

12 - 1 | 2024

Inteligência Artificial e Depressão: revisão sistemática

Artificial Intelligence and Depression: a systematic review

Júlio Gonçalves

Versão eletrónica

URL: <https://revistas.rcaap.pt/uiips/> ISSN: 2182-9608

Data de publicação: 05-08-2024 Páginas: 12

Editor

Revista UI_IPSantarém

Referência eletrónica

Gonçalves, J. (2024). Inteligência Artificial e Depressão: revisão sistemática. *Revista da UI_IPSantarém*. 12(1), e33936.
<https://doi.org/10.25746/ruiips.v12.i1.33936>

INTELIGÊNCIA ARTIFICIAL E DEPRESSÃO: REVISÃO SISTEMÁTICA

Artificial Intelligence and Depression: a systematic review

Júlio Gonçalves

Centro Universitário de Brusque (UNIFEBE), Brasil

contato@psicojulio.com

RESUMO

A revolução digital tem vindo a alterar as práticas em saúde de forma massiva, de modo a que a utilização de tecnologias como Inteligência Artificial (IA) integra a rotina da prática profissional em saúde há alguns anos, incluindo o Transtorno Depressivo Maior (TDM). Nesse sentido, a presente pesquisa tem como objetivo verificar o estado das investigações sobre a aplicação da IA na depressão. Para cumprir este objetivo, foi conduzida uma revisão sistemática nas bases de dados ACM, IEEE, PubMed, PsycINFO, SciELO, WoS, Scopus e BVS. Foram excluídos: a) artigos teóricos; b) sem aplicação tecnológica; c) sem relação com depressão. A maioria dos estudos incidiu sobre tecnologias para deteção da depressão (n=17). O algoritmo mais utilizado foi o Support Vector Machine (n=13). A média de desempenho dos algoritmos foi de 88,48 ($\Sigma=10,64$). A aplicação de IA na depressão tem sido promissora; contudo, são necessários mais estudos para refinar o desenvolvimento das tecnologias, bem como explorar a sua eficácia em diferentes contextos clínicos e populacionais.

Palavras-chave: Inteligência Artificial, Depressão, Computação.

ABSTRACT

The digital revolution has been massively altering healthcare practices, so much so that the use of technologies such as Artificial Intelligence (AI) has been integrated into the routine of professional healthcare practice for some years now, including Major Depressive Disorder (MDD). In this regard, the present research aims to assess the status of investigations into the application of AI in depression. To achieve this objective, a systematic review was conducted on the databases ACM, IEEE, PubMed, PsycINFO, SciELO, WoS, Scopus, and BVS. Exclusions were made for: a) theoretical articles; b) those without technological application; c) those unrelated to depression. The majority of studies focused on technologies for depression detection (n=17). The most commonly used algorithm was the Support Vector Machine (n=13). The average performance of the algorithms was 88.48 ($\Sigma=10.64$). The application of AI in depression has shown promise; however, more studies are needed to refine the development of technologies, as well as to explore their effectiveness in different clinical and population contexts.

Keywords: Artificial intelligence; Depression; Computing.

1 INTRODUÇÃO

O Transtorno Depressivo Maior (TDM) é caracterizado por um humor deprimido durante a maior parte do dia, anedonia, sentimentos de inutilidade e/ou culpa, ideação suicida, entre outros sintomas (American Psychiatric Association, 2023), com alta relevância em termos de saúde pública, já que detém: a) alto risco de recaída (Wang, 2004); b) relação direta com o suicídio (Omary, 2021); c) influencia o estado geral da saúde (Barlow & Durand, 2016); d) afeta taxas de mortalidade; e) é fator de risco para múltiplos problemas ao longo do desenvolvimento (Bevan Jones et al., 2018).

Tratamentos psicológicos/psiquiátricos são efetivos na remissão do quadro depressivo (Barlow & Durand, 2016), no entanto, um estudo longitudinal de seis (06) anos conduzido por Wang (2004), foi constatado que 49,8% dos pacientes tratados desenvolveram episódios subsequentes do transtorno, de modo que essa alta frequência de recaída aumenta as chances de suicídio nas manifestações severas do transtorno (Omary, 2021). Tais dados justificam a necessidade da identificação dos fatores de risco (e. g., cognições, eventos traumáticos) para que seja possível a elaboração de estratégias de prevenção para grupos ou de base populacional.

Por outro lado, a revolução digital tem vindo a alterar as práticas em saúde de forma massiva, de modo a que a detecção inteligente, análise de big data, inteligência artificial, computação em nuvem, controlo automático e execução autónoma e robótica já compõem a rotina da prática profissional em saúde há alguns anos (Lotttemberg, Silva & Klajner, 2019). Nas rotinas médicas, o uso de smartphones, dispositivos vestíveis, neuroimagem e redes sociais contribuem para atendimentos clínicos e pesquisas, através da produção e acesso a uma grande quantidade de dados (Renn, Schurr, Zaslavsky & Pratap, 2021).

A Inteligência Artificial (IA) integra o conjunto de tecnologias que justificam a Revolução Digital, com o seu início em 1950, a partir da união de vários cientistas, matemáticos e filósofos, dando início às discussões sobre a criação de máquinas que simulam o comportamento humano (Lotttemberg, Silva & Klajner, 2019). Uma das primeiras iniciativas aplicadas à saúde mental foi a Eliza, desenvolvida por Joseph Weizenbaum no laboratório de Inteligência Artificial do MIT entre os anos de 1964 e 1966, criada com a intenção de imitar um psicanalista (Taulli, 2020).

Atualmente, o campo de estudo em IA é amplo e várias outras áreas o integram, como o Aprendizado de Máquina (AM) e a Aprendizagem Profunda (AP). De forma resumida, tais áreas têm o objetivo de “ensinar a máquina a aprender” com os seus erros e corrigir as suas ações com base nesses erros, até alcançar o nível de previsão e/ou resposta de forma precisa (Taulli, 2020). Os principais tipos de aprendizagem de AM são a aprendizagem supervisionada, em que os dados de treino são selecionados e rotulados pelo programador, e a aprendizagem não-supervisionada, em que os dados não são rotulados e o algoritmo reconhece os padrões dos recursos inseridos e define a estrutura dos dados. Já na AP, os algoritmos aprendem diretamente a partir de dados brutos sem orientação humana (Graham et al., 2019).

Nas rotinas diárias, estas tecnologias permitem aos indivíduos o acesso a informações, interação através de redes sociais, operação de sistemas de segurança, planeamento de rotas, comandos de dispositivos digitais, estudos sobre comportamentos do consumidor, entre outros. Estas tecnologias têm-se inserido de modo prevalente nas áreas médicas, contudo, na saúde mental em específico, as tecnologias baseadas em IA atuam numa menor velocidade (Graham et al., 2019). Isto porque psicólogos e psiquiatras são centrados no paciente na sua prática clínica, com maior foco na relação terapêutica e observação direta dos comportamentos, emoções e cognições do paciente. Esta forma de prática clínica faz uso de declarações qualitativas e por meio de registos escritos do paciente, através de notas clínicas, formulários e questionários (Cummings et al., 2013).

Apesar destas especificidades, a IA beneficia a prática clínica em saúde mental, conforme já demonstram múltiplos estudos, tanto em áreas de diagnóstico quanto interventivas (Graham et al., 2019). Um exemplo disso é a Tess (<https://www.x2ai.com/>), mantida por uma combinação de tecnologias, algoritmos e técnicas de aprendizagem de máquina que oferece intervenções de saúde mental com a intenção de reduzir os sintomas de depressão e ansiedade (Fulmer, Joerin, Gentile, Lakerink & Rauws, 2018). Outro exemplo é a ADA (<https://ada.com/>), que foi desenvolvida com objetivo de gerar hipóteses diagnósticas.

Uma revisão recente conduzida por Sijia, Jingping e Lulu (2022) apontou resultados promissores da aplicação de IA no diagnóstico e intervenções psicológicas, com sugestão de estudos direcionados a fenômenos psicológicos específicos, como o TDM. Dessa forma, tendo em vista: 1) a importância de estratégias céleres para a identificação do TDM (Barlow & Durand, 2016; Bevan Jones et al., 2018); 2) a lenta inserção da IA na área da saúde mental (Graham et al., 2019); 3) a escassez de revisões de literatura relacionadas ao tema (Sijia, Jingping & Lulu, 2022); e 4) a necessidade de que estudos sejam conduzidos para estreitar e consolidar as relações entre tais áreas, procedeu-se com uma revisão sistemática da literatura mediada pela seguinte pergunta: Qual o estado das tecnologias digitais, baseadas em IA, utilizadas no tratamento do TDM?

2 MÉTODOS

O presente estudo configura-se como uma revisão sistemática, desenvolvida a partir da pergunta norteadora geral (PG): Qual o estado das tecnologias digitais, baseadas em IA, utilizadas no tratamento do TDM? Os componentes para a elaboração da pergunta de pesquisa da revisão seguiram o acrônimo PICOT, indicado para revisões sistemáticas que visam verificar os efeitos de uma variável independente sobre a dependente. Neste caso, P (População) foram indivíduos com diagnóstico de TDM, I (Intervenções) foi o uso de tecnologias baseadas em técnicas de IA, C (Comparadores) foram o uso de amostras clínicas com condições de saúde distintas, não clínicas ou lista de espera, O (Desfecho) desempenho da IA no que se refere ao diagnóstico e/ou intervenções para o TDM, e T (Tipo de estudo) foram estudos empíricos.

Além da pergunta geral (PG), há ainda quatro (04) perguntas estatísticas (PS) e cinco (05) perguntas específicas (PE). A pergunta geral aborda como os estudos aplicaram IA no tratamento do TDM. As perguntas estatísticas (PS) avaliam as fontes de publicações, a frequência por gênero, os departamentos envolvidos e as publicações por ano. As perguntas específicas referem-se às medidas de avaliação, aos tipos de dados utilizados, aos domínios de aplicação, ao desempenho do modelo de IA, ao tipo de algoritmo aplicado e aos recursos adicionais utilizados no estudo.

2.1 Protocolo e Registro

Utilizou-se como guia de descrição da revisão as diretrizes do Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), sem submissão prévia de protocolo. O mesmo pode ser acessado em <https://bit.ly/3wOghO8>

2.2 Ferramentas e procedimentos de busca

As pesquisas foram realizadas em junho de 2021. Dado que IA e saúde mental abrangem campos interdisciplinares, a busca foi conduzida em bancos de dados de saúde e Tecnologia da Informação (TI), nas bases de dados ACM Digital Library, IEEE Xplore, PubMed (via National Library of Medicine), PsycINFO (APA), SciELO.ORG, Web of Science - Social Sciences Citation Index (SSCI), Scopus (Elsevier) e Biblioteca Virtual em Saúde: BVS (BIREME). Os descritores e operadores booleanos foram definidos após testes em pré-buscas e consistiram em ("Inteligência artificial") OR ("Aprendizagem de Máquina") AND ("Transtorno Depressivo Maior") OR ("Depressão") em português, e suas respectivas traduções em inglês e espanhol. As pesquisas utilizaram palavras-chave conforme descritores em ciências da saúde (DeCS).

2.3 Critérios de Elegibilidade

Em relação aos critérios de elegibilidade, foram incluídos relatos em formato de artigos ou preprints de pesquisas que: a) apresentassem tecnologias descrevendo um método ou aplicação de IA no tratamento do TDM; b) o artigo avaliou o desempenho da IA utilizada. Foram excluídos: a) artigos de revisão e teóricos, textos publicados em congressos e conferências, estudos de caso, relatos de experiência e comentários; b) documentos que não apresentaram a descrição do método ou da

aplicação da tecnologia utilizada; c) que não tinham relação com o diagnóstico ou intervenção do TDM; d) indisponíveis na íntegra e/ou pagos.

2.4 Seleção dos Estudos

Os autores reviram de forma independente os dados extraídos através do software online de uso livre Rayyan® e discutiram os estudos a serem incluídos após as pesquisas independentes. Em casos de discordância, ambos os investigadores reviram os estudos para decidir sobre a sua inclusão ou exclusão.

2.5 Coleta de dados

A extração dos dados foi realizada com o auxílio do software Excel® e tabulada de forma independente, através de um formulário de extração pré-especificado e padronizado, no qual constavam os seguintes itens: metadados do artigo (referência completa, país de afiliação principal do primeiro autor, ano e departamento); tamanho da amostra; objetivos; medidas de avaliação; tipo de dados utilizados para prever resultados do tratamento; domínios de aplicação e desempenho do modelo concebido na previsão de resultados (por exemplo, precisão de classificação, especificidade); tipo de algoritmo de aprendizado de máquina aplicado (por exemplo, AM supervisionado versus não supervisionado, AP).

Quando os estudos reportaram mais de uma medida de precisão de classificação da IA, as medidas foram inicialmente extraídas e aquela que foi reportada de forma mais consistente entre os estudos foi utilizada nas análises quantitativas. Se um estudo relatasse resultados de vários modelos preditivos (por exemplo, usando diferentes métodos de aprendizado de máquina, conjuntos de variáveis preditoras), as proporções de precisão de classificação para todos os modelos relatados eram extraídas para análises qualitativas. Vale ressaltar que uma meta-análise não foi apropriada para esta revisão, dada a variedade de técnicas de IA e tipos de dados utilizados nos estudos identificados.

2.6 Avaliação da qualidade

Para a avaliação do risco de viés, foram utilizadas três (03) ferramentas do National Heart, Lung, and Blood Institute: a) Estudos de Pré e Pós-intervenção sem grupo controle; b) Estudos de Série de Casos; c) Estudos de Caso Controle. As ferramentas estão disponíveis no link <https://bit.ly/3AwyXoD>. As decisões sobre a pontuação foram consensuais entre todos os revisores antes da análise crítica, e os estudos foram classificados conforme o seguinte critério: risco de viés "alto" quando o estudo obteve até 0,5 de pontuação "sim"; "moderado" quando obteve pontuação entre 0,5 e 0,7 de "sim"; e "baixo" quando obteve pontuação "sim" superior a 0,7.

3 RESULTADOS

Aplicadas as estratégias de busca descritas, foram identificados 197 documentos. Após a remoção das 16 duplicidades, resultaram em 181 textos para a leitura de títulos e resumos, dos quais 32 documentos foram considerados elegíveis para a leitura na íntegra. Após a leitura e avaliação dos textos completos, seis (6) trabalhos foram excluídos. Destes, dois (2) não estavam relacionados à IA, três (3) eram resumos de congressos e um (1) não se referia ao tratamento do TDM. Assim, 26 trabalhos foram incluídos para a síntese e análise da presente revisão. O diagrama que representa este processo está apresentado na figura 1.

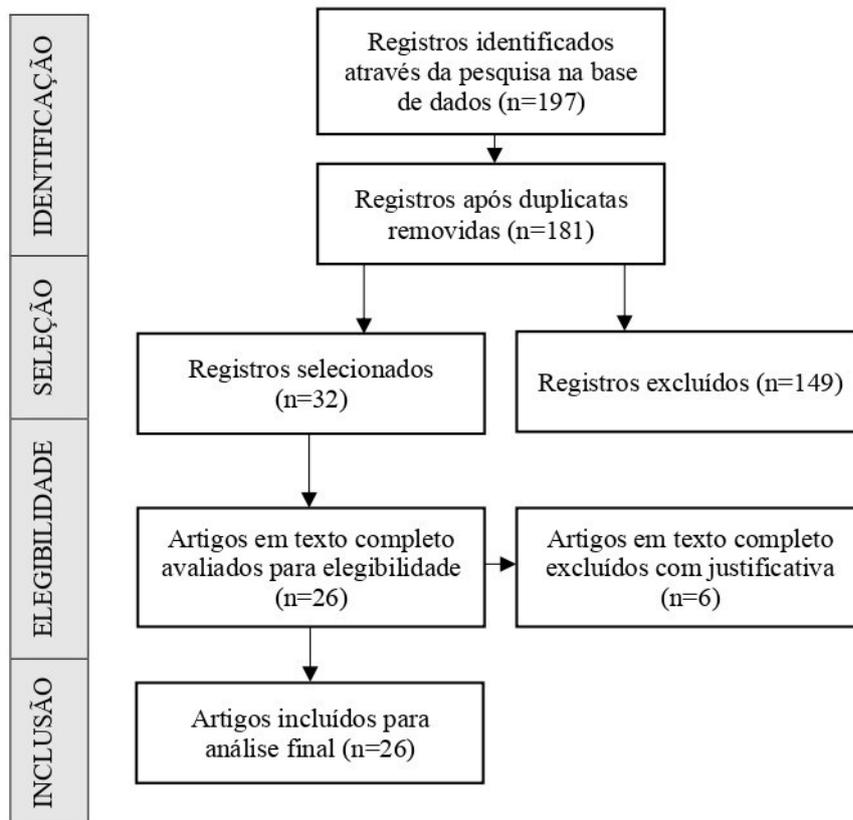


Figura 1 Sistematização da revisão

As informações sobre os motivos de exclusão dos estudos estão disponíveis no seguinte link: <https://bit.ly/3wP6YOU>. Além disso, é possível consultar a tabulação geral dos estudos no seguinte link: <https://bit.ly/38CNhRj>.

3.1 Características da amostra

Os artigos incluídos eram estudos transversais de acordo com o NHLBI (2014), com um total amostral de 11.825 participantes, variando de sete (07) a 5.398 participantes. Com base nos dados disponíveis, a amostra foi composta por 2.276 (19%) indivíduos com TDM e 9.369 (79%) sem TDM. As amostras de mulheres foram maiores (n=3468) do que de homens (n=2455), sendo que o diagnóstico de TDM era mais prevalente em mulheres (n=544). Além disso, em três (03) estudos, foram investigados outros transtornos mentais: Transtorno Bipolar (n=83) e Transtorno do Estresse Pós-Traumático (n=112).

Em termos de frequência de estudos por ano, 2015, 2017 e 2018 foram os anos com maior frequência de pesquisas sobre o tema, com quatro (4) estudos cada, seguido por 2012 e 2021 com três (3) estudos cada, e 2011, 2013, 2019 e 2020, com duas (2) publicações cada.

Os estudos foram conduzidos em departamentos de tecnologia e psicologia, sendo que 49% (n=23) estavam vinculados aos Departamentos de Engenharia, Comunicação e/ou Computação e 51% (n=24) estavam ligados aos Departamentos de Medicina, Neurociências e/ou Psicologia. A maioria dos estudos (n=21) conduziu suas pesquisas de forma interdisciplinar, com a integração de pesquisadores de ambos os departamentos.

Os estudos foram conduzidos, em sua maioria, na China (n=7), seguido pelos Estados Unidos (n=7), Canadá (n=4), Turquia (n=4), Reino Unido (n=3), Alemanha (n=2), Malásia (n=2), Arábia Saudita

(n=1), Austrália (n=1), Brasil (n=1), Holanda (n=1), Irã (n=1), Irlanda (n=1), Japão (n=1) e Portugal (n=1).

3.2 Domínio de Aplicação, Técnicas de IA e Desempenho

Os estudos concentraram-se em três (3) objetivos relacionados ao TDM: detecção, discriminação e previsão da eficácia do tratamento. A categoria "detecção" abrange artigos com o objetivo de identificar e diagnosticar o TDM, enquanto a "discriminação" envolve a distinção do TDM de outros transtornos mentais e, por fim, a previsão da eficácia do tratamento inclui artigos nos quais a IA prevê a eficácia de determinado tratamento para o TDM. A maioria dos estudos foi sobre tecnologias para detecção do TDM (n=17), seguida por tecnologias dedicadas à discriminação do TDM (n=5) do Transtorno Bipolar e do Transtorno do Estresse Pós-Traumático, e, por último, por tecnologias voltadas para a previsão da eficácia do tratamento (n=4) para o TDM.

A maioria dos estudos de "detecção" (n=8) utilizou escalas dicotômicas e Likert para gerar dados de treinamento para a AM, seguido por imagens de ressonância magnética (n=6). Quanto aos algoritmos de IA utilizados, destaca-se o Support Vector Machine (SVM) (n=8), seguido pela Artificial Neural Network (ANN) (n=4) e, em menor frequência, a Logistic Regression (LR) (n=2). Em termos de desempenho, obteve-se uma média geral de 85,12 ($\Sigma=12,20$), na qual a LR teve um desempenho médio de 94,05 ($\Sigma=5,02$), a ANN 91,23 ($\Sigma=12,45$) e a SVM 83,06 ($\Sigma=10,06$).

Os estudos de "discriminação" utilizaram, em maior medida, imagens de ressonância magnética (n=3), seguidas por escalas dicotômicas e Likert (n=3). Os algoritmos mais utilizados foram o SVM (n=3) e ANN (n=2). A média geral de desempenho foi de 79,54 ($\Sigma=9,35$), sendo que o desempenho do SVM ($x=82,36$ $\Sigma=8,73$) foi superior ao da ANN ($x=75,30$ $\Sigma=11,73$).

Da mesma forma, os estudos de previsão da eficácia do tratamento utilizaram, em maior frequência, imagens de ressonância magnética (n=3), seguidas por escalas dicotômicas e Likert (n=1). Os algoritmos mais utilizados foram a ANN (n=3), seguida por Mixture of Factor Analysis (MFA) (n=1). O desempenho geral foi de 86,52 ($\Sigma=4,38$), com um desempenho específico da ANN de 88,70 ($\Sigma=0,69$).

3.3 Medidas de Avaliação e Dados de Treinamento

Em relação aos instrumentos utilizados para identificação do TDM, duas categorias foram identificadas: Aplicação de Entrevistas e/ou Escalas (n=16) e Uso de Equipamentos (n=7). Na categoria relacionada às entrevistas e/ou escalas, a Hamilton Depression Rating Scale (n=9) foi utilizada com maior frequência, seguida pela Structured Clinical Interview for DSM Disorders (n=7), Beck Depression Inventory II (n=4) e o Mini International Neuropsychiatric Interview (n=3). Na categoria de Equipamentos, o eletroencefalograma foi o mais utilizado (n=14), seguido por sensores vestíveis (n=2) e sensores de Smartphone (n=2). Alguns instrumentos foram utilizados em menor frequência (n=1), como o Clinical Global Impression Scale-Severity (CGI-S), Ecological Momentary Assessments (EMA), Hospital Depression and Anxiety Scale (HADS), Geriatric Depression Scale (GDS), Ressonância Magnética Nuclear de alta resolução, Estimulação Magnética Transcraniana Repetitiva (EMTr), entre outros.

Os dados de treinamento coletados por esses instrumentos foram divididos em três (03) categorias: neuroanatômicos, psicossociais e orgânicos. A maioria dos estudos fez uso de dados neuroanatômicos (n=16), especificamente imagens de ressonância magnética. Dados psicossociais (n=8) foram utilizados em menor proporção e incluíram atividade física espontânea, nível de qualidade de vida, exposição a eventos traumáticos, estado de humor, interações sociais, exposição a ambientes iluminados, frequência cardíaca, características do sono, atividade eletrodérmica, índice de massa corporal e pressão do pulso. Por fim, apenas 2 estudos usaram dados orgânicos (plasma, urina e DNA).

3.4 Algoritmos de IA e desempenho geral dos modelos

O algoritmo utilizado com maior frequência foi o SVM (n=13), uma abordagem de Aprendizado de Máquina supervisionado, seguido pelo ANN (n=11), uma abordagem de Aprendizado Profundo. A média global de desempenho dos algoritmos foi de 88,48 ($\Sigma=10,64$). Especificamente, as técnicas com maior capacidade preditiva foram Logistic Regression ($x=87,89 \Sigma=9,44$), Artificial Neural Network ($x=86,99 \Sigma=10,48$), Support Vector Machine ($x=82,72 \Sigma=10,43$) e Decision Tree ($x=77,64 \Sigma=14,69$). Outras técnicas foram utilizadas, como Naïve Bayesian, Gaussian Process Classifier, Random Forest, Relevance Vector Machines e XGBoost; no entanto, devido à baixa frequência nos estudos, não foi possível estimar médias e desvios-padrão. Em relação ao tipo de aprendizagem, houve predominância de técnicas baseadas em Aprendizado de Máquina supervisionado (n=28), seguidas pelo Aprendizado Profundo (n=11).

3.5 Risco de viés dos estudos

A qualidade geral dos estudos incluídos obteve uma pontuação média total de 0,79 ($\Sigma=0,13$), considerada baixo risco de viés. Apenas dois estudos (7,7%) preencheram menos da metade ($<0,5$) dos critérios de qualidade avaliados, o que representa alto risco de viés. O risco de viés moderado ($>0,5$ a $0,7$) foi identificado em 10 estudos (38,4%), enquanto o baixo risco de viés ($>0,7$) foi observado em 14 estudos (53,8%). Os escores individuais de cada artigo podem ser visualizados na tabela 1.

Os problemas que justificam o alto e moderado risco de viés estiveram relacionados à amostra, uma vez que nenhum dos estudos incluídos realizou cálculos amostrais ou justificou o tamanho da amostra. Muitos estudos não relataram completamente os resultados dos testes estatísticos, como a omissão de estimativas de efeitos nomeados, valores de p ou medidas de precisão, quando apropriado (como erros padrão ou intervalos de confiança).

Tabela 1

Escore de viés por estudo

Estudos	Viés	Estudos	Viés
Khodayari et al. (2013)	0,70	Mumtaz et al. (2017)	0,80
Erguzel et al. (2015)	0,70	Schnyer et al. (2017)	0,82
Gao et al. (2017)	0,60	Fang et al. (2012)	0,82
Erguzel et al. (2015)	0,60	Zhu et al. (2017)	0,82
Delahunty et al. (2018)	0,40	Guo at al. (2012)	0,50
Burns et al. (2011)	0,70	Saeedi at al. (2021)	0,60
Pedrelli et al. (2020)	0,70	Gaoliang at al. (2018)	0,82
Erguzel et al. (2015)	0,70	Mwangi et al. (2012)	0,82
Mikus et al. (2018)	0,60	Mumtaz et al. (2015)	0,80
Khodayari et al. (2011)	0,60	Oliveira et al (2013)	0,82
Qian et al. (2019)	0,80	Zhang et al. (2019)	1,00
Tao et al. (2021)	0,80	Ucuz et al. (2020)	1,00
Zahan et al. (2018)	0,78	Nemesure et al. (2021)	0,88

Fonte: Autor (ano)

Os problemas que justificam o alto e moderado risco de viés relacionaram-se à amostra, uma vez que nenhum dos estudos incluídos realizou cálculos amostrais ou justificou os seus tamanhos amostrais. Muitos estudos não relataram totalmente os resultados de testes estatísticos, omitindo, por exemplo, estimativas de efeitos nomeadas, valores de p ou medidas de precisão, quando apropriado, como erros padrão ou intervalos de confiança.

4 DISCUSSÃO DE RESULTADOS

Atualmente, as tecnologias baseadas em aprendizagem de máquina aplicadas à psicologia e à psiquiatria concentram-se no diagnóstico, prognóstico, previsão de tratamento e detecção e monitorização de biomarcadores potenciais. A partir disso, o objetivo central da aprendizagem de máquina é gerar procedimentos que apoiem, principalmente, tomadas de decisão entre profissionais de saúde, oferecendo sugestões acerca do diagnóstico, prognóstico futuro, previsões acerca do tratamento, entre outros (Sijia, Jingping & Lulu, 2022). A presente revisão demonstrou que tais aplicações se enquadram nas pesquisas relacionadas ao TDM, tendo em vista as três (03) categorias encontradas, a saber: detecção do TDM, discriminação do TDM de outra condição de saúde e previsão da eficácia do tratamento para o TDM.

Historicamente, a aproximação precoce da IA com as Neurociências justifica o amplo uso de dados neuroanatômicos, conforme revelado no presente estudo, tendo em vista que um dos objetivos iniciais da IA é simular o comportamento humano, o que, para isso, requer entendimento sobre o funcionamento cerebral (Graham et al., 2019). Ainda, o uso de Tomografia por Emissão de Positrões (PET) ou de Ressonância Magnética (RM) permite uma investigação não invasiva da estrutura e função do cérebro humano na saúde e na patologia, o que permite a identificação de biomarcadores para diagnóstico precoce, planejamento de tratamento e monitorização da doença (Orrù et al., 2012). Outro fator que contribui para o amplo uso de dados neuroanatômicos é que o uso de outros subsídios, como dados genéticos ou linguísticos, é recente, tendo em vista o curto espaço de tempo de exploração da área (Dwyer, Falkai & Koutsouleris, 2018).

Os achados desta revisão são semelhantes aos resultados de revisões sistemáticas recentes baseadas em outros transtornos mentais, como a revisão conduzida por Shatte, Hutchinson e Teague (2019). Os investigadores concluíram que as aplicações de aprendizagem de máquina foram evidentes numa variedade de condições de saúde mental, incluindo depressão (n=88), doença de Alzheimer e outros declínios cognitivos (n=46), esquizofrenia (n=37), stress (n=30) e suicídio (n=20). Os tipos de dados usados para desenvolver modelos de AM incluíram, na sua maioria, dados de imagem (n=102), dados de investigação (n=40), dados de sensores móveis e vestíveis (n=29) e dados de redes sociais (n=28). Na revisão de Orrù et al. (2012), os achados já detinham como prevalência a doença de Alzheimer e a depressão nos seus resultados.

Outro aspeto importante destacado na presente revisão é a prevalência de algoritmos baseados em AM supervisionada (n=28), nos quais os dados são pré-rotulados. O operador define rótulos para os dados de treino (por exemplo, amostra com TDM e amostra não clínica) e o algoritmo aprende a associar esses recursos de entrada (por exemplo, medidas sociodemográficas, clínicas, biológicas etc.), para que, na fase de teste, consiga prever às cegas o rótulo a que uma nova observação pertence. Nesta categoria de aprendizagem supervisionada, a técnica mais comum é a Support Vector Machine (SVM), a mais utilizada na revisão (n=13) e aplicada de forma massiva nos estudos de distúrbios neurológicos e psiquiátricos (Dwyer, Falkai & Koutsouleris, 2018).

Contudo, tendo em vista que as práticas de saúde mental envolvem habilidades, em primeira ordem, de cunho qualitativo, como a observação direta de cognições, emoções e comportamentos, a relação terapêutica e as narrativas históricas, abordagens de processamento e análise da linguagem humana inserem-se como potenciais métodos de aplicação em detrimento às abordagens que utilizam exclusivamente rótulos definidos pelo operador (Cummings et al., 2013; Graham et al., 2019).

Estudos em psicopatologia destacam o comportamento humano como multifatorial, de maneira que transtornos mentais como o TDM, dependem de uma interação complexa entre fatores hereditários, exposição à adversidade no início da vida, personalidade, alterações hormonais e maturacionais, problemas familiares, hábitos de vida e aspectos psicossociais, o que transcende classificações

baseadas em imagens ou escalas e/ou questionários generalistas de sintomas e saúde global (Sheldon et al., 2021).

Uma abordagem em potencial para a resolução deste impasse é a AP, usada com menor frequência na presente revisão (n=11), já que é capaz de lidar de forma mais “humana” com dados de alta dimensão do que modelos de AM supervisionada, incluindo arquivos clínicos e informações de redes sociais para reconhecer fatores como traços de personalidade, eventos traumáticos ou estressantes, ou familiares com histórico de transtornos mentais, e assim por diante (Sijia, Jingping & Lulu, 2022).

Nesta abordagem, os algoritmos de AP aprendem diretamente a partir de dados brutos, aplicando redes neurais artificiais que processam a informação por meio de camadas, semelhante ao modo como o cérebro humano processa informações. As camadas são compostas por nós que avaliam os dados de entrada através de coeficientes (pesos) que reduzem ou amplificam a “força” desses dados em termos de saída. Dadas estas características, a AP é descrita como ideal para elucidar estruturas intrincadas nos seus dados, contudo, a interpretação dos dados de saída tende a ser difícil, já que não há clareza sobre como o algoritmo definiu esses dados (Graham et al., 2019).

Integrada à AP, um método em potencial para a resolução dos impasses referentes à classificação mediada por um operador é o Processamento de Linguagem Natural (PLN), uma subárea da IA que se destaca por utilizar a linguagem humana na forma de texto não estruturado, com uso de tradução de linguagem, compreensão semântica e extração de informações (Maldonado, Alulema, Morocho & Proaño, 2016). Como dados de entrada, são utilizadas notas clínicas, linguagem escrita, sessões transcritas, entre outros, medidas estas que são a base da prática em saúde mental. Desta forma, conceber um algoritmo que reconhece o significado subjacente da linguagem humana, por mais generalista que seja, é um grande avanço científico da IA aplicada às áreas de saúde mental (Graham et al., 2019).

5 CONCLUSÃO

Aplicações de IA na área da saúde estão a aumentar progressivamente, evidenciando a necessidade de os profissionais se familiarizarem com essa modalidade de tecnologia. Apesar da modesta integração da IA na saúde mental, o seu crescimento exponencial é visível, possibilitado pelas características próprias da IA, cujo desenvolvimento acelerado está relacionado ao uso contínuo de algoritmos cada vez mais eficazes. Isso gera um ciclo de soluções cada vez mais refinadas, aumentando a disposição da sociedade em utilizá-las.

Especificamente, a presente revisão revelou aspetos importantes sobre a aplicação da IA no TDM, exigindo atenção: a) um aumento nas pesquisas sobre o tema a partir de 2015; b) os Estados Unidos e a China são os países líderes em pesquisa sobre o tema; c) predominância de estudos voltados para a detecção do TDM; d) imagens de ressonância magnética são as mais utilizadas como dados de entrada; e) abordagens de AM supervisionada são as mais comuns, especialmente técnicas de classificação como o SVM; f) predição média de 88,48 ($\Sigma=10,64$).

Dada a complexidade do TDM e o que tem sido usualmente aplicado em termos de IA, abordagens de AP e PNL são cada vez mais necessárias. O principal desafio da AP é obter uma interpretação fiável dos dados de saída, o que será possível à medida que mais pesquisas sejam conduzidas (Sijia, Jingping & Lulu, 2022). Outro ponto importante é que os algoritmos de IA são desenvolvidos a partir de dados emergentes, ou seja, imagens de ressonância magnética captam apenas um dos fatores preditores da depressão; no entanto, devido ao transtorno ser multifatorial, a investigação biopsicossocial é primordial (Sheldon et al., 2021).

Além disso, apenas uma revisão contou com a participação de pesquisadores brasileiros, embora o estudo não tenha sido conduzido no país (Oliveira, Ladouceur, Phillips, Brammer & Mourao-Miranda, 2013), indicando escassez de estudos sobre o tema no Brasil. Esse atraso é corroborado pelo facto de que apenas recentemente a IA foi incluída como pauta nos assuntos do Ministério da Ciência, Tecnologia e Inovações, através do documento intitulado Estratégia Brasileira de

Inteligência Artificial (EBIA) mantido pela portaria GM n.º 4.617 (Ministério da Ciência Tecnologia e Inovações, 2021).

Ainda, com base no estudo apresentado, algumas limitações da revisão podem ser identificadas. Primeiramente, o viés de seleção de estudos pode ter influenciado os resultados, uma vez que critérios específicos foram estabelecidos para a seleção dos artigos, como a exclusão de estudos teóricos e de revisão, o que poderia restringir a diversidade de perspectivas consideradas na análise. Além disso, a busca foi conduzida em bancos de dados específicos e em três idiomas (português, inglês e espanhol), o que pode ter excluído estudos relevantes publicados em outras línguas ou em fontes não contempladas na busca. Ademais, a diversidade de técnicas de Inteligência Artificial (IA) e tipos de dados utilizados nos estudos pode ter dificultado a comparação e a síntese dos resultados de forma quantitativa, o que pode limitar a generalização dos achados. Por fim, a distribuição geográfica dos estudos incluídos foi desigual, com a maioria dos estudos conduzidos em alguns países específicos, o que pode afetar a representatividade dos resultados em contextos globais. Estas limitações destacam áreas que podem ser aprimoradas em futuras revisões sistemáticas para fornecer uma análise mais abrangente e robusta sobre o uso da IA no tratamento do TDM.

Por fim, a necessidade de pesquisa sobre os temas desta revisão é endossada pelas tendências futuras relacionadas à aplicação de IA na saúde mental, conforme revelado por estudos e comentários recentes, como os de Aktan, Turhan e Dolu (2022) e Kim, Jones e D'Angelo (2019), que destacam a preferência por uma psicoterapia baseada em IA devido à possibilidade de abordar assuntos embaraçosos confortavelmente, acessíveis a qualquer momento e lugar, e reduzindo estigmas. No entanto, os impactos relacionados ao uso excessivo de tecnologias necessitam de investigação, pois estudos sobre intervenções tecnológicas em saúde sem base empírica adequada tendem a reforçar problemas para os quais as intervenções são projetadas para solucionar. Estas divergências em relação aos impactos versus benefícios serão passíveis de resolução à medida que mais pesquisas forem conduzidas.

6 REFERÊNCIAS

- Aktan, T., & Dolu, R. (2022). Attitudes and perspectives towards the preferences for artificial intelligence in psychotherapy. *Computers in Human Behavior*, 133. <https://doi.org/10.1016/j.chb.2022.107273>
- American Psychiatric Association. (2023). *Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-5-TR)*. Artmed.
- Barlow, D. H. & Durand, V. M. (2016). *Psicopatologia: uma abordagem integrada (2ª ed.)*. Cengage Learning.
- Bevan Jones, R., Thapar, A., Stone, Z., Thapar, A., Jones, I., Smith, D., & Simpson, S. (2018). Psychoeducational interventions in adolescent depression: A systematic review. *Patient Education and Counseling*, 101, 804–816. <https://doi.org/10.1016/j.pec.2017.10.015>
- Cummings, C. M., Caporino, N. E., Settapani, C. A., Read, K. L., Compton, S. N., March, J., ... Kendall, P. C. (2013). The therapeutic relationship in cognitive-behavioral therapy and pharmacotherapy for anxious youth. *Journal of Consulting and Clinical Psychology*, 81(5), 859–864. <https://doi.org/10.1037/a0033294>
- Duckworth, A. L., Quirk, A., Gallop, R., Hoyle, R. H., Kelly, D. R., & Matthews, M. D. (2019). Cognitive and noncognitive predictors of success. *Proceedings of the National Academy of Sciences, USA*, 116(47), 23499–23504. <https://doi.org/10.1073/pnas.1910510116>
- Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine Learning Approaches for Clinical Psychology and Psychiatry. *Annual Review of Clinical Psychology*, 14(1), 91–118. <https://doi.org/10.1146/annurev-clinpsy-032816-045037>
- Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2018). Using Psychological Artificial Intelligence (Tess) to relieve symptoms of Depression and Anxiety: Randomized Controlled Trial. *JMIR Mental Health*, 5(4). <https://doi.org/10.2196/mental.9782>
- Graham, S., Depp, C., Lee, E.E., Nebeker, C., Tu, X., Kim, H., & Jeste D. V. (2019). Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. *Current Psychiatry Reports*, 21(11). <https://doi.org/10.1007/s11920-019-1094-0>

- Kim, J. W., Jones, K. L., & D'angelo, E. (2019). How to Prepare Prospective Psychiatrists in the Era of Artificial Intelligence. *Academic Psychiatry*, 43, 337-339. <https://doi.org/10.1007/s40596-019-01025-x>
- Lottemberg, C., Silva, P. E., & Klajner S. (2019). *A revolução digital na saúde: como a inteligência artificial e a internet das coisas tornam o cuidado mais humano, eficiente e sustentável*. Editora dos Editores.
- Maldonado, M., Alulema, D., Morocho, D., & Proaño, M. (2016). System for monitoring natural disasters using natural language processing in the social network twitter. *IEEE International Carnahan Conference on Security Technology*. <https://doi.org/10.1109/CCST.2016.7815686>
- Ministério da Ciência, Tecnologia e Inovações. (2019). *Estratégia Brasileira de Inteligência Artificial (EBIA)* Portaria GM n.º 4.617. https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivos/inteligenciaartificial/ebia-documento_referencia_4-979_2021.pdf
- Oliveira, L., Ladouceur, C. D., Phillips, M. L., Brammer, M., & Mourao-Miranda, J. (2013). What does Brain Response to Neutral Faces Tell Us about Major Depression? Evidence from Machine Learning and fMRI. *PLoS ONE*, 8(4). 2022. <https://doi.org/10.1371/journal.pone.0060121>
- Omary, A. (2021). Predictors and Confounders of Suicidal Ideation and Suicide Attempts among Adults with and without Depression. *Psychiatric Quarterly*, 92, 331–345. <https://doi.org/10.1007/s11126-020-09800-y>
- Orrù, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G., & Mechelli, A. (2012). Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. *Neuroscience & Biobehavioral Reviews*, 36(4), 1140-1152. <https://doi.org/10.1016/j.neubiorev.2012.01.004>
- Renn, B. N., Schurr, M. S., Zaslavsky, O., & Pratap, A. (2021). Artificial Intelligence: An Interprofessional Perspective on Implications for Geriatric Mental Health Research and Care. *Frontiers in Psychiatry*, 12. <https://doi.org/10.3389/fpsy.2021.734909>
- Shatte, A. B. R., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: a scoping review of methods and applications. *Psychological Medicine*, 49(9), 1426-1448. <https://doi.org/10.1017/S0033291719000151>
- Sheldon, E., Simmonds-Buckley, M., Bone, C., Mascarenhas, T., Chan, N., ... Barkham, M. (2021). Prevalence and risk factors for mental health problems in university undergraduate students: A systematic review with meta-analysis. *Journal of Affective Disorders*, 287, 282-292. <https://doi.org/10.1016/j.jad.2021.03.054>
- Taulli, T. (2020). *Introdução à Inteligência Artificial: uma Abordagem Não Técnica*. Novatec.
- Wang, J. (2004). A longitudinal population-based study of treated and untreated major depression. *Medical Care*, 42(6), 543-550. <https://doi.org/10.1097/01.mlr.0000128001.73998.5c>