologia, e não subjugado a ela. Essa distinção é fundamental: o objetivo não é criar profissionais que confiam na IA, mas profissionais que sabem quando e por que confiar na IA.

Na prática curricular, isso envolve incorporar estatística aplicada e fundamentos de aprendizado de máquina de forma integrada ao raciocínio clínico, e não como disciplinas tecnológicas isoladas. Também requer abordar as limitações dos algoritmos nos mesmos contextos em que se discutem as restrições de exames laboratoriais ou das evidências científicas. Significa criar oportunidades para que os estudantes interajam, em contextos supervisionados, com sistemas reais de apoio à decisão baseados em IA, desenvolvendo familiaridade crítica que permita avaliar esses sistemas de forma informada.

Os aspectos éticos da formação merece atenção particular. Os estudantes precisam desenvolver a capacidade de identificar quando um sistema de IA pode estar gerando resultados enviesados, quando sua aplicação pode comprometer a equidade do cuidado e quando a responsabilidade de questionar as recomendações da máquina é parte integrante do exercício profissional responsável. Essa capacidade não se desenvolve apenas com conhecimento técnico; depende de uma formação que valorize o pensamento crítico e incentive o questionamento de diferentes fontes de autoridade, inclusive a tecnológica.

23 O FUTURO DO DIAGNÓSTICO EM SAÚDE: INTEGRAÇÃO, PERSONALIZAÇÃO E PREVENÇÃO

As perspectivas futuras da IA no diagnóstico em saúde evidenciam uma integração crescente entre diferentes fontes de dados, sistemas tecnológicos e especialidades clínicas. O avanço da interoperabilidade entre sistemas de informação em saúde permitirá a construção de modelos preditivos mais abrangentes, capazes de considerar simultaneamente múltiplas dimensões da história clínica de cada paciente ao longo de sua vida.

A personalização do diagnóstico tende a se intensificar com o uso crescente de dados genômicos, de microbioma e de outros biomarcadores moleculares. A medicina de precisão, que busca adaptar diagnósticos e tratamentos às características moleculares específicas de cada paciente, depende essencialmente da capacidade de integrar e analisar esses dados em escala, o que só é possível com o auxílio de sistemas de IA. Nesse cenário, o diagnóstico deixa de ser um evento pontual e passa a constituir um processo contínuo de refinamento da compreensão do estado de saúde individual.

A expansão da telemedicina e das ferramentas de saúde digital amplia o alcance geográfico dessas tecnologias, criando a possibilidade de acesso a diagnóstico assistido por IA em regiões com escassez histórica de especialistas. Contudo, sua efetividade está condicionada à infraestrutura de conectividade, de políticas de inclusão digital, de modelos de validação que garantam equidade de desempenho entre populações distintas e de mecanismos de suporte aos profissionais que, muitas vezes sem formação especializada, utilizarão essas ferramentas no contexto de atenção primária.

24 REDEFINIÇÃO DO PAPEL DO PROFISSIONAL DE SAÚDE

A narrativa da substituição, em que a máquina toma o lugar do clínico, é recorrente mas não se sustenta diante dos dados disponíveis. O que a evidência disponível sugere é uma reconfiguração de papéis, não uma eliminação. O profissional de saúde na era da IA não faz menos do que antes: faz diferente, com ferramentas distintas, em um contexto de maior complexidade informacional e de novas responsabilidades.

As tarefas mais diretamente afetadas pela automação são aquelas que envolvem detecção e classificação de padrões em dados estruturados: análise de imagens, triagem de resultados laboratoriais, identificação de riscos em grandes populações. São precisamente as tarefas em que os algoritmos têm demonstrado desempenho mais consistente. Essa delegação progressiva libera tempo clínico para atividades que demandam julgamento contextual e que os sistemas atuais de IA não são capazes de executar adequadamente.

O profissional de saúde do futuro próximo será, crescentemente, um intérprete: alguém que recebe análises produzidas por sistemas de IA e as integra ao contexto humano, social e ético do paciente. Esse papel é tecnicamente mais exigente do que o de executor de procedimentos diagnósticos padronizados, requerendo maior profundidade de conhecimento, maior capacidade analítica e maior sofisticação ética. A IA não simplifica o trabalho do clínico: ela amplia sua complexidade ao mesmo tempo em que expande suas ferramentas.

25 MODELO CONCEITUAL DO DIAGNÓSTICO HÍBRIDO HUMANO-ALGORITMO

O diagnóstico híbrido pode ser compreendido como um sistema de três componentes que interagem de forma sequencial e iterativa: dados, algoritmo e julgamento clínico. Os dados clínicos, provenientes de exames, histórico do paciente, sinais clínicos e informações complementares, alimentam o sistema de IA. O algoritmo processa esses dados e produz estimativas probabilísticas sobre diagnósticos possíveis. O clínico interpreta essas estimativas à luz do contexto do paciente, de fatores que os dados não capturam, e de considerações éticas e de preferências do paciente, chegando então a uma decisão.

Esse modelo contribui para além da velocidade ou volume de processamento. É a estruturação explícita da incerteza e a separação entre o que os dados mostram objetivamente, o que o algoritmo infere estatisticamente e o que o clínico decide a partir dessas informações. Essa separação aumenta a transparência do processo e sua rastreabilidade, o que facilita o aprendizado a partir de erros e a responsabilização quando eles ocorrem.

O modelo híbrido não é uma solução definitiva para os desafios do diagnóstico em saúde. Ele desloca alguns problemas tradicionais e cria outros novos. Mas representa, no estado atual do

desenvolvimento tecnológico, a abordagem mais responsável para integrar as capacidades da IA à prática clínica sem comprometer a centralidade do julgamento humano no cuidado ao paciente.

26 IMPLICAÇÕES PRÁTICAS PARA A ROTINA CLÍNICA

Na prática do dia a dia, a incorporação da IA altera ritmos, distribuições de atenção e formas de interação com a informação. A análise inicial de uma radiografia, que antes consumia um tempo significativo do especialista, pode ser realizada pelo algoritmo em segundos, identificando alterações que merecem atenção e priorizando casos de maior urgência. O especialista humano recebe então um conjunto de casos pré-selecionados, com anotações sobre as regiões de maior interesse, e pode focar sua atenção na avaliação crítica e na interpretação contextual.

Entretanto, essa eficiência tem custos que precisam ser reconhecidos e gerenciados. O principal deles é o risco de que a velocidade e a aparente objetividade do algoritmo pressionem o profissional a aceitar suas conclusões sem reflexão adequada. Quando o sistema marca uma lesão como suspeita, o profissional tende a direcionar sua avaliação para confirmar o achado indicado pelo sistema, reduzindo a consideração de hipóteses alternativas. O tempo preservado pela automação precisa ser reinvestido na qualidade do raciocínio clínico, não apenas no aumento do volume de atendimentos.

Na odontologia, por exemplo, a utilização de IA em radiografias pode antecipar a identificação de alterações e direcionar a atenção do profissional para regiões específicas, permitindo que a consulta se concentre na integração clínica e no planejamento terapêutico personalizado. Isso tende a aumentar a qualidade diagnóstica, mas exige que o profissional preserve seu raciocínio clínico ativo, em vez de se tornar um validador passivo das conclusões do algoritmo.

27 DIRETRIZES PARA USO SEGURO E EFICAZ DA IA NO DIAGNÓSTICO

O uso responsável da IA no diagnóstico em saúde pode ser orientado por um conjunto de princípios que, embora não esgotem a complexidade do problema, oferecem um ponto de partida para a tomada de decisão institucional e individual. O primeiro e mais fundamental é que a IA deve ser utilizada como ferramenta de apoio ao julgamento clínico, nunca como substituta dele. O profissional mantém autonomia e responsabilidade decisória em todos os casos, independentemente do que o sistema recomende.

O segundo princípio é o da validação adequada. Nenhum sistema de IA deve ser implementado em contexto clínico sem evidências robustas de desempenho na população e no contexto específicos de uso pretendido. O desempenho demonstrado em estudos publicados, frequentemente realizados em condições ideais, não é garantia de desempenho equivalente no mundo real. A validação externa, em condições que reproduzam as do uso pretendido, é uma exigência metodológica essencial.

O terceiro princípio é o do monitoramento contínuo. O desempenho de um modelo de IA pode se degradar ao longo do tempo, à medida que as características dos dados de entrada se alteram. Mecanismos de monitoramento sistemático são necessários para detectar quedas de desempenho e permitir intervenção antes que erros sistemáticos causem danos significativos.

O quarto princípio é o da transparência. Profissionais de saúde e pacientes têm o direito de saber quando um sistema de IA participou do processo diagnóstico, de compreender em termos gerais como o sistema funciona e de ter acesso a informações sobre seu desempenho e seus limites conhecidos. Essa transparência é pré-requisito para uma confiança tecnicamente fundamentada, a única eticamente aceitável em contextos de saúde.

28 INTEGRAÇÃO COM PESQUISA CIENTÍFICA E PRODUÇÃO DE EVIDÊNCIA

A IA transforma também a forma como o conhecimento científico em saúde é produzido. A capacidade de analisar grandes bases de dados com múltiplas variáveis abre possibilidades para identificar associações e gerar hipóteses em escala que os métodos tradicionais simplesmente não permitem. Nesse contexto, análises integradas de inúmeros registros de dados clínicos, laboratoriais, genômicos e de imagem, estão se tornando tecnicamente viáveis e metodologicamente relevantes.

Mas esse potencial analítico exige rigor metodológico redobrado. A principal armadilha dos estudos baseados em grandes volumes de dados é a identificação de correlações estatisticamente robustas sem significado causal. Com dados suficientemente grandes, quase qualquer correlação se torna estatisticamente significativa, o que torna os testes de significância estatística tradicionais insuficientes como critério de relevância clínica. A distinção entre correlação e causalidade, sempre importante na epidemiologia, torna-se crítica quando algoritmos identificam associações em espaços de dados de altíssima dimensionalidade.

A reprodutibilidade dos modelos é outra dimensão essencial. Estudos que desenvolvem algoritmos de IA precisam descrever seus métodos e disponibilizar seus dados com transparência suficiente para que outros grupos possam replicar e validar os resultados. Essa exigência de ciência aberta é especialmente relevante na pesquisa de IA em saúde, onde a opacidade metodológica pode esconder problemas que só se tornam visíveis quando os modelos são aplicados em outros contextos.

29 LIMITES EPISTEMOLÓGICOS DA INTELIGÊNCIA ARTIFICIAL NO DIAGNÓSTICO

Por fim, é necessário delimitar com clareza o que a IA não é e o que ela não pode fazer, pelo menos com as abordagens atualmente disponíveis. Os sistemas de IA processam, mas não compreendem. Classificam, mas não interpretam no sentido hermenêutico. Identificam padrões, mas não raciocinam sobre causas. Essas não são insuficiências técnicas que serão superadas com mais

dados ou mais poder computacional: refletem diferenças estruturais entre inferência estatística e raciocínio clínico.

Um algoritmo treinado em dados históricos apresenta desempenho comprometido diante de um quadro clínico radicalmente novo, seja uma doença emergente como a que ocorreu durante a pandemia de COVID-19, uma apresentação atípica de uma condição conhecida ou uma combinação de fatores que nunca apareceu nos dados de treinamento. Além disso, ele não detecta dados incorretos, não identifica quando o contexto muda as regras da interpretação e não reconhece quando o paciente avaliado não se encaixa no perfil da população de treinamento. Essas são situações em que o julgamento clínico humano é insubstituível.

Reconhecer esses limites epistemológicos não representa pessimismo tecnológico, mas condição necessária para o uso responsável da IA. Uma ferramenta usada além de seus limites de competência não é apenas ineficaz: é perigosa. E no contexto da saúde, onde as consequências dos erros recaem sobre pessoas vulneráveis, a clareza sobre esses limites é uma exigência ética fundamental.

30 CONSIDERAÇÕES FINAIS

A Inteligência Artificial representa uma das transformações mais significativas na história do diagnóstico em saúde, comparável, em profundidade, à introdução do microscópio ou dos exames laboratoriais automatizados. Ela amplia a capacidade analítica, aumenta a sensibilidade diagnóstica, explicita a incerteza inerente ao processo clínico e cria possibilidades que a prática tradicional não poderia imaginar. Ignorar esse potencial contraria a evidência disponível e compromete o avanço científico da área.

Mas a IA não resolve os problemas fundamentais da clínica. Ela os reconfigura. A incerteza diagnóstica permanece, agora expressa em probabilidades algorítmicas. Os erros persistem, agora com a capacidade de se reproduzir em escala. As desigualdades em saúde continuam, agora com o risco de serem amplificadas por dados históricos que as refletem. E a responsabilidade pelo cuidado do paciente permanece, integralmente, com o ser humano que está à frente dele.

O futuro do diagnóstico em saúde depende da construção deliberada e eticamente fundamentada de uma parceria que preserve o que cada um tem de melhor: a capacidade analítica e a consistência do algoritmo, e a sensibilidade, o contexto, o julgamento ético e a presença humana do profissional. Construir essa parceria é o desafio central da medicina e da odontologia contemporâneas. É uma responsabilidade compartilhada entre os que desenvolvem as tecnologias, os que as regulam, os que formam os profissionais e os que cuidam, dia após dia, de quem precisa de cuidado.

REFERÊNCIAS

1. Foucault M. O nascimento da clínica. 7ª ed. Rio de Janeiro: Forense Universitária; 2011.
2. Topol EJ. Deep medicine: how artificial intelligence can make healthcare human again. New York: Basic Books; 2019.
3. 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*. 2017;542(7639):115-8.
4. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med*. 2019;380(14):1347-58.
5. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447-53.
6. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-44.
7. Goodfellow I, Bengio Y, Courville A. Deep learning. Cambridge: MIT Press; 2016.
8. Hinton G, Deng L, Yu D, Dahl GE, Mohamed AR, Jaitly N, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Process Mag*. 2012;29(6):82-97.
9. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402-10
10. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. *Nature*. 2020;577(7788):89-94.
11. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*; 2019. p. 4171-86.
12. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst*. 2012;25:1097-105.
13. Char DS, Shah NH, Magnus D. Implementing machine learning in health care - addressing ethical challenges. *N Engl J Med*. 2018;378(11):981-3.
14. 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*. 2018;28(4):689-707.
15. Beauchamp TL, Childress JF. Principles of biomedical ethics. 8th ed. New York: Oxford University Press; 2019.
16. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. *Nat Mach Intell*. 2019;1(9):389-99.
17. Shortliffe EH, Sepúlveda MJ. Clinical decision support in the era of artificial intelligence. *JAMA*. 2018;320(21):2199-200.

18. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ Digit Med.* 2020;3:17.
19. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. *Acad Med.* 2003;78(8):775-80.
20. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med.* 2005;165(13):1493-9.
21. Norman GR, Monteiro SD, Sherbino J, Ilgen JS, Schmidt HG, Mamede S. The causes of errors in clinical reasoning: cognitive biases, knowledge deficits, and dual process thinking. *Acad Med.* 2017;92(1):23-30.
22. Schuwirth LWT, Ash J. Can clinical decision support systems improve diagnostic performance? *BMJ.* 2013;346:f3444.
23. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell.* 2019;1(5):206-15.
24. Ribeiro MT, Singh S, Guestrin C. "Why should I trust you?": explaining the predictions of any classifier. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; 2016. p. 1135-44.
25. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. In: *Proceedings of the IEEE International Conference on Computer Vision*; 2017. p. 618-26.
26. Schwendicke F, Samek W, Krois J. Artificial intelligence in dentistry: chances and challenges. *J Dent Res.* 2020;99(7):769-74.
27. Shan T, Tay FR, Gu L. Application of artificial intelligence in dentistry. *J Dent Res.* 2021;100(3):232-44.
28. Lee JH, Kim DH, Jeong SN, Choi SH. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J Dent.* 2018;77:106-11.
29. Hung K, Montalvao C, Tanaka R, Kawai T, Bornstein MM. The use and performance of artificial intelligence applications in dental and maxillofacial radiology: a systematic review. *Dentomaxillofac Radiol.* 2020;49(1):20190107.
30. Burt BA. Prevention policies in the light of the changed distribution of dental caries. *Acta Odontol Scand.* 1998;56(3):179-86.
31. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential biases in machine learning algorithms using electronic health record data. *JAMA Intern Med.* 2018;178(11):1544-7.
32. Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: addressing ethical challenges. *PLoS Med.* 2018;15(11):e1002689.
33. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ.* 1996;312(7023):71-2.

34. Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? *BMJ*. 2014;348:g3725.
35. Price WN 2nd, Cohen IG. Privacy in the age of medical big data. *Nat Med*. 2019;25(1):37-43.
36. Murdoch TB, Detsky AS. The inevitable application of big data to health care. *JAMA*. 2013;309(13):1351-2.
37. Mittelstadt BD, Floridi L. The ethics of big data: current and foreseeable issues in biomedical contexts. *Sci Eng Ethics*. 2016;22(2):303-41.
38. Panch T, Szolovits P, Atun R. Artificial intelligence, machine learning and health systems. *J Glob Health*. 2018;8(2):020303.
39. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng*. 2018;2(10):719-31.
40. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol*. 2017;2(4):230-43.
41. Chen JH, Asch SM. Machine learning and prediction in medicine - beyond the peak of inflated expectations. *N Engl J Med*. 2017;376(26):2507-9.
42. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med*. 2019;25(1):30-6.
43. Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med*. 2019;25(9):1337-40.
44. Kerasidou A. Artificial intelligence and the ongoing need for empathy, compassion and trust in healthcare. *Bull World Health Organ*. 2020;98(4):245-50.
45. Briganti G, Le Meur O. Artificial intelligence in medicine: today and tomorrow. *Front Med (Lausanne)*. 2020;7:27.
46. Kahneman D. *Thinking, fast and slow*. New York: Farrar, Straus and Giroux; 2011.
47. Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. *J Am Med Inform Assoc*. 2012;19(1):121-7.
48. Parasuraman R, Manzey DH. Complacency and bias in human use of automation: an attentional integration. *Hum Factors*. 2010;52(3):381-410.
49. U.S. Food and Drug Administration. Artificial intelligence and machine learning in software as a medical device [Internet]. Silver Spring: FDA; 2021 [citado 2025 Abr 26]. Disponível em: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>
50. European Commission. Proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Brussels: European Commission; 2021.

51. Celi LA, Cellini J, Charpignon ML, Dee EC, Dernoncourt F, Eber R, et al. Sources of bias in artificial intelligence that perpetuate healthcare disparities—a global review. *PLoS Digit Health*. 2022;1(3):e0000022.
52. Irvin J, Rajpurkar P, Ko M, Yu Y, Ciurea-Ilcus S, Chute C, et al. CheXpert: a large chest radiograph dataset with uncertainty labels and expert comparison. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. 2019;33(01):590-7.
53. Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, et al. Language models are few-shot learners. *Adv Neural Inf Process Syst*. 2020;33:1877-901.
54. McCradden MD, Joshi S, Mazwi M, Anderson JA. Ethical limitations of algorithmic fairness solutions in health care machine learning. *Lancet Digit Health*. 2020;2(5):e221-3.
55. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med*. 2019;17(1):195.
56. Maslove DM, Podchiyska T, Lowe HJ. Discretization of continuous features in clinical datasets. *J Am Med Inform Assoc*. 2013;20(5):821-6.
57. Sendak MP, Gao M, Brajer N, Balu S. Presenting machine learning model information to clinical end users with model facts labels. *NPJ Digit Med*. 2020;3:41.
58. Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. *JAMA*. 2017;318(6):517-8.
59. Castelvechi D. Can we open the black box of AI? *Nature*. 2016;538(7623):20-3.
60. London AJ. Artificial intelligence and black-box medical decisions: accuracy versus explainability. *Hastings Cent Rep*. 2019;49(1):15-21.