

SAÚDE MENTAL

Desafios da Prevenção, Diagnóstico, Tratamento e
Cuidado na Sociedade Moderna

Edição XXX

Capítulo 59

INTELIGÊNCIA ARTIFICIAL NA PRÁTICA PSIQUIÁTRICA: APLICAÇÕES CLÍNICAS E DESAFIOS ÉTICOS

GABRIELA UCHÔA DA SILVEIRA¹
JÚLIA STRANO RODRIGUES¹
LUÍSA NARDIN GUTSCHOW¹
NATÁLIA GOLDANI SERRANO¹
NICOLLY PETER FARIAS¹
RAFAEL MACIEL HOMRICH¹
RENATO ASSMANN MACCARI¹

¹Discente – Medicina da Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS).

Palavras-chave: Transtornos Mentais; Diagnóstico Psiquiátrico; Ética em Saúde Digital.

DOI:

10.59290/1803219120

EP EDITORA
PASTEUR

INTRODUÇÃO

Os transtornos mentais representam uma das principais causas de incapacidade em escala global, respondendo por cerca de 32% dos anos vividos com incapacidade e estando associados a maior mortalidade, abuso de substâncias e isolamento social, cenário agravado pela pandemia de Covid-19 (LEE *et al.*, 2021). Estima-se que aproximadamente uma em cada oito pessoas no mundo seja afetada por algum transtorno psiquiátrico, ao passo que o acesso ao tratamento permanece limitado por custos elevados, estigma social, escassez de profissionais especializados e barreiras estruturais dos sistemas de saúde (WANG *et al.*, 2025). Esse descompasso entre necessidade assistencial e capacidade de resposta clínica tem estimulado a investigação de tecnologias capazes de ampliar o alcance e a eficiência do cuidado em saúde mental, entre as quais a inteligência artificial (IA) ocupa posição de destaque.

A incorporação da IA à prática médica já se encontra consolidada em especialidades como oncologia, radiologia e dermatologia, nas quais critérios diagnósticos relativamente objetivos e a disponibilidade de grandes bases de imagens favoreceram o desenvolvimento e a validação de algoritmos. Na psiquiatria, entretanto, essa trajetória tem sido mais lenta, e até o momento não existe qualquer ferramenta baseada em IA aprovada por agências regulatórias como o *Food and Drug Administration* (FDA) para uso clínico psiquiátrico (LEE *et al.*, 2021). Essa defasagem reflete particularidades próprias da especialidade: os diagnósticos psiquiátricos apoiam-se predominantemente em manifestações clínicas subjetivas e relatos verbais, sem biomarcadores amplamente estabelecidos; categorias nosológicas como as do DSM-5 frequentemente apresentam sobreposição sintomática entre transtornos distintos;

e os dados clínicos gerados no contexto da saúde mental são especialmente sensíveis do ponto de vista ético e legal (LEE *et al.*, 2021; KOLDING *et al.*, 2024).

Do ponto de vista técnico, as aplicações de IA em saúde mental abrangem abordagens de aprendizado de máquina (*machine learning*), aprendizado profundo (*deep learning*), aprendizado por transferência e modelos híbridos, além do processamento de linguagem natural e, mais recentemente, da IA generativa baseada em grandes modelos de linguagem. Revisões sistemáticas com metanálise indicam desempenho consistente dessas tecnologias, com acurácia diagnóstica agrupada em torno de 85% e eficácia terapêutica agrupada próxima de 84%, sendo os modelos de *machine learning* os que apresentam o melhor desempenho relativo, seguidos pelos modelos híbridos e pelos de *deep learning* (RONY *et al.*, 2025). Essas abordagens têm sido aplicadas a fontes de dados cada vez mais diversas, incluindo prontuários eletrônicos, neuroimagem, sensores vestíveis, padrões de fala e conteúdos de redes sociais, ampliando as possibilidades de caracterização multimodal do funcionamento psíquico para além do encontro clínico tradicional (LEE *et al.*, 2021).

A IA generativa introduziu, ainda, possibilidades específicas de interação direta com pacientes por meio de *chatbots* e agentes conversacionais. Ferramentas como o ChatGPT têm demonstrado capacidade de psicoeducação, respostas empáticas e razoável reconhecimento de conteúdo emocional em textos e imagens, além de auxiliar em tarefas como reestruturação cognitiva, entrevista motivacional e sumarização de sessões (KOLDING *et al.*, 2024; WANG *et al.*, 2025).

Entre jovens de 10 a 24 anos, uma metanálise recente identificou redução estatisticamente significativa, ainda que de magnitude pe-

quena, do sofrimento psicológico associado a intervenções por *chatbot*, sem efeito significativo correspondente sobre o bem-estar psicológico, sugerindo que os ganhos observados concentram-se mais na atenuação de sintomas do que na promoção positiva de saúde mental (LI *et al.*, 2025). Paralelamente, a mesma literatura documenta limitações relevantes dessas ferramentas, como geração de referências científicas inexistentes, desatualização de informações, vieses demográficos e subestimação do risco suicida em casos clínicos mais complexos (KOLDING *et al.*, 2024; WANG *et al.*, 2025).

O conjunto dessas evidências indica que a inteligência artificial possui potencial concreto para qualificar o cuidado em saúde mental, ampliando o acesso, aprimorando processos diagnósticos e prognósticos e personalizando intervenções terapêuticas. Esse potencial, contudo, ainda convive com uma distância relevante entre o desempenho observado em condições experimentais e sua efetiva validação em contextos assistenciais reais, o que impõe cautela quanto à incorporação apressada dessas tecnologias à rotina clínica. É nesse cenário de potencialidades relevantes e desafios ainda não plenamente resolvidos que se insere a proposta deste capítulo.

O objetivo desse estudo é analisar criticamente, com base na literatura científica atual, as aplicações da inteligência artificial na prática psiquiátrica, contemplando suas principais abordagens técnicas, seu desempenho em diagnóstico, prognóstico e intervenção terapêutica, bem como as limitações metodológicas e os desafios éticos que condicionam sua tradução para a prática clínica real.

MÉTODO

Trata-se de uma revisão de literatura, reali-

zada em junho de 2026, com o objetivo de analisar as principais aplicações da inteligência artificial na prática psiquiátrica, com ênfase em suas possibilidades clínicas e nos desafios éticos relacionados à sua utilização em saúde mental. A busca bibliográfica foi realizada na base de dados PubMed, utilizando-se, de forma isolada e combinada, os termos “*artificial intelligence*”, “*machine learning*”, “*deep learning*”, “*generative artificial intelligence*”, “*generative AI*”, “*chatbot*”, “*mental health*”, “*psychiatry*”, “*psychiatric*” e “*mental health care*”. A estratégia de busca foi estruturada a partir da combinação dos descritores relacionados à inteligência artificial e à saúde mental por meio dos operadores booleanos “AND” e “OR”.

Foram priorizados artigos publicados nos últimos cinco anos, considerando-se o período de 2021 a 2026, sem, contudo, excluir publicações anteriores quando consideradas relevantes para a compreensão e contextualização do tema proposto. Foram incluídas publicações disponíveis no idioma inglês, que abordassem diretamente a utilização da inteligência artificial no contexto da psiquiatria e da saúde mental, contemplando aplicações clínicas, diagnóstico, monitoramento, intervenções terapêuticas, *chatbots*, inteligência artificial generativa e aspectos éticos relacionados ao uso dessas tecnologias. Foram considerados diferentes tipos de estudos, incluindo revisões de literatura, revisões sistemáticas, metanálises e ensaios clínicos, desde que apresentassem contribuições pertinentes ao objetivo da pesquisa.

Após a realização da busca, os artigos foram selecionados de acordo com sua relevância para o tema e submetidos à leitura dos títulos, resumos e, posteriormente, do texto completo. Ao final, foram selecionados oito artigos considerados pertinentes à discussão proposta. A partir da leitura e análise das publicações, as

informações foram organizadas de forma descritiva e agrupadas em categorias temáticas, contemplando as principais aplicações da inteligência artificial na prática psiquiátrica, seus potenciais benefícios clínicos, limitações e os principais desafios éticos relacionados à privacidade e segurança dos dados, vieses algorítmicos, transparência, responsabilidade profissional e segurança dos pacientes.

RESULTADOS E DISCUSSÃO

Caracterização dos estudos analisados

Foram analisados oito artigos publicados entre 2021 e 2026: duas revisões sistemáticas com metanálise, três revisões sistemáticas qualitativas, uma revisão narrativa histórica, uma revisão sobre barreiras, facilitadores e sabedoria artificial e um ensaio clínico randomizado. As publicações contemplaram acurácia diagnóstica, eficácia terapêutica, intervenções com *chatbots*, IA generativa, diagnóstico, monitoramento, governança ética e apoio a psicoterapeutas (LEE *et al.*, 2021; KOLDING *et al.*, 2024; RONY *et al.*, 2025; WANG *et al.*, 2025; LI *et al.*, 2025; CRUZ-GONZALEZ *et al.*, 2025; ZAVALETA-MONESTEL *et al.*, 2026; SADEH-SHARVIT *et al.*, 2023).

Chama atenção, já nessa caracterização inicial, o predomínio de sínteses de evidência sobre estudos primários de intervenção: apenas um ensaio clínico randomizado foi identificado entre as publicações selecionadas. Esse desequilíbrio parece refletir um padrão descrito pelos próprios autores das revisões, segundo os quais a maior parte da produção científica concentra-se em estudos de desenvolvimento e validação de modelos, com escassez de investigações conduzidas em condições reais de assistência (CRUZ-GONZALEZ *et al.*, 2025; ZAVALETA-MONESTEL *et al.*, 2026).

Abordagens técnicas e desempenho global

As abordagens técnicas descritas incluem aprendizado de máquina, aprendizado profundo, aprendizado por transferência, modelos híbridos, processamento de linguagem natural e IA generativa. A metanálise de Rony *et al.* (2025), que reuniu 14 estudos com amostras de 21 a 142.432 participantes, encontrou acurácia diagnóstica agrupada de 85% (IC: 80–87%) e eficácia terapêutica de 84% (IC: 82–86%). Os modelos de machine learning apresentaram o melhor desempenho (0,85), seguidos pelos modelos híbridos (0,84) e de *deep learning* (0,82).

Esses valores precisam ser interpretados com cautela. Entre os 28 estudos que informaram acurácia na revisão de Cruz-Gonzalez *et al.* (2025), os resultados variaram de 51% a 97,54%, demonstrando importante heterogeneidade. Em metanálises anteriores reunidas por Zavaleta-Monestel *et al.* (2026), modelos de neuroimagem alcançaram sensibilidade e especificidade próximas a 80% na esquizofrenia e, na depressão maior, sensibilidade de 77% e especificidade de 78%. Um modelo aplicado à coorte STAR*D obteve acurácia de 64,6% para remissão, reduzida para 59,6% na validação externa ($p = 0,043$). Assim, o desempenho depende da tarefa, da qualidade dos dados e do rigor da validação.

Entre essas abordagens, o processamento de linguagem natural ocupa posição particular na psiquiatria, uma vez que a linguagem e a fala são elementos centrais tanto da anamnese quanto da própria expressão sintomática, o que distingue a especialidade de áreas mais dependentes de dados de imagem (LEE *et al.*, 2021; KOLDING *et al.*, 2024).

Suas aplicações incluem a análise de prontuários e de textos clínicos não estruturados, com extração de sintomas, sinais de risco, dados de funcionamento e trajetórias terapêuti-

cas, além da identificação de comorbidades e da predição de risco de suicídio e de reinternações. A análise da fala amplia esse escopo ao incorporar prosódia, fluência, coerência semântica e expressão afetiva como fontes de informação clínica (LEE *et al.*, 2021; ZAVALETA-MONESTEL *et al.*, 2026).

Aplicações no diagnóstico e na avaliação clínica

No campo diagnóstico, as aplicações mais consistentemente descritas envolvem a diferenciação entre transtornos com apresentações clínicas sobrepostas, como depressão unipolar e bipolar, esquizofrenia, transtornos de ansiedade, TDAH e diferentes subtipos de demência. Além da classificação, a IA tem sido utilizada na identificação de possíveis subtipos fenotípicos a partir da integração de dados genéticos, de neuroimagem e de biomarcadores, o que interessa particularmente a uma especialidade cujas categorias nosológicas são síndromicas e apenas parcialmente ancoradas em mecanismos biológicos definidos (LEE *et al.*, 2021; RONY *et al.*, 2025; ZAVALETA-MONESTEL *et al.*, 2026).

Os modelos utilizaram informações clínicas e sociodemográficas, escalas psicométricas, prontuários eletrônicos, neuroimagem, eletroencefalograma, biomarcadores e características da fala e da linguagem (CRUZ-GONZALEZ *et al.*, 2025). Também foram empregados sensores vestíveis, smartphones e redes sociais, permitindo monitoramento fora da consulta e detecção precoce de alterações, mas ampliando preocupações relativas ao consentimento e à privacidade (LEE *et al.*, 2021).

Predição de risco e prognóstico

A predição de risco de suicídio configura-se como a aplicação mais estudada dentro da literatura sobre IA generativa em saúde men-

tal, com o GPT-4 alcançando acurácia de até 0,96 nessa tarefa e superando modelos tradicionais como o SVM em determinados contextos. Modelos baseados em prontuários eletrônicos também demonstraram, em alguns estudos, desempenho superior à avaliação clínica isolada na estratificação de risco suicida de curto prazo. Cabe ressaltar, entretanto, que boa parte desses estudos utilizou postagens de redes sociais e casos simulados, e não pacientes avaliados prospectivamente em serviços reais (WANG *et al.*, 2025; ZAVALETA-MONESTEL *et al.*, 2026).

O monitoramento longitudinal constituiu um dos campos mais desenvolvidos: dos 85 estudos revisados por Cruz-Gonzalez *et al.* (2025), 40 abordaram monitoramento e 25 buscaram prever resposta ou eficácia terapêutica, o que aproxima a IA do conceito de psiquiatria de precisão. Foram utilizados dados de sono, atividade física, *smartphones*, frequência cardíaca, fala, neuroimagem e informações genômicas. Modelos longitudinais também foram empregados para estimar trajetória da depressão, abuso de substâncias e desfechos funcionais, desde que submetidos a validação interna e externa rigorosa (LEE *et al.*, 2021).

Aplicações terapêuticas e apoio ao profissional

No campo terapêutico, a literatura descreve predição de resposta a antidepressivos, antipsicóticos, eletroconvulsoterapia, estimulação magnética transcraniana repetitiva, cetamina, psilocibina e psicoterapia, com potencial de reduzir tentativas terapêuticas ineficazes e o tempo até a resposta clínica (LEE *et al.*, 2021; CRUZ-GONZALEZ *et al.*, 2025). Ferramentas de IA generativa têm sido testadas ainda no planejamento terapêutico, na formulação de casos clínicos e no apoio à decisão, com planos coerentes com diretrizes e, em algumas

comparações, desempenho superior ao de clínicos gerais (WANG *et al.*, 2025).

O único ensaio clínico randomizado analisado oferece a evidência mais próxima da prática assistencial. Sadeh-Sharvit *et al.* (2023) acompanharam 47 adultos com depressão ou ansiedade, randomizados para tratamento usual ou terapia apoiada por uma plataforma que resumizava sessões, oferecia *feedback* aos terapeutas e auxiliava na documentação. O grupo assistido por IA apresentou frequência 67% maior às sessões, maior redução dos sintomas e notas de evolução concluídas mais rapidamente (14 contra 69 horas), sem diferença na satisfação. O benefício decorreu do apoio ao terapeuta, e não da substituição profissional; contudo, a amostra pequena, o seguimento de dois meses e o centro único limitam a generalização.

Chatbots e interação direta com o paciente

As intervenções conversacionais dirigidas diretamente ao usuário constituem uma linha distinta e apresentam resultados mais modestos. Wang *et al.* (2025) identificaram que cerca de um quarto dos estudos revisados investigava *chatbots* e agentes conversacionais para suporte emocional, com relatos de redução da ansiedade, melhora do engajamento e aumento da autoeficácia, mas também de respostas genéricas, baixa personalização, repetitividade e mecanismos de segurança insuficientes. A revisão sistemática com metanálise de Li *et al.* (2025) incluiu 11 estudos, conduzidos em 12 países, com jovens de 10 a 24 anos. Os resultados demonstraram redução estatisticamente significativa, embora de pequena magnitude, do sofrimento psicológico, sem efeito significativo sobre o bem-estar psicológico. Essa diferença pode ser explicada pelo fato de a maioria das intervenções ter sido desenvolvida

para reduzir sintomas, e não para promover bem-estar positivo.

Nas análises de subgrupo, os *chatbots* baseados em IA apresentaram efeito superior aos sistemas baseados em regras, cujo efeito não foi significativo; as interações multimodais superaram aquelas exclusivamente textuais; e as intervenções entregues por mensagens instantâneas, plataformas já familiares aos jovens, mostraram-se mais efetivas do que as veiculadas por outros meios. Os benefícios concentraram-se em participantes com sintomas clínicos ou subclínicos. Os usuários valorizaram acessibilidade, comunicação empática e ausência de julgamento, mas relataram respostas pouco personalizadas, dificuldade de compreensão e falhas técnicas (LI *et al.*, 2025).

Vale registrar que Cruz-Gonzalez *et al.* (2025) advertem que os resultados obtidos com *chatbots* tradicionais não devem ser extrapolados automaticamente para sistemas generativos, sobre os quais ainda não há evidência suficiente de eficácia terapêutica. Essa ressalva é importante porque a maior parte da experiência acumulada com intervenções conversacionais provém de sistemas anteriores aos grandes modelos de linguagem, tecnologicamente distintos daqueles hoje disponíveis ao público.

Desempenho da IA generativa e percepções de usuários e profissionais

Os estudos que avaliaram ferramentas como o ChatGPT descreveram desempenho desigual. Foram observadas respostas claras e empáticas, boa capacidade de psicoeducação e reconhecimento emocional; em contrapartida, ocorreram referências inexistentes, informações desatualizadas, vieses, subestimação do risco suicida e recomendações potencialmente perigosas em casos complexos, inclusive em

temas como uso de substâncias (KOLDING *et al.*, 2024; WANG *et al.*, 2025). A fluidez das respostas, portanto, não equivale a confiabilidade clínica.

Quanto às percepções, pacientes valorizam o suporte contínuo e sem julgamento e relatam melhora na autoconfiança, ao mesmo tempo em que notam inconsistências e conteúdos inadequados. Psiquiatras, por sua vez, reconhecem o potencial de redução de tarefas burocráticas, mas manifestam preocupação com a privacidade dos dados e com o fato de muitos pacientes consultarem a IA antes de procurar atendimento médico (KOLDING *et al.*, 2024).

Limitações metodológicas

As principais fragilidades metodológicas foram baixa reprodutibilidade, descrição insuficiente dos algoritmos e dados de treinamento, amostras pequenas, heterogeneidade, ausência de avaliação pós-implementação e predominância de cenários simulados. Também foram identificados *overfitting*, *data leakage*, variação entre centros e dificuldade de calibração (KOLDING *et al.*, 2024; RONY *et al.*, 2025; WANG *et al.*, 2025; ZAVALETA-MONESTEL *et al.*, 2026).

O dado mais eloquente a esse respeito provém de uma revisão de 584 estudos de modelos preditivos individualizados citada por Zavaleta-Monestel *et al.* (2026): apenas 10,4% apresentavam validação interna, 4,6% validação externa e somente 0,2% haviam progredido para implementação clínica. Esse achado ajuda a compreender por que, apesar de mais de cinco décadas de desenvolvimento técnico, a implementação clínica dessas ferramentas permanece limitada. Parte dessa dificuldade é intrínseca ao objeto: modelos treinados para reproduzir rótulos diagnósticos heterogêneos e parcialmente sobrepostos podem acabar aprendendo padrões da prática clínica local, e

não mecanismos estáveis de doença (ZAVALETA-MONESTEL *et al.*, 2026).

Desafios éticos

Os desafios éticos mais frequentes foram vieses raciais, sociais e de gênero, privacidade de dados sensíveis, opacidade dos modelos, ausência de consentimento e auditoria e indefinição sobre responsabilidade legal. Também se destacaram o risco de atribuição excessiva de autoridade à IA e a dificuldade de reconhecer crises suicidas. Kolding *et al.* (2024) sugerem evitar o termo “alucinação” para designar erros dos sistemas, devido ao possível efeito estigmatizante.

Em saúde mental, o uso de dados de redes sociais e *digital phenotyping* tensiona a fronteira entre cuidado e vigilância (ZAVALETA-MONESTEL *et al.*, 2026). As estratégias propostas incluem supervisão humana, sistemas híbridos, encaminhamento diante de risco, diversificação das bases, ajuste multilíngue, verificação factual, auditorias independentes e uso de fontes clínicas confiáveis. Os arcabouços GenAI4MH e TRIPOD-LLM podem contribuir para padronizar a qualidade ética e metodológica das pesquisas (KOLDING *et al.*, 2024; WANG *et al.*, 2025).

Resumo dos achados

Os resultados analisados apontam para uma assimetria persistente entre desempenho algorítmico e utilidade clínica. Há evidências de que modelos de IA podem classificar transtornos, estratificar riscos e prever respostas terapêuticas com desempenho superior ao acaso e, em situações específicas, superior à avaliação clínica isolada. Há também evidência de que ferramentas de apoio ao profissional podem melhorar desfechos assistenciais concretos, como demonstrado no único ensaio randomizado analisado. Por outro lado, a evidên-

cia de que essas tecnologias funcionam de modo seguro, equitativo e generalizável em serviços reais permanece escassa, e praticamente inexistente no caso das ferramentas generativas de uso direto pelo paciente em situações de crise.

Nesse sentido, o eixo interpretativo mais útil para a prática talvez não seja perguntar o quanto a IA acerta, mas em que condições seus acertos se convertem em decisões clínicas melhores. Os achados aqui reunidos sugerem que isso ocorre com maior clareza quando a tecnologia é posicionada como instrumento de apoio, organizando informação, ampliando o monitoramento entre consultas e reduzindo a carga administrativa, e sob supervisão profissional contínua, e não quando é investida de autoridade diagnóstica ou terapêutica autônoma.

Essa leitura converge com a proposta de Lee *et al.* (2021), para quem os sistemas atuais, ainda que eficientes em tarefas cognitivas específicas, carecem de atributos centrais do cuidado em saúde mental, como empatia, compaixão, moralidade e autorreflexão. Os autores defendem a evolução em direção a uma “sabedoria artificial”, na qual princípios éticos e maior sensibilidade humana seriam incorporados ao próprio desenvolvimento dessas tecnologias. Tal horizonte desloca o critério de sucesso da acurácia isolada para a capacidade de sustentar um cuidado seguro, personalizado e centrado no paciente.

CONCLUSÃO

As evidências apresentadas ao longo deste capítulo mostram que a inteligência artificial vem ocupando um espaço cada vez mais relevante na psiquiatria e na saúde mental. Suas

aplicações abrangem o apoio ao diagnóstico e ao prognóstico, a identificação de padrões clínicos, a personalização de intervenções, o monitoramento de pacientes e a redução de tarefas administrativas dos profissionais. Tecnologias baseadas em *machine learning*, processamento de linguagem natural e, mais recentemente, inteligência artificial generativa têm apresentado resultados promissores, inclusive em intervenções voltadas à redução de sintomas e no suporte à tomada de decisão clínica. Entretanto, esses avanços ainda coexistem com limitações importantes relacionadas à validação dos modelos em diferentes populações e contextos, à presença de vieses, à dificuldade de interpretação de algumas decisões algorítmicas e a questões de privacidade e segurança dos dados.

Dessa forma, a incorporação da inteligência artificial à prática psiquiátrica deve ocorrer de maneira criteriosa, priorizando seu papel como ferramenta de apoio ao profissional e não como substituta do julgamento clínico ou da relação terapêutica. Para que os benefícios observados em contextos experimentais possam ser efetivamente traduzidos para a assistência, são necessários novos estudos prospectivos e realizados em condições reais de cuidado, com populações mais amplas e diversas, que avaliem não apenas o desempenho dos modelos, mas também seu impacto sobre os desfechos clínicos, sua segurança, equidade, aceitabilidade e integração aos serviços de saúde. O avanço da inteligência artificial em psiquiatria dependerá, portanto, do desenvolvimento de tecnologias transparentes, clinicamente validadas e eticamente responsáveis, capazes de complementar o cuidado humano preservando a autonomia, a segurança e as necessidades dos pacientes.

REFERÊNCIAS BIBLIOGRÁFICAS

- CRUZ-GONZALEZ, P. *et al.* Artificial intelligence in mental health care: a systematic review of diagnosis, monitoring, and intervention applications. *Psychological Medicine*, v. 55, 2025. doi: 10.1017/S0033291724003295.
- KOLDING, S. *et al.* Use of generative artificial intelligence (AI) in psychiatry and mental health care: a systematic review. *Acta Neuropsychiatrica*, v. 37, 2024. doi: 10.1017/neu.2024.50.
- LEE, E.E. *et al.* Artificial intelligence for mental healthcare: clinical applications, barriers, facilitators, and artificial wisdom. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, v. 6, p. 856, 2021. doi: 10.1016/j.bpsc.2021.02.001.
- LI, J. *et al.* Chatbot-delivered interventions for improving mental health among young people: a systematic review and meta-analysis. *Worldviews on Evidence-Based Nursing*, v. 22, 2025. doi: 10.1111/wvn.70059.
- RONY, M.K.K. *et al.* Artificial intelligence in psychiatry: a systematic review and meta-analysis of diagnostic and therapeutic efficacy. *Digital Health*, v. 11, 2025. doi: 10.1177/20552076251330528.
- SADEH-SHARVIT, S. *et al.* Effects of an artificial intelligence platform for behavioral interventions on depression and anxiety symptoms: randomized clinical trial. *Journal of Medical Internet Research*, v. 25, e46781, 2023. doi: 10.2196/46781.
- WANG, X. *et al.* The application and ethical implication of generative AI in mental health: systematic review. *JMIR Mental Health*, v. 12, e70610, 2025. doi: 10.2196/70610.
- ZAVALETA-MONESTEL, E. *et al.* Artificial intelligence in psychiatry: five decades of progress and persistent translational challenges. *Psychiatric Research And Clinical Practice*, 2026. doi: 10.1176/appi.prcp.20260059.