

APLICABILIDADE DA INTELIGÊNCIA ARTIFICIAL NA PSIQUIATRIA: UMA REVISÃO DE ENSAIOS CLÍNICOS

APPLICABILITY OF ARTIFICIAL INTELLIGENCE IN PSYCHIATRY: A REVIEW OF CLINICAL TRIALS

Resumo

A inteligência artificial (IA), possibilitada pelo avanço tecnológico e aperfeiçoamento dos sistemas computadorizados, visa permitir que as máquinas mimetizem capacidades humanas com mais rapidez e acurácia e tem se mostrado uma ferramenta útil e eficiente dentro da área da saúde, representando potencial de aplicabilidade significativa na psiquiatria. O objetivo deste estudo foi apresentar e discutir ensaios clínicos que tenham estudado a IA, especificamente na área de psiquiatria, para esclarecer sua relevância na prática médica. Realizou-se um levantamento de ensaios clínicos publicados a partir do século XXI nas bases de dados PubMed, Biblioteca Virtual de Saúde e MEDLINE com o uso da IA na psiquiatria. Observou-se que os ensaios clínicos utilizando a IA foram realizados principalmente na área de predição de tratamento farmacológico, seguida de outras áreas, como no desenvolvimento de habilidades sociais e análise de alterações estruturais do sistema nervoso central. A aplicação da IA na medicina representa inovações e avanços significativos na área da psiquiatria, no entanto não substitui a avaliação clínica até o presente momento.

Palavras-chave: Inteligência artificial, psiquiatria, ensaios clínicos.

Abstract

Artificial intelligence (AI), enabled by technological advancement and improvement of computerized systems, aims to allow machines to mimic human capacities at increased speed and accuracy. AI has been shown to be as a useful and efficient tool within the health care scenario,

with potentially significant applicability in psychiatry. The objective of this manuscript was to present and discuss the applicability of AI specifically in psychiatry, to help clarify its relevance in the medical practice. In order to do that, a data survey of clinical trials describing the use of AI in psychiatry was conducted, focusing on studies published in the 21st century on the following virtual databases: PubMed, Biblioteca Virtual de Saúde, and MEDLINE. We observed that most of the clinical trials employing AI in psychiatry were conducted to predict pharmacological treatment, followed by other areas such as development of social skills and analysis of structural changes of the central nervous system. The application of AI to medicine represents significant innovations and advances in the field of psychiatry; however, so far, it is not yet a substitute for clinical evaluation.

Keywords: Artificial intelligence, psychiatry, clinical trials.

INTRODUÇÃO

Os transtornos mentais são um dos principais causadores de incapacidade funcional nos dias de hoje. No Brasil, considera-se que, a cada 100 pessoas, 30 tenham ou venham a desenvolver algum transtorno mental¹. O desenvolvimento de quadros psíquicos ou psicossomáticos pode provocar dificuldades nas relações sociais, com reflexo em problemas emocionais ou de relacionamento, de forma isolada ou em combinação², o que implica não apenas na qualidade de vida do paciente, mas, também, na economia pública.

Atualmente, sabe-se que o diagnóstico dentro da área da psiquiatria é realizado por meio de sistemas classificatórios



e descritivos, como a 5ª edição do Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-5) e a 10ª edição da Classificação Internacional de Doenças (CID-10), que determinam de forma clara, objetiva e detalhada os critérios para cada quadro clínico.

Devido à ausência de marcadores biológicos específicos para demonstrar a presença ou ausência dos transtornos mentais, utiliza-se instrumentos diagnósticos por meio de escalas, entrevistas, inventários ou outros para maior precisão diagnóstica³. A dificuldade para a assertividade clínica, entretanto, consiste não apenas na análise comportamental do paciente, uma vez que restringir-se à análise topográfica de comportamentos não oferece informações referentes às variáveis que controlam os mesmos⁴, mas também consiste na necessidade da imparcialidade na avaliação de pacientes psiquiátricos, de modo a serem evitados juízos distorcidos.

O DSM-5, por exemplo, por meio de uma visão mais ampla, permite afirmar que pacientes com um mesmo transtorno mental apresentem traços e dividam comportamentos semelhantes⁴.

Associando sintomas parecidos à inexperiência médica ou à própria complexidade do caso psiquiátrico, existe, ainda, considerável chance de erro diagnóstico, o que resulta em implicações diretas na efetividade do tratamento a ser instituído, inclusive no âmbito farmacológico. Na tentativa de solucionar esse problema, projetos de investigação empírica como o Research Domain Criteria (RDoC) – que, por meio da incorporação genética, neuroimagem e ciência cognitiva, pretende originar uma nova classificação que acompanhe os desenvolvimentos tecnológicos na área da saúde⁵ – aparecem com o objetivo de tornar a psiquiatria mais precisa.

A necessidade de maior assertividade diagnóstica, portanto, tem se tornado imprescindível no campo da saúde mental. De acordo com Lobo⁶, a inteligência artificial (IA) “é um ramo da ciência da computação que se propõe a desenvolver sistemas que simulem a capacidade humana na percepção de um problema, identificando seus componentes e, com isso, resolver problemas e propor/tomar decisões”.

Por meio do *machine learning* (ML), um ramo da IA em que programas de computação avançados têm a capacidade de processar grandes conjuntos de dados

(*big data*) de diferentes fontes e padrões e diferenciá-los entre múltiplas variáveis não selecionadas⁷, realiza-se, cada vez mais, estudos de casos clínicos com o intuito de avaliar a efetividade dessa tecnologia no tratamento de distúrbios psíquicos.

Essa abordagem tem potencial de representar uma alternativa de importante impacto no desenvolvimento tecnológico e na área da saúde, interferindo tanto no trabalho do médico, através da economia de tempo e energia, bem como na captação de detalhes imperceptíveis que possam ser efeitos de inúmeros fatores⁸.

Atualmente, a IA tem sido utilizada em diversas áreas médicas: pneumologia, em casos de hipertensão pulmonar⁹; endocrinologia, em casos de retinopatia diabética¹⁰; dermatologia, em casos de melanoma¹¹; radiologia, de maneira ampla¹²; entre outros.

O presente estudo tem como objetivo a realização de um levantamento de ensaios clínicos que tenham estudado a IA, especificamente na área da psiquiatria, abrangendo sua importância, impacto, vantagens e desvantagens, a fim de esclarecer a relevância da sua instituição na prática médica.

METODOLOGIA

O trabalho desenvolvido seguiu os preceitos estipulados para a realização de uma revisão bibliográfica por meio da busca de ensaios clínicos nas bases de dados PubMed, Biblioteca Virtual em Saúde (BVS) e MEDLINE. Para isso, foram utilizadas as seguintes palavras-chave: *artificial intelligence, psychiatry e clinical trial*.

A seleção dos artigos baseou-se em critérios de inclusão, sendo eles: artigos de acesso gratuito presentes nas bases de dados acima citadas, referentes a ensaios clínicos randomizados em humanos, publicados no século XXI e que tenham avaliado a aplicação da IA no ramo da psiquiatria. Foram excluídos todos os artigos encontrados, mediante busca inicial através das palavras-chave mencionadas, que não tenham preenchido todos os critérios de inclusão citados.

A busca primária levou a um total de 69 publicações, das quais 68 foram encontradas através do PubMed e uma através da BVS. Após análise de cada um deles, 14 estudos foram selecionados para a presente revisão bibliográfica.

RESULTADOS E DISCUSSÃO

A partir da revisão bibliográfica realizada, pôde-se constatar que os algoritmos-base de todas as vertentes da IA foram utilizados com inúmeros propósitos na prática clínica, entre eles o de avaliar variáveis individuais, avaliar a linguagem, detectar alterações estruturais no sistema nervoso central (SNC), desenvolver habilidades sociais em pacientes com diagnósticos psiquiátricos, entre outros. A seguir, serão relatados estudos relacionados a essas vertentes.

IA na avaliação de variáveis

Na psiquiatria, a resposta terapêutica varia de um indivíduo para outro, e atualmente não há marcadores biológicos objetivos que possam prever, de modo assertivo, a eficácia do tratamento medicamentoso¹³. Muitos estudos abordam a interferência do meio ambiente, além do impacto genético, como fatores de importante relevância na resposta terapêutica apresentada.

Uma das áreas mais amplas e mais bem estudadas da IA na prática médica é o seu uso para a predição do tratamento farmacológico por meio da análise de múltiplas variáveis que individualizem a escolha terapêutica para cada paciente.

Sabe-se que cerca de 1/4 dos pacientes com transtorno de déficit de atenção e hiperatividade/impulsividade (TDAH) não mostra melhora com o tratamento inicial à base de psicoestimulantes, tal como o metilfenidato. Kim et al.¹³, em 2015, realizaram um estudo que teve como objetivo o uso do *machine learning* (ML) com essa finalidade. Para avaliar a resposta terapêutica do metilfenidato em pacientes diagnosticados com TDAH, foram aplicados algoritmos capazes de realizar um cruzamento de dados de variáveis incluindo aspectos demográficos, questionários clínicos, envolvimento ambiental/neuropsicológico, informação genética e neuroimagens.

Este estudo foi realizado em quatro etapas, por meio da aplicação de quatro algoritmos distintos. Na primeira etapa, analisaram-se apenas aspectos demográficos. Na segunda, aspectos demográficos e informações clínicas. Na terceira, foram analisadas a segunda etapa e medidas neurofisiológicas. Já na quarta, associaram-se à terceira etapa aspectos genéticos, ambientais e medidas de neuroimagem. Foi possível observar que a agregação

dos aspectos individuais na quarta etapa resultou em maior acurácia dos algoritmos para prever a resposta ao metilfenidato, de tal forma que apenas marcadores neurobiológicos se mostram insuficientes para análise da resposta terapêutica para pacientes com TDAH.

Outros dois estudos com metodologias semelhantes foram publicados no *The Lancet Psychiatry* no ano de 2016. Koutsouleris et al.¹⁴ desenvolveram um método de avaliação por IA que tinha como objetivo prever o sucesso terapêutico de diferentes drogas em pacientes com episódios iniciais de psicose. Para isso, levaram em consideração informações clínicas prévias, variáveis psicossociais, sociodemográficas e psicométricas.

Foi avaliada a recuperação de pacientes após 4 ou 52 semanas de tratamento, sendo que em pacientes com tratamento de 4 semanas a acurácia alcançou um grau de 75%, enquanto em pacientes com 52 semanas foi demonstrada acurácia de 73,5%. O uso de algoritmos demonstrou ser possível prever, com base no histórico prévio do paciente, os fatores associados com insucesso terapêutico e as melhores drogas de escolha para casos individuais.

Já Chekroud et al.¹⁵ reconheceram que a maioria dos pacientes em tratamento por transtorno depressivo não alcançam remissão com a primeira terapia medicamentosa e, por isso, realizaram um estudo, baseado em ML, com o objetivo de avaliar se pacientes com esse diagnóstico alcançariam a remissão de sintomas. Para isso, treinaram o algoritmo com 25 variáveis que foram preditivas do tratamento medicamentoso em populações de dois outros estudos: o *Sequenced Treatment Alternatives to Relieve Depression* (STAR*D), estudo randomizado prospectivo em pacientes ambulatoriais com transtorno depressivo; e o *Combining Medications to Enhance Depression Outcomes* (CO-MED), estudo randomizado duplo-cego controlado com placebo para avaliação da eficácia medicamentosa em pacientes com esse transtorno.

O algoritmo foi aplicado em três grupos de pessoas com esse diagnóstico em uso de diferentes medicações. O grupo 1 utilizava escitalopram e placebo, o grupo 2, escitalopram e bupropiona, e o grupo 3, venlafaxina e mirtazapina. O estudo demonstrou que o algoritmo apresentou significância para os grupos sob uso de escitalopram/placebo e escitalopram/bupropiona, mas não demonstrou significância para o grupo em uso de



venlafaxina e mirtazapina ($p > 0,05$), demonstrando limitação da aplicabilidade da IA, conforme critérios de treinamento do ML.

Em 2017, Bak et al.¹⁶ optaram por identificar subgrupos de pacientes esquizofrênicos em seu primeiro episódio com base em medidas eletrofisiológicas e cognitivas, além da análise da resposta terapêutica em monoterapia com amissulprida nesses mesmos subgrupos. Para isso, o ML foi aplicado à população selecionada (44 pacientes e 53 controles, com características equivalentes de idade, sexo e *status* socioeconômico), avaliando 26 variáveis cognitivas e 19 variáveis eletrofisiológicas.

Como resultado, foram encontrados dois subgrupos, um com 26 pacientes e outro com 18, sugerindo a existência de subgrupos distintos que variam de acordo com o processamento precoce de informações e de funções cognitivas. Posteriormente, aplicou-se a Positive and Negative Syndrome Scale (PANSS) aos subgrupos, com o objetivo de avaliar a resposta terapêutica. Como resultado, o estudo demonstrou uma precisão de 74,3%,

significativamente maior que a basal (54,2%), no que se refere à resposta terapêutica e a diferentes impactos do fármaco de acordo com o subgrupo em questão.

No mesmo ano, Fabian Lenhard¹⁷ optou por aplicar o ML com o objetivo de prever o resultado da abordagem terapêutica cognitiva *internet-delivered* em pacientes pediátricos com transtorno obsessivo-compulsivo. Para isso, foram utilizados quatro algoritmos distintos que avaliaram variáveis coletadas em questionários prévios com abordagens demográficas e clínicas, como sexo, idade, medicamentos em uso, entre outros.

Como resultado, observou-se que os algoritmos foram capazes de prever a resposta terapêutica com boa a excelente acurácia (75-83%), demonstrando ser uma ferramenta que confirma a utilidade dessa terapia em pacientes sem condições socioeconômicas para acesso à terapêutica presencial.

A Tabela 1 apresenta de forma global os estudos relacionados à IA na análise das variáveis abordadas nesta revisão.

Tabela 1 - Estudos com inteligência artificial na avaliação de variáveis

Referências	Diagnóstico psíquico	Objetivo	População	Método	Resultados
Kim et al. ¹³	TDAH	Predizer pacientes com TDAH que terão boa resposta com o uso de metilfenidato	83 pacientes diagnosticados, recrutados do Seoul National University Hospital, na Coreia do Sul, com exclusão de pacientes com outros transtornos psíquicos, uso de estimuladores por mais de 6 meses ou uso recente menor que 4 semanas	Aplicaram-se quatro tipos de <i>machine learning</i> em situações distintas, sendo eles: Support Vector Machine, J48, Random Forest, Logistic Ridge Regression.	Os algoritmos aumentaram o seu grau de acurácia na medida em que foram adicionadas mais variáveis para cruzamento de dados.
Koutsouleris et al. ¹⁴	Psicose	Predizer a eficácia do tratamento de 4 ou 52 semanas em pacientes com primeiro episódio de psicose	1.047 pacientes entre 18-40 anos, com diagnóstico de esquizofrenia pelo DSM-IV, transtorno esquizofrênico-afetivo ou esquizofreniforme	Aplicaram-se algoritmos nos pacientes que demonstraram recuperação em 4 semanas e em 52 semanas.	Maior acurácia em pacientes com recuperação de 4 semanas em comparação com os de 52 semanas, avaliando interferências psicossociais nos resultados. Pacientes com episódios depressivos, sexo masculino e suicídio indicaram baixa recuperação após tratamento de 1 ano.

Continua na próxima página

Tabela 1 - Continua

Referências	Diagnóstico psíquico	Objetivo	População	Método	Resultados
Chekroud et al. ¹⁵	DDM	Desenvolver um algoritmo capaz de prever a resposta terapêutica para remissão de pacientes depressivos sintomáticos	Pacientes de idade entre 18-75 anos, com um diagnóstico prévio de transtorno depressivo maior não psicótico pelo DSM-IV, que tiveram recorrência ou depressão crônica e escore de 16 ou mais na HAM-D	Realizou-se treinamento do algoritmo por meio de validação cruzada de variáveis de pacientes de dois estudos: STAR*D e CO-MED. Posteriormente, aplicou-se o algoritmo em grupos distintos sob uso de medicações distintas.	O algoritmo apresentou significância para grupo sob uso de escitalopram-plaço e escitalopram-bupropiona, mas não mostrou significância para o grupo em uso de venlafaxina-mirtazapina.
Bak et al. ¹⁶	Esquizofrenia	Identificar potenciais subgrupos de pacientes esquizofrênicos com base em medidas eletrofisiológicas e de cognição, além de analisar a resposta terapêutica da monoterapia com amisulprida nos mesmos subgrupos	44 pacientes e 53 controles com características equivalentes de idade, sexo e <i>status</i> socioeconômico	Aplicou-se um algoritmo que avaliou 26 variáveis cognitivas e 19 variáveis eletrofisiológicas para identificação dos subgrupos, e utilizou-se a PANSS para avaliar o tratamento terapêutico.	Identificaram-se dois subgrupos, um com 26 participantes e outro com 18 participantes. A utilização da PANSS nos subgrupos demonstrou uma precisão de 74,3% maior que a precisão basal.
Lenhard et al. ¹⁷	TOC	Predizer os resultados da abordagem terapêutica cognitiva <i>internet-delivered</i> em pacientes pediátricos com TOC	67 pacientes com TOC entre 12-17 anos de idade que participaram de estudo clínico randomizado e receberam abordagem terapêutica cognitiva <i>internet-delivered</i> por 12 semanas	Aplicação da abordagem terapêutica cognitiva <i>internet-delivered</i> na população selecionada. Posteriormente, analisou-se a resposta terapêutica com aplicação de quatro algoritmos distintos.	Os resultados demonstraram a capacidade do <i>machine learning</i> em prever a resposta da abordagem terapêutica com acurácia de até 83%.

CO-MED = Combining Medications to Enhance Depression Outcomes; DDM = distúrbio depressivo maior; DSM-IV = 4ª edição do Manual Diagnóstico e Estatístico de Transtornos Mentais; HAM-D = Escala de Depressão de Hamilton; PANSS = Positive and Negative Syndrome Scale; STAR*D = Sequenced Treatment Alternatives to Relieve Depression; TDAH = transtorno de déficit de atenção com hiperatividade; TOC = transtorno obsessivo-compulsivo.

IA na avaliação da linguagem

Outra linha de estudo refere-se à aplicação do ML para análise do discurso em pacientes com diagnósticos psiquiátricos. Carrillo et al.¹⁸ publicaram um estudo com o objetivo de avaliar se a linguagem analisada por ML poderia prever respostas clínicas de tratamento com psilocibina (substância psicoativa capaz de produzir alterações perceptocognitivas com importância psicoterapêutica e biomédica em pacientes com transtorno depressivo resistente).

Para isso, pacientes foram submetidos a um teste de memória autobiográfico, no qual realizaram-se relatos de lembranças desencadeadas por palavras-chave. Os

relatos foram transcritos e posteriormente analisados pelo ML por meio da presença de palavras de caráter positivo ou negativo. O estudo foi dividido em dois momentos: primeiro, avaliou-se a capacidade do ML em diferenciar os pacientes do grupo controle, enquanto o segundo consistiu no uso do ML para prever a efetividade do tratamento farmacológico.

O tratamento foi realizado com duas doses de psilocibina, com uma posologia total de 35 mg, e considerou-se o paciente responsivo ao tratamento aquele que apresentasse uma redução de pelo menos 50% no Quick Inventory of Depressive Symptoms (QIDS-16). O uso do ML no estudo demonstrou uma acurácia



de 85% na resposta terapêutica entre os pacientes. Demonstrou, ainda, que pacientes que responderam bem ao tratamento tiveram um discurso com menos palavras positivas no início do estudo em comparação com o final, fato que corrobora para a melhor resposta terapêutica nesse grupo e motivo pelo qual o uso das palavras positivas no discurso é uma variável mais sensível em comparação ao uso de palavras negativas, sendo esta uma das características analisadas ao longo do diálogo médico-paciente.

Essa aplicação não se restringiu apenas à rotina clínica. Pestian et al.¹⁹ aplicaram o ML com o objetivo de examinar a linguagem de adolescentes suicidas em departamentos de emergência. Para isso, foram selecionados 60 adolescentes entre março e agosto de 2011, entre 13 e 17 anos, e que apresentaram ideação suicida em departamentos de emergência. Como controles, foram selecionados pacientes ortopédicos, por serem considerados de menor risco para perturbações neurobiológicas entre os pacientes presentes nos departamentos, com exclusão daqueles com retardo mental, história de transtorno de humor

ou histórico familiar de ideação suicida em parentes de primeiro grau.

Os dois grupos foram submetidos a questionários como Columbia Suicide Severity Rating Scale (C-SSRS), Suicidal Ideation Questionnaire – Junior (SIQ-Jr) e Ubiquitous Questionnaire (UQ), sendo que os discursos do UQ foram registrados e analisados por uma matriz estatística que converteu as palavras em valores matemáticos. As palavras ditas pelos grupos eram colocadas em colunas da matriz, sem considerar comentários dos entrevistados nem as questões realizadas. Como resultado, avaliou-se que somente o método com o C-SSRS foi significativo nos achados, demonstrando que a contagem das palavras é importante para a classificação, além de características como velocidade, pausa e espaçamento vocal.

Além disso, esses autores ressaltaram a importância da análise individual e momentânea de cada paciente, impedindo a aplicação tecnológica como ferramenta autônoma para qualquer tomada de decisão ou conclusão frente ao quadro.

Os estudos relacionados à IA na análise da linguagem estão apresentados na Tabela 2.

Tabela 2 - Estudos com inteligência artificial na avaliação da linguagem

Referências	Diagnóstico psíquico	Objetivo	População	Método	Resultados
Carrillo et al. ¹⁸	Transtorno depressivo	Avaliar se a linguagem poderia prever respostas clínicas de tratamento com psilocibina para pacientes com depressão resistente	17 pacientes com transtorno depressivo resistente e 18 pacientes para controle com características semelhantes de idades e sexo	Aplicou-se o algoritmo que avalia a linguagem na diferenciação dos grupos de pacientes e controle, bem como na predição da efetividade do tratamento com psilocibina nos pacientes resistentes a tratamento.	Demonstrou-se que palavras positivas no discurso é uma variável mais sensível para distinguir pacientes do grupo controle e pacientes responsivos e não responsivos ao tratamento farmacológico.
Pestian et al. ¹⁹	Ideação suicida	Analisar o discurso com características suicidas de adolescentes em departamentos de emergência por meio da aplicação de questionários (C-SSRS, SIQ-Jr e UQ)	60 adolescentes entre 13-17 anos com ideações suicidas em departamentos de emergência. Como controles: pacientes ortopédicos do departamento de emergência, com exclusão daqueles com retardo mental, história de transtorno de humor ou histórico familiar de ideação suicida em parentes de primeiro grau	Utilizou-se o algoritmo afim de promover o reconhecimento das palavras colhidas por questionários para a distinção entre o grupo controle e o grupo de pacientes.	O algoritmo apresentou relevância apenas com um dos questionários aplicados.

C-SSRS = Columbia Suicide Severity Rating Scale; SIQ-Jr = Suicidal Ideation Questionnaire – Junior; UQ = Ubiquitous Questionnaire.

IA na análise de alterações estruturais do sistema nervoso central

Como já dito, a dificuldade diagnóstica de muitos transtornos mentais está relacionada à ausência de marcadores biológicos específicos que façam a distinção entre quadros de acometimento psíquico. Como alternativa, estudos emergiram com base no uso de neuroimagem, visando obter melhores resultados terapêuticos.

Estudos recentes demonstraram a interferência do ML em casos de psicose. É estabelecido que indivíduos que cursam com quadros psicóticos possuem alterações cerebrais relacionadas a marcadores neuroanatômicos. Há estudos que demonstram não haver marcadores biológicos capazes de realizar a distinção entre os diagnósticos, mas afirmam que as alterações permitem a estratificação de pacientes a fim de avaliar a resposta terapêutica²⁰.

Outros estudos demonstraram que, por meio do ML, os métodos de imagem possuem alta acurácia na distinção entre pessoas com desordens psíquicas e indivíduos saudáveis. Zhang et al.²¹ publicaram um estudo no BioMed Research International com esse objetivo: a análise da homogeneidade de regiões cerebrais (ReHo) em pacientes diagnosticados com transtorno de ansiedade social e pessoas saudáveis. Foi demonstrado que o uso de algoritmos na análise de neuroimagens apresentou acurácia de mais de 76% na distinção entre os dois grupos.

Na mesma linha de pesquisa, um estudo publicado por Schnack et al.²² foi realizado com o objetivo de avaliar a capacidade de um algoritmo reconhecer padrões da neuroimagem e distinguir pessoas saudáveis de pacientes com transtornos psiquiátricos. Nele, analisaram-se três grupos de pesquisa: um de pacientes saudáveis, um de pacientes com esquizofrenia e um de pacientes com transtorno bipolar.

O resultado demonstrou que o algoritmo apresentou uma acurácia de 86% para diferenciar pacientes esquizofrênicos de todos os outros participantes, superior à acurácia entre pacientes controles e com transtorno bipolar, possuindo valor significativo no diagnóstico diferencial da esquizofrenia, mas sem resultados significativamente promissores para os pacientes com transtorno bipolar.

A fim de prever a eficácia de eletroconvulsoterapia (ECT) em pacientes diagnosticados com transtorno depressivo maior e associar o alívio sintomático com biomarcadores de regiões cerebrais, Redlich et al.²³ utilizaram o ML. Para o estudo, foram recrutados 47 pacientes com diagnóstico prévio de transtorno depressivo maior por meio da Escala de Depressão de Hamilton (HAM-D) e verificados através da 4ª edição do Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-IV).

Entre os 47 pacientes, subdividiram-se dois grupos, sendo que um grupo foi submetido à terapia eletroconvulsiva e ao tratamento medicamentoso, e o segundo, submetido apenas ao tratamento medicamentoso. Como controle, criou-se um grupo de 24 participantes de características compatíveis (idade, sexo e nível educacional). A ECT foi aplicada inicialmente três vezes por semana ou até que o alívio sintomático fosse alcançado. Como consequência, os pacientes submetidos à ECT foram divididos em respondedores e não respondedores, sendo estes últimos os pacientes com menos de 50% de alívio sintomático de acordo com a HAM-D.

Foi avaliada a capacidade do ML em prever a resposta terapêutica com base nas alterações anatômicas das regiões cerebrais relacionadas ao transtorno psíquico, tal como o giro cingulado. Como resultado, observou-se uma acurácia maior de 70% na capacidade de o algoritmo diferenciar pacientes respondedores e não respondedores, com uma sensibilidade alta, porém com baixa especificidade (40-50%).

Entretanto, o achado mais significativo na análise anatômica foi a associação entre o volume do giro cingulado e a resposta à ECT, de modo que quanto maior o giro cingulado no pré-tratamento, maior foi a eficácia terapêutica. Além disso, pacientes submetidos a essa terapia mostraram um aumento do volume da região hipocampal, achado este ausente no grupo sob tratamento exclusivamente medicamentoso, o que parece apontar para a capacidade neuroplástica da ECT.

A Tabela 3 apresenta os estudos relacionados aos estudos apresentados, nesta revisão, envolvendo a IA para análise de alterações estruturais do sistema nervoso central.



Tabela 3 - Estudos com inteligência artificial na análise de alterações estruturais do sistema nervoso central

Referência	Diagnóstico psíquico	Objetivo	População	Metodologia	Resultados
Zhang et al. ²¹	Transtorno de ansiedade social	Analisar a discriminação entre pacientes com transtornos psíquicos e pacientes sem diagnóstico prévio e avaliar quais regiões cerebrais estão envolvidas na patologia	Participantes foram divididos em dois grupos, sendo um com pacientes diagnosticados e outro para pacientes sem patologias	Aplicou-se um algoritmo para análise da neuroimagem em pacientes diagnosticados com transtorno ansioso e bipolar e grupo de controle.	O estudo apresentou uma acurácia de mais de 76%, sendo especificidade de 82,5% para o algoritmo aplicado. Além disso, demonstrou discriminação neuroanatômica, principalmente nas regiões frontal, temporal e occipital.
Schnack et al. ²²	Esquizofrenia e transtorno bipolar	Classificar pacientes com esquizofrenia, transtorno bipolar e saudáveis por meio de algoritmos de neuroimagem que reconhecem a presença de substância cinza	Foram recrutados 198 pacientes, sendo 72 homens, divididos em 3 grupos correspondentes em tamanho, gênero e idade (idade 37±11 anos)	Realizou-se um cruzamento de dados e análise da estrutura anatômica por meio de exames de imagem nos pacientes do estudo pela aplicação do algoritmo.	Confirmou-se melhor resultado com pacientes esquizofrênicos, com acurácia de 86%. A separação de pacientes com bipolaridade e pacientes saudáveis apresentou uma menor acurácia: 67% dos pacientes saudáveis foram corretamente classificados.
Redlich et al. ²³	Transtorno depressivo maior	Prever a eficácia da terapia eletroconvulsiva (ECT) em pacientes com transtorno depressivo maior e identificar regiões cerebrais por biomarcadores associados ao grau de alívio dos sintomas em pacientes tratados com ECT	47 pacientes com diagnóstico de transtorno depressivo maior pela HAM-D, posteriormente subdivididos, e 21 participantes para grupo controle	Aplicou-se o algoritmo para análise de alterações estruturais neuronais em pacientes submetidos à terapia medicamentosa associada à ECT e em pacientes submetidos exclusivamente à terapia medicamentosa.	Encontraram-se alterações anatômicas significativas que predizem a resposta ao tratamento com ECT, principalmente na região do giro cingulado. Além disso, pacientes submetidos à ECT apresentavam aumento no volume hipocampal.

ECT = terapia eletroconvulsiva; HAM-D = Escala de Depressão de Hamilton.

IA no desenvolvimento de habilidades sociais

Outra área de aplicação recentemente muito estudada é o uso da IA para o desenvolvimento das habilidades sociais, principalmente em crianças diagnosticadas com transtorno do espectro autista (TEA).

Yun et al.²⁴, no ano de 2016, avaliaram o treinamento das habilidades sociais em crianças diagnosticadas com TEA por meio de um sistema de intervenção comportamental robótica. Para isso, realizaram um estudo randomizado com crianças de 4 a 7 anos, com QI > 60, e analisaram o seu desempenho social por meio do uso de robôs que

eram capazes de interagir, reconhecer atividade humana e elaborar uma resposta, a fim de estimular a interação comportamental e emocional das crianças em questão.

O estudo, publicado no Journal of the American Medical Association (JAMA), demonstrou que a utilização de robôs melhora a interação do contato olho-olho, além de maior reconhecimento de expressões emocionais, pontos-chave em crianças autistas.

Em 2019, o foco se manteve no desenvolvimento das habilidades sociais. Um estudo randomizado, realizado por um grupo de pesquisadores da Universidade de

Stanford, teve como objetivo a utilização de um óculo baseado em IA para melhoria da socialização de crianças com TEA.

Os óculos eram capazes de realizar reconhecimento facial e de emoções de parceiros sociais da criança e estimulavam, por meio de cores e emojis, a interação da mesma. O estudo observou que pacientes em terapia comportamental associada ao uso dos óculos demonstraram melhora no instrumento Vineland Adaptive Behavioral Scales (VABS-II), que mede a comunicação, vivência diária, socialização e habilidades comportamentais²⁵.

Estudos ainda em curso preveem foco similar. Mengoni et al.²⁶ publicaram no BMJ Open um estudo com o objetivo de analisar a viabilidade de um ensaio clínico baseado no uso de um robô humanoide para o auxílio no desenvolvimento das habilidades sociais em crianças com TEA. Os resultados destes estudos podem corroborar a aplicabilidade da IA no desenvolvimento das habilidades sociais em crianças autistas.

O desenvolvimento de habilidades sociais com uso de IA não se restringiu à baixa faixa etária ou ao diagnóstico de TEA. Wesemann et al.²⁷ realizaram um estudo com o objetivo de avaliar a atitude de soldados da Força Internacional de Assistência à Segurança

no Afeganistão, considerando a possibilidade do diagnóstico de transtorno de estresse pós-traumático (TEPT).

Buscou-se avaliar o conhecimento desses sujeitos acerca do TEPT, a ocorrência de sintomas desse diagnóstico e o estado mental desses indivíduos antes e depois do treinamento pré-admissão e no término de turno de plantão. Para isso, foram selecionados 67 participantes, sendo 31 do grupo controle. O grupo experimental (n = 36) foi treinado por uma plataforma de treinamento interativo (CHARLY), projetado para fornecer habilidades sociais e treinamento de pré-implantação mental como parte da prevenção primária durante 1 dia e meio.

Os resultados foram comparados com o grupo controle submetido ao treinamento rotineiro e demonstraram que o CHARLY se mostrou superior nos quesitos psicoeducação e mudança de atitude. Os achados com base no uso de uma ferramenta tecnológica, baseada em aprendizado adaptativo, demonstraram importante impacto na prevenção primária na realidade de soldados de carreira militar.

Os estudos de IA no desenvolvimento de habilidades sociais analisados, neste trabalho, se encontram apresentados na Tabela 4.

Tabela 4 - Estudos com inteligência artificial no desenvolvimento de habilidades sociais

Referência	Diagnóstico psíquico	Objetivo	População	Método	Resultados
Yun et al. ²⁴	TEA	Avaliar o uso de um sistema de intervenção comportamental robótica no desenvolvimento de habilidades sociais em crianças diagnosticadas com TEA	Crianças com TEA entre 4 e 7 anos, com QI maior que 60	Utilizaram-se robôs capazes de interagir, reconhecer atividade humana e elaborar uma resposta, a fim de estimular a interação comportamental e emocional das crianças em questão	Os resultados mostraram que a interação entre humanos e robôs tiveram efeitos positivos sobre o contato olho-olho e o reconhecimento de expressão emocional das crianças com TEA.
Voss et al. ²⁵	TEA	Promover melhora da sociabilização em crianças com espectro autista	Crianças com TEA, de idade entre 6 e 12 anos, em tratamento de análise comportamental por pelo menos duas vezes por semana e escore maior que 15 no questionário de comunicação social	Utilizou-se um óculos, baseado no <i>machine learning</i> , capaz de realizar reconhecimento facial e de emoções de parceiros sociais da criança e estimular sua interação	Observou-se que pacientes em terapia comportamental associada ao uso dos óculos demonstraram melhora no VABS-II. Além disso, o estudo demonstrou melhora no ECG.

Continua na próxima página



Tabela 4 - Continua

Referência	Diagnóstico psíquico	Objetivo	População	Método	Resultados
Mengoni et al. ²⁶	TEA	Avaliar a viabilidade do uso de um robô humanoide (Kaspar) na intervenção de habilidades sociais com crianças com TEA	40 crianças entre 5 e 10 anos, com TEA confirmado, QI maior que 70 e com compreensão da linguagem inglesa	Submeteu-se um grupo de crianças ao tratamento tradicional, enquanto outro randomizado foi submetido a sessões com Kaspar por 8 semanas	Resultados não publicados.
Wesemann et al. ²⁷	TEPT	Comparar parâmetros como atitude em relação ao TEPT, conhecimento de TEPT, sintomas específicos de TEPT e estado mental atual antes e depois do treinamento pré-admissão e no término de turno de plantão	67 participantes, sendo 36 do grupo experimental e 31 do grupo controle. Os participantes consistiam em soldados da Força Internacional de Assistência à Segurança no Afeganistão	O grupo experimental foi treinado por uma plataforma de treinamento interativo (CHARLY), projetado para fornecer habilidades sociais e treinamento de pré-implantação mental como parte da prevenção primária durante 1 dia e meio e comparado com os participantes sob o treinamento de rotina	Diferença na mudança de atitude entre os grupos experimental e controle ($p = 0,045$) que pode ser explicada por CHARLY. Ambos os grupos mostraram um aumento significativo no estresse após a implantação, com um aumento menor significativo no CHARLY. Para sintomas específicos de TEPT, CHARLY alcançou uma superioridade significativa.

ECG = Emotion Guessing Game; TEA = Transtorno do espectro autista; TEPT = Transtorno do estresse pós-traumático; VABS-II = Vineland Adaptive Behavioral Scales.

CONCLUSÃO

A IA tem sido cada vez mais estudada e aplicada na prática médica. Sua utilização representa avanços significativos na área da psiquiatria, tanto para tratamento quanto para diagnósticos ou desenvolvimento de habilidades sociais.

Entre os quadros psiquiátricos, é possível avaliar estudos com a IA em distintas situações, desde pacientes com diagnósticos estabelecidos como TDAH, transtorno depressivo, esquizofrenia, transtorno obsessivo-compulsivo, transtorno do espectro autista, até quadros isolados de psicose e ideação suicida em situação de emergência.

É importante ressaltar que o ML e seus algoritmos são elaborados por seres humanos e, portanto, passíveis de erros, podendo ter implicações significativas na área da saúde. Por isso, fiscalizações acerca da elaboração dos algoritmos, além da clara definição da finalidade de cada estudo, são necessárias para maior confiabilidade. Novas pesquisas são necessárias para aprimorar os resultados

obtidos por estudos pioneiros, bem como possibilitar a identificação de novas áreas de atuação da IA.

Artigo submetido em 10/02/2020, aceito em 28/02/2020. Os autores informam não haver conflitos de interesse associados à publicação deste artigo.

Fontes de financiamento inexistentes.

Correspondência: Luiza Megid de Oliveira, Campus Londrina, Rua Pará, 966/07, CEP 86010-450, Londrina, PR. Tel.: (43) 99188.1960. E-mail: lmegid@hotmail.com

Referências

1. Brasil, Ministério da Saúde. Saúde mental: o que é, doenças, tratamentos e direitos [Internet]. [cited 2019 Apr 24]. <https://saude.gov.br/saude-de-a-z/saude-mental>
2. Falceto OG, Busnello ED, Bozzetti MC. Validação de escalas diagnósticas do funcionamento familiar para utilização em serviços de atenção primária à saúde. Rev Panam Salud Publica. 2000;7:255-63.

3. Del-Ben CM, Rufino AC, Azevedo-Marques JM, Menezes, PR. [Differential diagnosis of first episode psychosis: importance of optimal approach in psychiatric emergencies]. *Braz J Psychiatry*. 2010;32 Suppl 2:S78-86.
4. Araújo AC, Lotufo Neto F. A nova classificação americana para os transtornos mentais – o DSM-5. *Rev Bras Ter Comp Cogn*. 2014;16:67-82.
5. Insel T. Thomas Mental Health Information: transforming diagnosis [Internet]. 2013 Apr 29 [cited 2019 Apr 25]. www.nimh.nih.gov/about/directors/thomas-insel/blog/2013/transforming-diagnosis.shtml
6. Lobo Luiz Carlos Inteligência artificial, o Futuro da Medicina e a Educação Médica. *Rev Bras Educ Med*. 2018; 42: 44-46.
7. Scott IA. Machine learning and evidence-based medicine. *Ann Intern Med*. 2018;169:44-6.
8. Albu A, Stanciu L. Benefits of using artificial intelligence in medical predictions [Internet]. 2015 [cited 2019 Apr 27]. ieeexplore.ieee.org/document/7391610
9. Dawes TJ, de Marvao A, Shi W, Fletcher T, Watson GM, Wharton J, et al. Machine learning of three-dimensional right ventricular motion enables outcome prediction in pulmonary hypertension: a cardiac MR imaging study. *Radiology*. 2017;283:381-90.
10. Cheung CY, Tang F, Ting DS, Tan GS, Wong TY. Artificial intelligence in diabetic eye disease screening. *Asia Pac J Ophthalmol (Phila)*. 2019;8:158-64.
11. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542:115-8.
12. Saba L, Biswas M, Kuppili V, Cuadrado Godia E, Suri HS, Edla DR, et al. The present and future of deep learning in radiology. *Eur J Radiol*. 2019;114:14-24.
13. Kim JW, Sharma V, Ryan ND. Predicting methylphenidate response in ADHD using machine learning approaches. *Int J Neuropsychopharmacol*. 2015 May 10;18(11):pyv052. doi: 10.1093/ijnp/pyv052.
14. Koutsouleris N, Kahn RS, Chekroud AM, Leucht S, Falkai P, Wobrock T, et al. Multisite prediction of 4-week and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach. *Lancet Psychiatry*. 2016;3:935-46.
15. Chekroud AM, Zotti RJ, Shehzad Z, Gueorguieva R, Johnson MK, Trivedi MH, et al. Cross-trial prediction of treatment outcome in depression: a machine learning approach. *Lancet Psychiatry*. 2016;3:243-50.
16. Bak N, Ebdrup BH, Oranje B, Fagerlund B, Jensen MH, Düring SW, et al. Two subgroups of antipsychotic-naïve, first-episode schizophrenia patients identified with a Gaussian mixture model on cognition and electrophysiology. *Transl Psychiatry*. 2017;7:e1087.
17. Lenhard F, Sauer S, Andersson E, Månsson KN, Mataix-Cols D, Rück C, et al. Prediction of outcome in internet-delivered cognitive behaviour therapy for paediatric obsessive-compulsive disorder: a machine learning approach. *Int J Methods Psychiatr Res*. 2018 Mar;27(1). doi: 10.1002/mpr.1576. Epub 2017 Jul 28.
18. Carrillo F, Sigman M, Fernández Slezak D, Ashton P, Fitzgerald L, Stroud J, et al. Natural speech algorithm applied to baseline interview data can predict which patients will respond to psilocybin for treatment-resistant depression. *J Affect Disord*. 2018;230:84-6.
19. Pestian JP, Grupp-Phelan J, Bretonnel Cohen K, Meyers G, Richey LA, Matykiewicz P, et al. A controlled trial using natural language processing to examine the language of suicidal adolescents in the emergency department. *Suicide Life Threat Behav*. 2016;46:154-9.
20. Dazzan P. Neuroimaging biomarkers to predict treatment response in schizophrenia: the end of 30 years of solitude? *Dialogues Clin Neurosci*. 2014;16:491-503.
21. Zhang W, Yang X, Lui S, Meng Y, Yao L, Xiao Y, et al. Diagnostic prediction for social anxiety disorder via multivariate pattern analysis of the regional homogeneity. *Biomed Res Int*. 2015;2015:763965.



22. Schnack HG, Nieuwenhuis M, van Haren NE, Abramovic L, Scheewe TW, Brouwer RM, et al. Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects. *Neuroimage*. 2014; 84:299-306.
23. Redlich R, Opel N, Grotegerd D, Dohm K, Zaremba D, Bürger C, et al. Prediction of individual response to electroconvulsive therapy via machine learning on structural magnetic resonance imaging data. *JAMA Psychiatry*. 2016;73:557-64.
24. Yun SS, Choi J, Park SK, Bong GY, Yoo H. Social skills training for children with autism spectrum disorder using a robotic behavioral intervention system. *Autism Res*. 2017;10:1306-23.
25. Voss C, Schwartz J, Daniels J, Kline A, Haber N, Washington P, et al. Effect of wearable digital intervention for improving socialization in children with autism spectrum disorder: a randomized clinical trial. *JAMA Pediatr*. 2019;173:446-54.
26. Mengoni SE, Irvine K, Thakur D, Barton G, Dautenhahn K, Guldberg K, et al. Feasibility study of a randomized controlled trial to investigate the effectiveness of using a humanoid robot to improve the social skills of children with autism spectrum disorder (Kaspar RCT): a study protocol. *BMJ Open*. 2017;7:e017376.
27. Wesemann U, Kowalski JT, Jacobsen T, Beudt S, Jacobs H, Fehr J, et al. Evaluation of a technology-based adaptive learning and prevention program for stress response -- a randomized controlled trial. *Mil Med*. 2016;181:863-71.