

Psicoterapia e inteligência artificial: limites tecnológicos na promoção da saúde

Sophia Ivantes-Rodrigues¹, Marília da Matta Silva¹, Leonardo Pestillo de Oliveira¹, Lucas França Garcia¹

1. Universidade Cesumar, Maringá/PR, Brasil.

Resumo

A integração da inteligência artificial (IA) na psicologia tem registrado um crescimento significativo, porém as questões bioéticas associadas ainda são pouco discutidas. Este ensaio crítico analisa os desafios bioéticos, biojurídicos e os que se relacionam com o Código de Ética do Profissional da Associação Americana de Psicologia no contexto do uso de inteligências artificiais para psicoterapia. Duas categorias principais foram desenvolvidas: “Inteligência artificial e psicoterapia: funcionalidade, integração e aplicação” e “Desafios e limites bioéticos na integração da inteligência artificial na psicologia”. O ensaio examina o caráter dessas intervenções, destacando seus riscos e benefícios potenciais, bem como os limites éticos envolvidos. Conclui-se que a integração de inteligências artificiais na psicoterapia apresenta riscos e limitações que demandam estudos mais aprofundados e maior atenção e discussão por parte da comunidade científica e dos profissionais da área.

Palavras-chave: Inteligência artificial. Bioética. Psicologia. Confidencialidade. Privacidade.

Resumen

Psicoterapia e inteligencia artificial: límites tecnológicos en la promoción de la salud mental

La integración de la inteligencia artificial en psicología ha crecido significativamente, aunque las cuestiones bioéticas relacionadas con esta práctica siguen siendo poco exploradas. Este ensayo crítico analiza los desafíos bioéticos y biojurídicos, así como aquellos relacionados con el Código de Ética del Profesional de la Asociación Americana de Psicología respecto al uso de inteligencias artificiales en psicoterapia. Se desarrollaron dos categorías principales: “Inteligencias artificiales y psicoterapia: funcionalidad, integración y aplicación” y “Desafíos y límites bioéticos en la integración de la inteligencia artificial en la psicología”. Este ensayo analizó la naturaleza de estas intervenciones destacando sus posibles riesgos y beneficios, así como los límites éticos involucrados. Se concluye que la integración de inteligencias artificiales en psicoterapia presenta riesgos y limitaciones que requieren estudios más profundos, además de una mayor atención y debate por parte de la comunidad científica y los profesionales del área.

Palabras clave: Inteligencia artificial. Bioética. Psicología. Confidencialidad. Privacidad.

Abstract

Psychotherapy and artificial intelligence: the limits of technology in the promotion of mental health

The integration of artificial intelligence in psychology has been growing significantly, although bioethical issues related to this practice remain underexplored. This study aims to analyze the bioethical and biolegal challenges, including those related to the American Psychological Association's Code of Ethics, concerning the use of artificial intelligence in psychotherapy. This critical article addresses the topic from a reflective perspective. Two main categories were developed: “Artificial intelligence and psychotherapy: functionality, integration, and application” and “Bioethical challenges and limits in the integration of artificial intelligence in psychology.” The nature of these interventions was analyzed, highlighting their potential risks and benefits, as well as the ethical boundaries involved. It is concluded that the integration of artificial intelligence into psychotherapy practice presents risks and limitations that require further studies, increased scrutiny, and discussion by the scientific community and professionals in the field.

Keywords: Artificial intelligence. Bioethics. Psychology. Confidentiality. Privacy.

Declararam não haver conflito de interesse.

Com o crescimento do uso das redes sociais e o avanço tecnológico no século XXI, a integração entre psicologia, ciências do comportamento e inteligências artificiais (IA) tornou-se uma realidade inerente^{1,2}. Apesar das contribuições significativas dessa integração, como a ampliação do acesso à psicoterapia, surgem preocupações éticas e morais. Essas questões incluem possíveis infrações ao Código de Ética Profissional da Associação Americana de Psicologia (APA), além de desafios aos princípios da bioética e do biodireito³.

As preocupações centrais estão nos impactos positivos e negativos para pacientes e psicólogos, bem como no uso de dados para marketing e na natureza ainda pouco conhecida das intervenções realizadas por IA, frequentemente vistas como uma psicoterapia não autêntica. Nesse cenário, princípios bioéticos como a beneficência e a não maleficência são tensionados por riscos como vieses algorítmicos e decisões automatizadas imprecisas, enquanto a justiça é questionada em termos de acessibilidade e distribuição equitativa dessas tecnologias^{4,5}. Holohan e Müller⁶, nesse contexto, apresentam o conceito de “*significant otherness*”, destacando a relevância da alteridade e do reconhecimento da singularidade do paciente nas psicoterapias assistidas por IA.

Sob a perspectiva do biodireito, a crescente aplicação da IA na psicoterapia demanda uma vigilância rigorosa quanto à privacidade, confidencialidade e segurança dos dados pessoais em conformidade com regulamentações como a Lei Geral de Proteção de Dados (LGPD)⁷ e o General Data Protection Regulation (GDPR)⁸. Aspectos cruciais, tais como o consentimento informado, a salvaguarda de informações sensíveis e a transparência nas operações algorítmicas, são essenciais para assegurar os direitos dos pacientes^{3,9}. Este ensaio, portanto, realiza uma análise crítica das implicações bioéticas do emprego da IA em psicoterapia, explorando os princípios bioéticos, a importância da alteridade e os desafios jurídicos que acompanham essa inovação.

Método

Tipo de estudo

Este ensaio crítico aborda as principais questões bioéticas vinculadas à integração de IA no contexto da psicoterapia, baseando-se em uma análise

reflexiva da literatura disponível. Essa metodologia possibilita uma avaliação aprofundada das diversas pesquisas sobre o assunto, com o objetivo de discutir os desafios e as implicações éticas da incorporação da tecnologia de IA na psicologia.

Fontes de informação e estratégia de busca

As buscas foram conduzidas nas bases de dados PubMed, Scopus, Web of Science e SciELO em janeiro de 2025, utilizando a combinação dos termos “*artificial intelligence*”, “*psychotherapy*” e “*bioethics OR ethics*”.

CrITÉrios de inclusão e exclusão

O ensaio incluiu artigos publicados em periódicos científicos de 2020 a 2025, em inglês, espanhol e português, focados nos aspectos éticos, bioéticos e jurídicos da integração de IA na psicoterapia. Artigos que não estavam alinhados com o objetivo central foram desconsiderados.

Triagem e seleção dos estudos

Após a coleta de dados nas bases mencionadas, os artigos selecionados na primeira etapa de triagem foram examinados integralmente. As divergências entre os avaliadores foram solucionadas por consenso ou, quando necessário, com a intervenção de um terceiro revisor.

Análise e síntese dos dados

Após a conclusão da triagem, os artigos foram baixados em sua totalidade, e seus dados foram organizados e codificados utilizando o *software* QSR NVivo 15 para Windows¹⁰. O estudo seguiu o método de análise de conteúdo proposto por Bardin, que incluiu a categorização temática dos dados e sua interpretação para identificar padrões e tendências recorrentes¹¹. A síntese dos dados foi feita de maneira narrativa, enfatizando os principais resultados e explorando suas relações com as questões bioéticas do uso de IA na psicoterapia. Essa estratégia visou identificar padrões, contrastes e lacunas nos estudos avaliados¹². Além disso, dois pesquisadores (SRI e LFG) realizaram análises e codificações independentes, agrupando as unidades de registro em unidades de significado (conceitos-chave), que foram consolidadas em categorias temáticas.

Tabela 1. Categorias temáticas

Unidade temática central	Categorias temáticas	Unidades de significado (exemplos)	Evidências nos artigos
IA e psicoterapia: funcionalidade, integração e aplicação	Eficácia clínica	Chatbots aumentam adesão terapêutica ¹³	12/17 artigos (70,6%)
		Limitações na empatia artificial ¹⁴	
	Digital therapeutic alliance (DTA)	Replicação de vínculos via personalização ¹⁵	9/17 artigos (52,9%)
		Falta de aprofundamento emocional ⁶	
Desafios e limites bioéticos	Autonomia e consentimento informado	Opacidade algorítmica compromete decisões ¹⁶	14/17 artigos (82,4%)
		Necessidade de linguagem acessível ¹⁷	
	Proteção de dados e conformidade legal	Riscos do <i>digital phenotyping</i> sob LGPD/GDPR ¹⁸	11/17 artigos (64,7%)
		Falta de criptografia robusta ¹⁹	

Resultados e discussão

Nesta seção, serão discutidos os principais achados do ensaio crítico acerca da implementação de inteligências artificiais na psicoterapia, estruturados em categorias temáticas resultantes da análise de conteúdo dos 17 artigos examinados. Os artigos são provenientes de regiões como Europa, América do Norte e Ásia, e não foram encontrados estudos brasileiros que satisfizessem os critérios de inclusão.

As implicações éticas do uso de IA na prática psicoterapêutica foram destacadas, abordando seus impactos tanto positivos quanto negativos para psicólogos e pacientes. As categorias desenvolvidas refletem os principais debates e desafios bioéticos presentes na literatura, incluindo: “IA e psicoterapia: funcionalidade, integração e aplicação” e “Desafios e limites bioéticos na integração das IA na psicologia”. Além disso, subcategorias foram definidas para explorar questões específicas, como “Potenciais violações e orientações do Código de Ética do Profissional da Associação Americana de Psicologia” e “Implicações jurídicas”. Essa estrutura possibilita uma análise mais detalhada dos limites normativos e legais associados à implementação dessa tecnologia no ambiente clínico.

IA e psicoterapia: funcionalidade, integração e aplicação

A utilização da IA na psicoterapia está em rápido desenvolvimento, proporcionando novas ferramentas para melhorar a saúde mental, ao mesmo tempo

que introduz desafios significativos. Grodniewicz e Hohol¹⁴ apontam três obstáculos principais: a primeira, o “terapeuta confuso”, ressalta a dificuldade da IA em compreender nuances emocionais específicas; a segunda, o “terapeuta não humano”, sublinha as limitações nas relações interpessoais e a ausência de empatia genuína; e a terceira, o “terapeuta de inteligência restrita”, demonstra a dificuldade da IA em fornecer intervenções abrangentes e integradas. Alfano e colaboradores¹³, assim como Khawaja e Bélisle-Pipon¹⁶, reconhecem o papel da IA para expandir o acesso a tratamentos psicológicos, particularmente em comunidades vulneráveis, mas advertem sobre os riscos associados, incluindo a excessiva padronização e o comprometimento da humanização no atendimento.

Alfano e colaboradores¹³ destacam que, como resultado da pandemia de covid-19, observou-se um crescimento notório no número de empresas que passaram a oferecer serviços de psicoterapia utilizando IA. Essas empresas flexibilizaram o acesso à psicoterapia, eliminando barreiras como limitações geográficas, vulnerabilidade e custos associados. Embora os dados fornecidos por essas empresas sejam restritos, eles revelam que os usuários mais assíduos têm entre 16 e 25 anos, uma faixa etária que cresceu imersa em tecnologia e vivenciou sua adoção em larga escala durante a pandemia.

A crescente demanda por psicoterapia mediada por IA, particularmente entre adolescentes, está associada a fatores como o constrangimento em compartilhar eventos e sentimentos com

psicólogos humanos, a familiaridade com plataformas digitais e a superação de barreiras físicas e econômicas. No entanto, apesar desses benefícios, aspectos fundamentais da prática psicológica, como a desistência da psicoterapia e a falta de um ambiente terapêutico apropriado, suscitam dúvidas relevantes. Entre essas preocupações estão os interesses financeiros das empresas, a carência de provas que atestem a eficácia dos serviços e a ausência de qualidades humanizadas nos *chatbots*, tais como empatia e confidencialidade. Adicionalmente, há questionamentos sobre o funcionamento dos algoritmos, especialmente quanto aos critérios de avaliação da qualidade dos serviços e às possíveis transformações no mercado de trabalho¹³.

De fato, *chatbots* já estão sendo implementados em aplicativos que utilizam algoritmos para detectar padrões linguísticos indicativos de risco de suicídio e depressão em jovens adultos e adolescentes. Nesse cenário, a IA se destaca pela capacidade de processar grandes volumes de dados em poucos minutos, enquanto o tratamento em si continua a depender fortemente do psicólogo e do valor intrínseco do relacionamento humano na intervenção e no contato com o paciente¹³. A qualidade da relação terapêutica mediada por IA é, portanto, crucial. Malouin-Lachance e colaboradores¹⁵ introduzem o conceito de aliança terapêutica digital, enfatizando a tendência de formar vínculos terapêuticos por meio de interações digitais, desde que estas sejam personalizadas, responsivas e adaptáveis ao histórico emocional do paciente. Entretanto, essas interações digitais são meramente uma reprodução dos mecanismos terapêuticos tradicionais, manifestando-se de maneira significativamente diferente das interações humanas, especialmente no que tange à profundidade emocional.

Holohan e Müller⁶ ampliam essa discussão com o conceito de *significant otherness*, argumentando que as experiências terapêuticas mediadas por IA podem ter um impacto emocional considerável quando a tecnologia é desenvolvida com empatia digital, interatividade adaptativa e linguagem afetiva. Eles sugerem que a relação entre paciente e IA, nesse contexto, pode ser comparada ao conceito psicanalítico de transferência, em que as projeções emocionais do paciente sobre a tecnologia podem ser mais bem exploradas por meio de estudos aprofundados e avaliadas com um olhar crítico por psicólogos e pesquisadores.

No âmbito das inovações práticas, Özden³ investiga o potencial das terapias imersivas com IA, que englobam realidade aumentada, ambientes simulados e intervenções sensoriais personalizadas, ressaltando sua eficácia no tratamento de fobias, traumas e transtornos de ansiedade. Zhang² enfatiza a importância do reconhecimento emocional em tempo real, detalhando como algoritmos de IA, ao analisar variações de linguagem, tom de voz e expressões faciais, tornam as interações mais responsivas, ajustando o diálogo conforme as emoções detectadas. Dehbozorgi, Zangeneh, Khooshab, Nia e Hanif¹⁷ reforçam a necessidade de cocriação tecnológica, argumentando que a participação ativa de psicoterapeutas e pacientes no processo de desenvolvimento dessas ferramentas é fundamental para garantir maior adesão, confiança e segurança, evitando assim falhas éticas no desenho das intervenções.

Assim, os progressos da IA em psicoterapia apresentam tanto oportunidades quanto desafios significativos. Isso implica não só o desenvolvimento de algoritmos mais atentos ao contexto emocional dos pacientes, mas também a adoção de práticas eticamente adequadas, que incluam supervisão humana constante, auditoria transparente dos processos algorítmicos e formação ética para os profissionais da área. Além disso, é essencial incentivar estudos longitudinais para analisar os efeitos a longo prazo das terapias assistidas por IA, garantindo que elas não apenas aumentem a eficiência tecnológica, mas também preservem a segurança emocional e a qualidade do atendimento psicoterapêutico.

Desafios e limites bioéticos na integração das IA na psicologia

A incorporação da IA em psicoterapia levanta profundos questionamentos bioéticos, particularmente em relação aos princípios da ética biomédica estabelecidos por Beauchamp e Childress⁵. Esses princípios incluem: a autonomia, que respeita a autodeterminação e a capacidade do paciente de tomar decisões informadas; a beneficência, que requer a promoção ativa do bem-estar do paciente; a não maleficência, que impõe a obrigação de evitar causar danos; e a justiça, que demanda uma distribuição justa e equitativa de benefícios e riscos.

A autonomia é um pilar central na relação terapêutica, envolvendo o direito do paciente de tomar decisões informadas, sem coerção ou manipulação. Khawaja e Bélisle-Pipon¹⁶ e Zhang² apontam que a IA pode minar esse princípio ao gerar uma percepção enganosa de vínculo terapêutico, especialmente quando não é evidente o papel da tecnologia na interação. O consentimento informado é um aspecto crucial: Dehbozorgi e colaboradores¹⁷ destacam que ele deve ser transparente, acessível e compreensível, evitando mal-entendidos sobre as limitações e capacidades da IA. Holohan e Müller⁶ também indicam que a opacidade dos algoritmos prejudica a compreensão do paciente sobre os processos que norteiam as decisões automatizadas. Portanto, para promover a autonomia, é necessário que as plataformas de IA adotem práticas como a transparência algorítmica, o consentimento adaptativo e o uso de linguagem clara. Esse comprometimento com a autodeterminação digital reforça o papel do paciente como agente ativo em seu processo terapêutico.

O princípio da beneficência envolve a promoção ativa do bem-estar do paciente, ao passo que a não maleficência impõe o dever de evitar danos. Stoll e colaboradores⁴ e Grodniewicz e Hohol¹⁴ alertam que, apesar de a IA ampliar o acesso a terapias, algoritmos enviesados ou treinados com dados limitados podem, na verdade, gerar mais prejuízos do que benefícios. Isso inclui intervenções inadequadas ou potencialmente nocivas, além de perpetuar estigmas e preconceitos. Além disso, Özden³ enfatiza que a ausência de supervisão adequada em terapias imersivas pode resultar em desconforto emocional ou dependência tecnológica.

Adicionalmente, surgem preocupações éticas relacionadas ao uso da IA: Dehbozorgi e colaboradores¹⁴ apontam que sistemas automatizados que fornecem diagnósticos sem validação humana podem gerar falsas expectativas ou tratamentos inadequados. Outro ponto é que a personalização inadequada das interações com IA pode levar a respostas genéricas, comprometendo o cuidado, conforme sugerem Holohan e Müller⁶. Para mitigar esses riscos, é essencial implementar práticas de auditoria algorítmica, testes contínuos e o monitoramento humano das decisões automatizadas. A integração de protocolos de segurança, o treinamento de IA com dados representativos e a validação constante da eficácia das terapias

digitais são, portanto, fundamentais para garantir o cumprimento dos princípios de beneficência e não maleficência.

A justiça na bioética exige que os benefícios e riscos da IA em psicoterapia sejam distribuídos de forma equitativa, promovendo acessibilidade e inclusão². Nesse sentido, Bhatt²⁰ e Gutierrez e colaboradores²¹ apontam que a justiça se concretiza quando as inovações tecnológicas alcançam diferentes grupos sociais, evitando disparidades no acesso. Grodniewicz e Hohol¹⁴ alertam que algoritmos treinados com dados homogêneos podem perpetuar desigualdades raciais e culturais, ressaltando a necessidade de conjuntos de dados diversos e auditáveis. Stoll e colaboradores⁴ também enfatizam que desigualdades digitais, como falta de acesso à internet, intensificam a exclusão de grupos vulneráveis. Portanto, a justiça demanda uma visão multidimensional, levando em conta os impactos sociais, culturais e econômicos da IA em psicoterapia.

A alteridade, que significa reconhecer a singularidade do outro, é fundamental para criar relações terapêuticas significativas⁶. Tavory⁹ também defende que a IA deve ser desenvolvida com *response-ability*, ou seja, a habilidade de responder com sensibilidade às subjetividades de cada pessoa. Dehbozorgi e colaboradores¹⁷ acrescentam que a falta de diversidade cultural nos algoritmos pode resultar em interações insensíveis, reforçando estereótipos e estigmas sociais ao invés de proporcionar um cuidado verdadeiramente inclusivo.

Potenciais violações e orientações do Código de Ética do Profissional da Psicologia

O Código de Princípios Éticos da APA orienta os psicólogos a diagnosticar após uma avaliação minuciosa do paciente. Isso cria uma zona cinzenta, pois não especifica se a atuação de IA na prática clínica é eticamente adequada ou não. A APA, no entanto, permite que os psicólogos possam adotar essas tecnologias, desde que eles esclareçam as limitações desses métodos e como elas podem influenciar a confiabilidade e a validade de suas conclusões¹.

Quanto à confidencialidade e à privacidade ao consentimento informado, o psicólogo deve informar ao paciente que a sua privacidade pode estar em risco, principalmente se os dados forem

para fins de marketing. Além disso, é crucial explicar quais são os potenciais danos e o objetivo do uso dessas informações e obter o consentimento do paciente de forma clara. Vale lembrar que o uso dessas práticas não atende às conformidades da *Health Insurance Portability and Accountability Act of 1996*, o que pode comprometer a privacidade do paciente. A APA também ressalta que usar redes sociais para avaliar pacientes pode levar ao uso excessivo de informações que não são necessárias para a avaliação, especialmente quando essas redes são exploradas por meio de IA¹.

Embora os *chatbots* ofereçam serviços personalizados, como atendimento ilimitados em questão de horários e locais, consideração de características culturais e pessoais e integração de abordagens como a terapia cognitivo-comportamental (TCC) e a psicanálise, eles têm suas limitações. Essas limitações – como a potencial exploração de usuários, conselhos com vieses algorítmicos e recursos terapêuticos que podem ser prejudiciais – podem levar ao equívoco terapêutico (ET) quando mal compreendidas ou ignoradas¹⁴.

O ET, tema amplamente debatido na bioética, refere-se à subestimação das capacidades da IA em fornecer um suporte terapêutico adequado, ético e eficaz. Esse equívoco pode surgir de quatro maneiras: por meio de marketing impreciso, formação de uma aliança terapêutica digital, design inadequado dos *chatbots* e limitação da autonomia do usuário¹⁴.

Ademais, a falta de empatia genuína dos *chatbots*, devido à ausência de emoções pode causar danos ao usuário e impedir a formação de um vínculo terapêutico, que é fundamental na psicoterapia. As IA, embora sejam capazes de identificar e manipular emoções e vulnerabilidades, podem colocar grupos vulneráveis e minoritários em risco de danos e expectativas irreais⁹. Sem a presença de um psicólogo humano, não há desenvolvimento de confiança ou decisão compartilhada, o que pode comprometer a autonomia do paciente^{3,14}. Além disso, embora as IA demonstrem eficácia a curto prazo, não apresentem resultados positivos em serviços de longo prazo. Questões como a retenção de memória podem afetar a continuidade do processo terapêutico³.

Implicações jurídicas

Quanto aos desafios jurídicos e de proteção de dados, Gerke e colaboradores¹⁸ ressaltam que

o uso de IA em saúde mental deve seguir marcos regulatórios como a LGPD e o GDPR. Eles enfatizam a importância de ter protocolos rigorosos para coletar, armazenar e processar dados sensíveis. Jin e colaboradores¹⁹ indicam que os riscos da coleta contínua de dados por agentes conversacionais e plataformas de telepsicologia são elevados, especialmente quando há compartilhamento não autorizado de informações. Stoll e colaboradores⁴ apontam a necessidade de políticas claras de consentimento e de criptografia robusta para proteger a confidencialidade do paciente.

Dehbozorgi e colaboradores¹⁷ alertam para os riscos do *digital phenotyping* – a análise de padrões digitais para inferir o estado emocional do paciente – e recomendam que esses sistemas estejam em conformidade com as diretrizes da LGPD e do GDPR. Gerke e colaboradores¹⁸ defendem a realização de auditorias algorítmicas para garantir que o uso secundário de dados seja devidamente regulamentado. Já Holohan e Müller⁶, assim como Grodniewicz e Hohol¹⁴, destacam que a complexidade dos sistemas de IA pode tornar o processo de consentimento menos transparente, dificultando a compreensão dos riscos envolvidos no uso dessas tecnologias em terapias digitais. Stoll e colaboradores⁴ sugerem a adoção de formulários de consentimento interativos e adaptativos, capazes de informar de maneira clara aos pacientes sobre o funcionamento das tecnologias utilizadas.

Essas reflexões evidenciam que o uso ético da IA em psicoterapia vai além da mera inovação tecnológica; exige práticas centradas nos direitos humanos, com foco em responsabilidade, empatia e inclusão. É fundamental que o desenvolvimento de sistemas de IA inclua processos participativos, envolvendo psicoterapeutas, pacientes e especialistas em bioética, para garantir que os princípios bioéticos sejam aplicados de forma prática e efetiva. Os desafios jurídicos e de proteção de dados destacam a necessidade de equilibrar inovação tecnológica com o respeito aos direitos fundamentais do paciente. O alinhamento entre avanços tecnológicos e conformidade regulatória será crucial para garantir a ética e a segurança nas práticas psicoterapêuticas mediadas por IA.

O sucesso dessas tecnologias será avaliado não apenas pela sua eficiência técnica, mas também pela sua capacidade de promover dignidade, segurança e cuidado centrado no paciente. Além disso,

a regulação clara e a formação de comitês interdisciplinares para avaliar continuamente os impactos éticos da IA são passos essenciais. Isso garante que os avanços tecnológicos contribuam para uma prática psicoterapêutica verdadeiramente ética.

Considerações finais

As implicações do uso da inteligência artificial em psicoterapia, analisadas sob as dimensões éticas, jurídicas e relacionais, revelam a complexidade desse avanço e seus desafios fundamentais. A IA tem o potencial de democratizar o acesso à psicoterapia e personalizar as intervenções. No entanto, Grodniewicz¹⁴ e Dehbozorgi e colaboradores¹⁷ alertam para o risco de vínculos ilusórios e para as limitações dos algoritmos em compreender a complexidade das experiências humanas. A análise sob os princípios da bioética expõe os seguintes dilemas: a ameaça à autonomia diante de consentimentos opacos¹⁶, os riscos à beneficência com vieses algorítmicos⁴ e as desigualdades no acesso à tecnologia, comprometendo a justiça²⁰. Ao mesmo tempo, Gerke e colaboradores¹⁸ e Jin e colaboradores¹⁹ enfatizam a urgência de adequação à LGPD e ao GDPR, destacando que o consentimento informado e a proteção de dados são pilares inegociáveis para a ética no uso de IA em saúde.

Holohan e Müller⁶, assim como Tavery⁹, ressaltam que, para respeitar a alteridade, a IA deve ser projetada com *response-ability*, ou seja, respondendo com sensibilidade às singularidades culturais e emocionais do usuário. Bhatt²⁰ reforça a importância da diversidade nos conjuntos de dados, para evitar vieses culturais.

As IA destinadas aos serviços de saúde mental ainda estão em estágio inicial e não devem

substituir os psicólogos humanos. Ao invés disso, devem ser utilizadas como ferramentas para ampliar as capacidades de intervenção, especialmente em questões como a memória humana, a sobrecarga de profissionais e a rápida identificação de sinais precoces de problemas de saúde mental, como ideações suicidas e episódios de depressão^{2,3,16}. Logo, para evitar a normalização de uma dependência excessiva dessas tecnologias e reduzir os impactos relacionados à busca por apoio e suporte social, é fundamental destacar recorrentemente que as IA não substituem o psicólogo humano¹³.

O investimento governamental em serviços de saúde mental, juntamente com a colaboração multiprofissional entre psicólogos, desenvolvedores de IA e profissionais do direito, é essencial para criar regulamentações que garantam os direitos dos usuários com transparência e clareza. Além disso, é crucial melhorar as IA, para que elas se adaptem às diferenças culturais, sociais e econômicas. Sugere-se que esses serviços ofereçam a opção de atendimento humano em situações de crise e necessidades de suporte mais complexas.

A integração da IA à psicoterapia só será eticamente adequada e eficaz se unir inovação tecnológica com responsabilidade ética e legal. É imperativo que as práticas sejam reguladas, transparentes e focadas no respeito à dignidade humana, buscando um equilíbrio entre o potencial técnico e o cuidado humano que caracteriza a psicoterapia. Dessa forma, conclui-se que as IA não devem substituir os serviços prestados pelos profissionais da psicologia, mas devem ser utilizadas como ferramentas complementares. É necessário respeitar os princípios éticos, os direitos do paciente e garantir a humanização nos atendimentos clínicos.

Referências

1. Fleming MN. Considerations for the ethical implementation of psychological assessment through social media via machine learning. *Ethics Behav* [Internet]. 2021 [acesso 11 ago 2025];31(3):181-92. DOI: 10.1080/10508422.2020.1817026
2. Zhang Z, Wang J. Can AI replace psychotherapists? Exploring the future of mental health care. *Front Psychiatry* [Internet]. 2024 [acesso 11 ago 2025];15:1-7. DOI: 10.3389/fpsy.2024.1444382
3. Özden HC. The use of artificial intelligence in psychotherapy: practical and ethical aspects. *Turk J Psychiatry* [Internet]. 2024 [acesso 11 ago 2025];35(4):336-9. Disponível: <https://bit.ly/46WIOTr>

4. Stoll J, Müller JA, Trachsel M. Ethical issues in online psychotherapy: a narrative review. *Front Psychiatry* [Internet]. 2020 [acesso 11 ago 2025];10:1-16. DOI: 10.3389/fpsy.2019.00993
5. Beauchamp TL, Childress JF. Principles of biomedical ethics. 5ª ed. New York: Oxford University Press; 2001.
6. Holohan M, Müller R. Beyond humanism: telling response-able stories about significant otherness in human-chatbot relations. *Front Psychol* [Internet]. 2024 [acesso 11 ago 2025];15. DOI: 10.3389/fpsyg.2024.1357572
7. Brasil. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). *Diário Oficial da União* [Internet]. Brasília, 15 ago 2018 [acesso 11 ago 2025]; Disponível: <https://bit.ly/4mZaduf>
8. European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). *Official Journal of the European Union* [Internet]. Brussels, 27 abr 2016 [acesso 11 ago 2025]; Disponível: <https://bit.ly/3VVXT2F>
9. Tavory T. Regulating AI in mental health: ethics of care perspective. *JMIR Ment Health* [Internet]. 2024 [acesso 11 ago 2025];11. DOI: 10.2196/58493
10. Houghton C, Murphy K, Meehan B, Thomas J, Brooker D, Casey D. From screening to synthesis: using NVivo to enhance transparency in qualitative evidence synthesis. *J Clin Nurs* [Internet]. 2017 [acesso 11 ago 2025];26(5-6):873-81. DOI: 10.1111/jocn.13443
11. Bardin L. Análise de conteúdo. São Paulo: Edições 70; 2008.
12. Khurshid Z, De Brún A, Martin J, McAuliffe E. A systematic review and narrative synthesis: determinants of the effectiveness and sustainability of measurement-focused quality improvement trainings. *J Contin Educ Health Prof* [Internet]. 2021 [acesso 11 ago 2025];41(3):210-20. DOI: 10.1097/CEH.0000000000000331
13. Alfano L, Malcotti I, Ciliberti R. Psychotherapy, artificial intelligence and adolescents: ethical aspects. *J Prev Med Hyg* [Internet]. 2023 [acesso 11 ago 2025];64(4):438-42. Disponível: <https://bit.ly/4mU8AOk>
14. Grodniewicz JP, Hohol M. Waiting for a digital therapist: three challenges on the path to psychotherapy delivered by artificial intelligence. *Front Psychiatry* [Internet]. 2023 Jun [acesso 11 ago 2025];14. DOI: 10.3389/fpsy.2023.1190084
15. Malouin-Lachance A, Capolupo J, Laplante C, Hudon A. Does the digital therapeutic alliance exist?: integrative review. *JMIR Ment Health* [Internet]. 2025 Feb 7 [acesso 11 ago 2025];12. DOI: 10.2196/69294
16. Khawaja Z, Bélisle-Pipon JC. Your robot therapist is not your therapist: understanding the role of AI-powered mental health chatbots. *Front Digit Health* [Internet]. 2023 [acesso 11 ago 2025];5. DOI: 10.3389/fdgth.2023.1278186
17. Dehbozorgi R, Zangeneh S, Khooshab E, Nia DH, Hanif HR, Samian P *et al*. The application of artificial intelligence in the field of mental health: a systematic review. *BMC Psychiatry* [Internet]. 2025 [acesso 11 ago 2025];25(1):132. DOI: 10.1186/s12888-025-06483-2
18. Gerke S, Minssen T, Cohen IG. Ethical and legal challenges of artificial intelligence-driven healthcare. In: Bohr A, Memarzadeh K. *Artificial intelligence in healthcare*. London: Academic Press; 2020. p. 295-336. DOI: 10.1016/B978-0-12-818438-7.00012-5
19. Jin KW, Li Q, Xie Y, Xiao G. Artificial intelligence in mental healthcare: an overview and future perspectives. *Br J Radiol* [Internet]. 2023 [acesso 11 ago 2025];96(1150):20230213. DOI: 10.1259/bjr.20230213
20. Bhatt S. Digital mental health: role of artificial intelligence in psychotherapy. *Ann Neurosci* [Internet]. 2024 [acesso 11 ago 2025];32(2):117-27. DOI: 10.1177/09727531231221612
21. Gutierrez G, Stephenson C, Eadie J, Asadpour K, Alavi N. Examining the role of AI technology in online mental healthcare: opportunities, challenges, and implications, a mixed-methods review. *Front Psychiatry* [Internet]. 2024 [acesso 11 ago 2025];15. DOI: 10.3389/fpsy.2024.1356773

Sophia Ivantes Rodrigues – Graduanda – sophiaivantes@hotmail.com

 0009-0008-9967-6049

Marília da Mata Silva – Doutoranda – marilia.mata1991@gmail.com

 0000-0001-7346-2386

Leonardo Pestillo de Oliveira – Doutor – leopestillo@gmail.com

 0000-0001-5278-0676

Lucas França Garcia – Doutor – lucasfgarcia@gmail.com

 0000-0002-5815-6150

Correspondência

Lucas França Garcia – Programa de Pós-Graduação em Promoção da Saúde, Unicesumar. Av. Guedner, 1.610, bloco 7. CEP 87050-390. Maringá, Paraná, Brasil.

Contribuições dos autores (CRediT)

Sophia Ivantes Rodrigues foi responsável pela concepção do estudo, coleta de dados e redação do manuscrito. Marília da Mata Silva contribuiu com a concepção do estudo. Lucas França Garcia participou da concepção do estudo e redação do manuscrito. Leonardo Pestillo de Oliveira participou da redação do manuscrito. Todos os autores fizeram análise dos dados e revisão crítica do manuscrito.

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

Editora responsável: Dilza Teresinha Ambrós Ribeiro

Recebido: 25.2.2025

Revisado: 11.8.2025

Aprovado: 22.8.2025