Journal of Health Informatics

J Health Inform. 2026;18:e1849
DOI: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1849

ARTIGO COMPLETO
CBIS 2026

Da pesquisa à aplicação clínica: uma revisão sistemática sobre aprendizagem de máquina no diagnóstico do transtorno do déficit de atenção e hiperatividade

From research to clinical application: a systematic review of machine learning in the diagnosis of attention-deficit/hyperactivity disorder

Mayra Daher de Carvalho Pereira[1], Viviane Borges[2], Clauirton de Albuquerque Siebra[1]

ORCID: 0009-0003-2991-0593 · 0000-0003-0753-9423 · 0000-0002-9742-2175

[1] Centro de Informática, Universidade Federal da Paraíba, João Pessoa (PE), Brasil.

[2] Centro de Ciências da Saúde, Universidade Federal da Paraíba, João Pessoa (PE), Brasil.

FICHA DO ARTIGO

TIPO
Artigo completo · CBIS 2026

IDENTIFICADOR
e1849

DOI
10.59681/2175-4411.v18.iEspecial.2026.1849

RECEBIDO
17/05/2026

APROVADO
20/07/2026

PROCESSO DE AVALIAÇÃO
Double Blind Review

EDITOR RESPONSÁVEL
Juliano de Souza Gaspar

ACESSO
Aberto, CC BY 4.0

AUTOR CORRESPONDENTE
Mayra Daher de Carvalho Pereira
mayra.daher@academico.ufpb.br

LICENÇA

Este é um artigo publicado em acesso aberto sob a licença Creative Commons CC BY 4.0, que permite uso, distribuição e reprodução em qualquer meio, desde que o trabalho original seja corretamente citado.

CC BY 4.0

RESUMO

Objetivo: Este trabalho visa identificar lacunas e limitações dos modelos de Aprendizagem de Máquina (AM) desenvolvidos para o diagnóstico do Transtorno do Déficit de Atenção e Hiperatividade (TDAH). **Métodos:** Trata-se de uma Revisão Sistemática da Literatura conduzida a partir das orientações do PRISMA. As buscas foram realizadas nas bases de dados ACM Digital Library, IEEE Xplore Digital Library e PubMed e procuraram estudos que aplicassem AM no diagnóstico de TDAH com explicabilidade. **Resultados:** Nove estudos foram incluídos no corpus final. CNN e XGBoost são os algoritmos mais utilizados. Acurácia e sensibilidade são as métricas mais reportadas. Dentre as estratégias de validação, a 5-fold cross-validation foi a mais utilizada. Quanto à explicabilidade, a técnica SHAP foi a mais recorrente. **Conclusão:** Apesar dos avanços, os modelos propostos ainda não conseguem superar as barreiras para implementação clínica, pois se limitam ao desempenho técnico. São necessárias novas soluções para o desenvolvimento de soluções aplicáveis na clínica.

Descritores: Aprendizado de máquina; Explicabilidade; Prática clínica.

ABSTRACT

Objective: This work aims to identify gaps and limitations of Machine Learning models developed for Attention Deficit Hyperactivity Disorder diagnosis. **Methods:** This is a Systematic Literature Review conducted according to the PRISMA guidelines. The searches were performed in the ACM Digital Library, IEEE Xplore Digital Library, and PubMed databases, targeting studies that applied Machine Learning to ADHD diagnosis with explainability. **Results:** Nine studies were included in the final corpus. CNN and XGBoost are the most commonly used algorithms. Accuracy and sensitivity are the most frequently reported metrics. Among the validation strategies, 5-fold cross-validation was the most utilized. Regarding explainability, the SHAP technique was the most recurrent. **Conclusion:** Despite progress, the proposed models fail to overcome the barriers to clinical implementation, remaining limited to technical performance. New approaches are required to develop tools that are truly applicable in clinical practice.

Keywords: Machine Learning; Explainability; Clinical practice.

CBIS 2026 XXI CONGRESSO BRASILEIRO DE INFORMÁTICA EM SAÚDE

Artigo apresentado no XXI Congresso Brasileiro de Informática em Saúde (CBIS 2026), realizado de 23 a 25 de setembro de 2026, em Brasília (DF), e publicado na seção CBIS 2026 do Journal of Health Informatics.

COMO CITAR ESTE ARTIGO: Pereira MDC, Borges V, Siebra CA. Da pesquisa à aplicação clínica: uma revisão sistemática sobre aprendizagem de máquina no diagnóstico do transtorno do déficit de atenção e hiperatividade. J Health Inform 2026;18:e1849. doi: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1849

1 INTRODUÇÃO

De acordo com o Manual Diagnóstico e Estatístico de Transtornos Mentais [1], o Transtorno do Déficit de Atenção e Hiperatividade (TDAH) é uma neurodivergência que apresenta os primeiros sintomas ainda na infância, segue pela vida adulta e, se não for adequadamente tratado, pode acarretar prejuízos expressivos na vida social, acadêmica e profissional do indivíduo, como mau desempenho no trabalho, dificuldades na gestão das finanças pessoais e propensão ao uso de substâncias viciantes. Segundo Fernández-Delgado *et al.* [2], a incidência de sintomas de transtornos do desenvolvimento cresceu nos últimos anos, com destaque para o TDAH. O TDAH pode se apresentar em três subtipos: predominantemente desatento, predominantemente hiperativo/impulsivo e combinado [3]. Além disso, outras comorbidades podem se apresentar associadas, como o Transtorno do Espectro Autista [4], Transtorno de ansiedade generalizada e Transtorno de oposição desafiante [3].

De acordo com [5], existem diferenças em partes estruturais e funcionais nos cérebros de indivíduos neurodivergentes quando comparados aos neurotípicos. Essas variações estão associadas a mudanças expressivas nas funções executivas e cognitivas, assim como na regulação dopaminérgica e atencional. No entanto, evidências científicas mostram que essas alterações não são estáticas, o que reforça a importância da detecção precoce e do tratamento adequado para proporcionar melhorias na qualidade de vida. A estimativa mundial de indivíduos com TDAH é de 6,76% entre adultos e 7,2% entre crianças, o que evidencia o impacto desse transtorno na saúde pública [6].

No sistema público de saúde brasileiro (SUS), o diagnóstico e tratamento deste transtorno devem ser feitos nos Centros de Atenção Psicossocial (CAPS) ou Centros de Atenção Psicossocial Infantojuvenil (CAPSi). Somente em 2022 foram realizados 229.872 atendimentos ambulatoriais de crianças com sintomas de TDAH no SUS [6]. Porém, dificuldades de acesso e horários reduzidos dos atendimentos constituem barreiras relevantes ao diagnóstico e à continuidade do tratamento [7].

No Brasil, o diagnóstico de TDAH é feito principalmente através da aplicação de testes, questionários e entrevistas que podem ser conduzidos por psicólogos, mas precisam de validação médica. Contudo, por serem aplicados e interpretados manualmente, esses instrumentos permanecem sujeitos à subjetividade e à variação no nível de experiência do profissional [8]. Diante desse cenário, estudos [4] reforçam a importância de projetar soluções automatizadas voltadas ao auxílio do diagnóstico e da condução terapêutica.

O crescente volume de dados na área da saúde tem favorecido a utilização de técnicas de aprendizagem de máquina (AM) para diagnóstico de diversas doenças e transtornos, incluindo o TDAH. No entanto, a adesão clínica exige que as ferramentas de AM sejam capazes de produzir resultados robustos e reproduzíveis. Além disso, aspectos como a explicabilidade das predições e a capacidade de integração delas às práticas diagnósticas já existentes estão diretamente relacionados à confiança dos profissionais nos algoritmos e em seus resultados [9]. Nesse contexto, o cenário atual das aplicações de AM lida com um *trade-off*: os modelos que apresentam melhor desempenho técnico tendem a ser menos transparentes. Por outro lado, aqueles mais interpretáveis geralmente oferecem menor eficácia nas métricas de avaliação [9].

Apesar do crescimento nas pesquisas e do aumento na criação de soluções baseadas em inteligência artificial para auxiliar na triagem e diagnóstico de TDAH, até onde sabemos, não foi encontrado na literatura nenhum artigo que tenha realizado uma revisão sistemática da literatura (RSL) abrangendo AM, explicabilidade e implementação real no cotidiano clínico. Logo, o presente trabalho justifica-se pela necessidade de mapear e analisar as abordagens diagnósticas mais recentes para o TDAH que considerem a explicabilidade e a integração a fluxos de trabalho clínicos. O objetivo é identificar as limitações e lacunas das pesquisas atuais e, assim, contribuir com o desenvolvimento de ferramentas que superem as barreiras na transição clínica. Para tanto, conduziu-se uma RSL sobre o tema, ressaltando aspectos como atributos técnicos intrínsecos aos modelos que enriquecem a confiança, a explicabilidade e os padrões da prática médica atual.

2 MÉTODOS

Esta RSL foi desenvolvida com o objetivo de investigar o estado atual da validação e implantação clínica dos modelos de AM aplicados ao diagnóstico do TDAH, e seguiu as seguintes etapas: estruturação do protocolo de pesquisa; busca, seleção dos estudos e extração dos dados; e análise qualitativa. A RSL foi delineada a partir das diretrizes do PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*), e para auxiliar na documentação e organização dos resultados utilizou-se a ferramenta *Rayyan*. Com o objetivo de eliminar vieses, a etapa de busca, seleção e extração dos dados foi conduzida por duas pesquisadoras e de forma independente.

A estruturação do protocolo de pesquisa iniciou-se com a definição da pergunta de pesquisa que rege este estudo: “Como avaliar e validar algoritmos de aprendizagem de máquina aplicados ao diagnóstico de TDAH, para apoiar profissionais de saúde na obtenção de diagnósticos mais precisos e consistentes, no contexto de práticas clínicas?” Além disso, perguntas secundárias auxiliares também foram utilizadas, são elas: (RQ1) Quais os modelos de AM e aprendizagem

profunda mais propostos para o diagnóstico de TDAH? (RQ2) Quais os atributos técnicos intrínsecos aos modelos que contribuem para a confiança clínica? (RQ3) De qual forma a explicabilidade (XAI) é integrada aos modelos? (RQ4) Como os modelos propostos são alinhados aos padrões da prática clínica atual?

Na definição da estratégia de pesquisa, foi utilizado o critério PICo, e estabeleceu-se a seguinte *string* de busca: ("attention deficit hyperactivity disorder" OR "ADHD") AND ("machine learning" OR "deep learning") AND ("explainability" OR "XAI" OR "interpretability") AND ("diagnosis" OR "clinical"). A *string* de busca foi aplicada nas seguintes bases de dados: ACM Digital Library, IEEE Xplore Digital Library e PubMed.

Para refinar a pesquisa foram empregados critérios de inclusão e critérios de exclusão. Os critérios de inclusão aplicados foram: CI01: Estudos primários revisados por pares; CI02: Estudos publicados entre 01/01/2020 e 31/03/2026; CI03: Estudos nos idiomas inglês, português ou espanhol. Já os critérios de exclusão foram: CE01: Trabalhos que não apresentaram as métricas de avaliação do modelo; CE02: Tipo da publicação (capítulos de livros, teses, dissertações, *posters*, editoriais, anais ou resumos expandidos, *survey*); CE03: Trabalhos onde o modelo de AM foi utilizado para outras aplicações que não sejam o diagnóstico/classificação de TDAH; CE04: Artigos cujo texto completo não está disponível ou que não estão escritos em inglês, português ou espanhol; CE05: Trabalhos que não apresentam aplicação de explicabilidade; CE06: Trabalhos que não apresentam nenhuma estratégia de alinhamento dos modelos ao contexto da prática clínica.

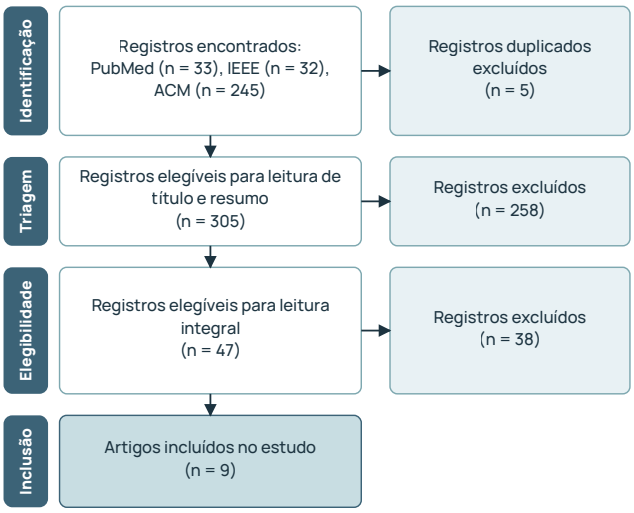
A etapa de busca e seleção dos estudos foi composta por: busca nas bases de dados previamente definidas, exclusão de duplicatas, triagem, leitura integral dos trabalhos candidatos e, por fim, inclusão dos selecionados no *corpus* final de artigos. Então, após a inclusão, a coleta e a organização de dados foram feitas com o auxílio de uma planilha de extração de dados. Foram coletados: (a) título do artigo, (b) nomes dos autores, (c) ano de publicação, (d) países dos autores, (e) revista onde foi publicado, (f) algoritmo utilizado, (g) métricas de avaliação do modelo, (h) técnica de validação do modelo, (i) se o algoritmo foi utilizado para classificação binária, (j) se o modelo foi utilizado para classificação dos subtipos de TDAH, (k) se o modelo foi utilizado para diagnóstico somente de TDAH, ou TDAH e comorbidades correlatas, (l) características dos participantes, (m) características do *dataset*, (n) técnica de explicabilidade utilizada, (o) padrões da prática clínica atual utilizados no trabalho, (p) avaliação clínica realizada por profissional da saúde, (q) utilização de biomarcadores específicos do TDAH, (r) técnica longitudinal aplicada.

Finalmente, a análise qualitativa dos dados foi conduzida realizando correlações entre os artigos, identificando tendências, padrões e lacunas na literatura selecionada. Em seguida, foi realizada uma síntese narrativa com os principais achados de maneira a responder às perguntas de pesquisa definidas e discutir o estado da arte acerca do tema.

3 RESULTADOS

A busca foi iniciada no dia 10 de abril de 2026 nas 3 bases de dados supracitadas. Foram identificados 310 artigos, dos quais 5 duplicatas foram excluídas. Na seleção, a triagem inicial foi realizada aplicando os critérios de inclusão e exclusão nos títulos e resumos, resultando em 47 artigos candidatos. Posteriormente, a leitura integral dos trabalhos foi realizada. Seguindo os devidos critérios de inclusão e exclusão, 38 foram excluídos, sendo o CE06 o principal motivo, com dezesseis trabalhos, seguido por CE03 com dez, CE04 e CE05 com cinco cada e CE01 e CE02 com uma pesquisa cada. Por fim, nove trabalhos foram incluídos no *corpus* final de artigos.

Figura 1. Fluxograma da etapa de busca e seleção de artigos



Fonte: Elaborado pelos autores (2026).

A coleta e organização de dados foram feitas com o auxílio da planilha de extração de dados, e os artigos selecionados podem ser brevemente visualizados no Quadro 1 abaixo.

A capacidade de generalização dos modelos também foi analisada nos artigos selecionados. A técnica de validação frequente foi: *5-fold cross-validation* (n=4), nos trabalhos [10], [8], [12] e [16]; outras abordagens identificadas incluíram a validação *10-fold cross-validation* no trabalho [11], a estratégia *Leave-One-Subject-Out* (LOSO) no estudo [15] e a validação *5-fold stratified cross-validation* no artigo [17]. Ademais, vale destacar que duas pesquisas não utilizaram nenhuma técnica de validação cruzada.

Quadro 1. Síntese dos estudos incluídos na revisão.

Artigo	Base de dados/ modalidade	Amostra e Perfil demográfico	Algoritmo principal	Métricas de avaliação	explicabilidade	Aspecto facilitador da translação clínica
(10)	Base de dados: Healthy Brain Network (HBN - Child Mind Institute) Modalidades: EEG (105 canais) em tarefa de aprendizado de sequência	Amostra: 121 participantes (62 com TDAH e 59 Controles/ Não-TDAH) Perfil: Faixa etária de 16 a 21 anos	CNN - EEGNet architecture	Acurácia Precisão Sensibilidade Especificidade F1-score	Análise de ablação, Mapa de saliências, Análise de filtros espaço-temporais	Utilização do biomarcador event-related potentials (ERPs) e equipe multidisciplinar
(11)	Base de dados: SASSI Trial / Child Guidance Clinic (Cingapura) Modalidade: Eletrocardiograma de 1 canal (1D-ECG bruto, segmentos de 2s)	Amostra: 123 participantes (45 TDAH, 62 TDAH+TC [Transtorno de Conduta] e 16 TC) Perfil: Crianças e adolescentes (7 a 16 anos)	convolutional neural network (CNN) - 1D	Acurácia Precisão Sensibilidade F1-score	Grad-CAM	equipe multidisciplinar Uso de sinal de ECG
(8)	Base de dados: testes psicológicos de inteligência em prontuários eletrônicos do *Catholic Medical Center* (Coreia do Sul, 2011–2022) Modalidade: Texto livre / Linguagem Natural Não Estruturada	Amostra: 1.475 laudos/ pacientes (461 Deficiência Intelectual [DI] e 1.014 TDAH) Perfil: Idade de 0 a 90 anos, 67,2% do sexo masculino; diagnósticos médicos de DI e TDAH	Natural Language Processing, Naive Bayes, Random Forest, eXtreme Gradient Boosting, Light Gradient Boosting Machine	Acurácia Sensibilidade F1-score	NOTS (*n-gram feature-based original text search*) Importância por permutação e SHAP	Utilização de testes psicológicos e equipe multidisciplinar
(12)	Base de dados: Plataforma de telemedicina MedicalBIT (IRCCS Eugenio Medea, Itália, 2018–2022) Modalidade: Questionário socioanamnésico/ fenotípico	Amostra: 1.688 crianças e adolescentes (270 TDAH, 270 TEA, e demais transtornos/ sem diagnóstico) Perfil: 2 a 16 anos; 35% do sexo feminino	Random forest, Decision trees, Logistic Regression	Acurácia Sensibilidade	SHAP	Utilização de questionário socioanamnésico / fenotípico e equipe multidisciplinar
(13)	Base de dados: ADHD-200 Modalidades: Trimodal (fMRI, sMRI e dados fenotípicos/ texto)	grupos TDAH vs. Controle/ Não-TDAH Perfil: Idade, gênero, manualidade e QI médio	3D ResNet-50 + BioBERT + DenseNet	Acurácia Precisão Sensibilidade, F1-score	Estudo de ablação	Utilização de dados fenotípicos, QI, biomarcadores de neuroimagem e equipe multidisciplinar
(14)	Base de dados: ADHD-200 Modalidade: EEG (19–32 canais no sistema 10-20)	Amostra: 375 crianças (185 TDAH e 190 Controles) Perfil: Faixas etárias entre 7 e 13 anos; multicentros	dual-transformer (com Vision Transformer e Time-Series Transformer)	Acurácia Precisão Sensibilidade F1-score	Grad-CAM, SHAP e estudo de ablação	Utilização de biomarcadores eletrofisiológicos e avaliação dos resultados por neurologistas pediátricos
(15)	Base de dados: Dados coletados em três hospitais (Daegu Catholic University Medical Center, Seoul National University Hospital e Hanyang University Seoul Hospital) Modalidade: resting-state EEG	Amostra: 168 participantes (107 no grupo TDAH e 61 no grupo Neurotípico/ Controle). Perfil: Crianças e adolescentes de 8 a 15 anos.	Extreme Gradient Boosting	Acurácia Precisão Sensibilidade Especificidade F1-score	SHAP	Utilização de biomarcadores e equipe multidisciplinar
(16)	Base de dados: Dados clínicos do *South West Yorkshire Partnership NHS Foundation Trust*, Reino Unido (2019 a 2024). Modalidade: questionários socioanamnésico e psicométricos	Amostra: 786 participantes adultos (302 com TDAH / 39% e 484 Controles / 61% sem TDAH) Perfil: Adultos (18 a 40 anos e maior proporção de homens)	Extreme Gradient Boosting	Acurácia AUC - ROC	SHAP	questionário socioanamnésico, questionários psicométricos e equipe multidisciplinar
(17)	Base de dados: Peking University Sixth Hospital / AHBA (Allen Human Brain Atlas) / 1000 Genomes Project Modalidade: Multimodal (fMRI em repouso + dados genômicos)	Amostra: 484 participantes (247 adultos com TDAH e 237 Controles/Não-TDAH) Perfil demográfico: Adultos	Edge-Variational Graph Convolution Network	Acurácia AUC - ROC	Análise de Sensibilidade (SA), SHAP e Gradientes Integrados	Utilização de dados genéticos (polimorfismos de nucleotídeo único) e biomarcadores

Fonte: Elaborado pelos autores (2026).

Os países com maior volume de publicações foram China, Coreia do Sul, Índia, Austrália e Itália, com 2 cada, o que indica uma crescente relevância do continente asiático na área. Os estudos selecionados são de 2025 (n=5), o que sugere um aumento do interesse pelo assunto e da maturidade das pesquisas. No que diz respeito ao perfil dos participantes, a maioria (6 trabalhos) estudaram crianças e adolescentes. No que concerne à aplicação, cerca de 67% dos modelos foram utilizados para classificação binária, 33% para diagnóstico de TDAH e outras comorbidades associadas, e nenhum deles teve como objetivo distinguir os subtipos de TDAH, o que aponta uma possível lacuna na literatura.

3.1 (RQ1) Quais os modelos de AM e aprendizagem profunda mais propostos para o diagnóstico de TDAH?

A análise dos resultados revelou que tanto Redes Neurais Convolucionais (CNNs) (n=3) quanto o eXtreme Gradient Boosting (XGBoost) (n=3) aparecem como os modelos mais empregados. Observa-se que as CNNs foram utilizadas em conjunto com biomarcadores, o que valida sua robustez para processar dados como imagens e sinais. Por outro lado, o XGBoost demonstrou versatilidade sendo aplicado tanto em dados clínicos/comportamentais quanto em biomarcadores neurológicos. Essa flexibilidade do XGBoost contribui para torná-lo um dos algoritmos mais amplamente adotados para classificação na literatura selecionada.

3.2 (RQ2) Quais os atributos técnicos intrínsecos aos modelos que contribuem para a confiança clínica?

Sob a ótica da confiança clínica, os atributos mais reportados foram as métricas de desempenho, a capacidade de generalização e a capacidade de interpretação do modelo ou seu resultado. Dentre os indicadores de desempenho, os mais recorrentes são acurácia (n=9) e sensibilidade (n=7), que validam a eficácia dos algoritmos. Já a capacidade de generalização foi avaliada através de estratégias de validação cruzada, dentre as quais a *5-fold cross-validation* foi a mais utilizada. Ademais, observa-se que tanto os trabalhos que empregam modelos tradicionais quanto aqueles que utilizam arquiteturas complexas ("caixa-preta") manifestam preocupação crescente com a capacidade de interpretação, oferecendo métodos *post hoc* para garantir transparência na tomada de decisão.

3.3 (RQ3) De qual forma a explicabilidade (XAI) é integrada aos modelos?

Ao investigar a literatura selecionada, constatou-se que mesmo em algoritmos mais tradicionais e transparentes, como Regressão Logística e Árvore de Decisão, a explicabilidade foi utilizada para enriquecer a compreensão humana sobre o comportamento ou predições dos modelos. A técnica SHAP (*SHapley Additive exPlanations*) foi a mais recorrente, aparecendo em seis estudos. Em seguida, vem a Análise de Ablação (n=3), adotada por dois modelos de CNN. Observou-se também que na maioria dos modelos de arquitetura híbrida é utilizada mais de uma técnica de explicabilidade, o que sugere uma tendência nas estratégias de validação de modelos mais complexos e mitigação de vieses.

3.4 (RQ4) Como os modelos propostos são alinhados aos padrões da prática clínica atual?

A análise das pesquisas revelou os aspectos mais empregados que facilitam a translação clínica: equipes de caráter multidisciplinar e o emprego de biomarcadores. Predominantemente, encontramos times compostos por pelo menos um integrante da área de saúde (n=8). Adicionalmente, a utilização de biomarcadores de neuroimagem ou eletrofisiológicos foi identificada em seis estudos. Dados oriundos de testes psicológicos, questionários socioanamnésicos e questionários psicométricos também foram utilizados, porém com menor incidência. Contudo, é importante ressaltar que apenas um trabalho teve seus resultados avaliados por uma equipe externa de profissionais da área de saúde.

4 DISCUSSÃO

Os resultados sugerem que além do desempenho das métricas de avaliação e da explicabilidade, a capacidade de generalização dos resultados é um atributo relevante na construção da confiança clínica. Outrossim, indicam também que a adoção de equipes multidisciplinares e a utilização de marcadores biológicos constituem os recursos mais empregados que facilitam a translação para o contexto clínico.

Evidenciou-se que a escolha da arquitetura depende do tipo dos dados que serão utilizados, e os biomarcadores aparecem em mais da metade dos trabalhos, apontando uma tendência: priorização de dados objetivos em detrimento de dados subjetivos oriundos de questionários e testes. Da mesma forma, constatou-se a adoção da combinação de mais de uma técnica de explicabilidade em arquiteturas híbridas para incrementar a explicabilidade e facilitar a adesão por profissionais da saúde.

Apesar dos avanços, algumas lacunas expressivas foram detectadas a partir da análise dos estudos. Apenas uma equipe submeteu seu trabalho à avaliação por profissionais externos da área de saúde, o que demonstra que, embora os modelos apresentem alto desempenho técnico (alta acurácia, por exemplo), eles ainda carecem de avaliação real. Além disso, identificou-se que, a despeito do uso de marcadores biológicos ou da adoção de equipe multidisciplinar, as soluções ainda são insuficientes para superar os gargalos da implementação clínica. Adicionalmente, constatou-se que nenhum estudo propôs modelos voltados à classificação dos subtipos de TDAH (predominantemente desatento, predominantemente

hiperativo/impulsivo e combinado), o que constitui uma barreira crítica para o uso médico.

Este trabalho apresenta algumas limitações pertinentes. A busca restringiu-se a apenas três bases de dados, ACM Digital Library, IEEE Xplore Digital Library e PubMed, o que pode influenciar o perfil dos dados obtidos. Por conseguinte, a pequena amostra de trabalhos selecionados para a RSL (n=9) revela a necessidade de mais estudos que investiguem não apenas o desempenho técnico dos modelos, mas também suas qualidades para facilitar a adesão clínica.

Dessa forma, trabalhos futuros podem focar na distinção dos subtipos de TDAH, na detecção de comorbidades correlatas, no desenvolvimento de ferramentas que entregam saídas mais interpretáveis, e na validação clínica dos modelos por meio de ensaios prospectivos ou avaliação multidisciplinar independente. Sugerem-se, ainda, estudos que promovam a avaliação da qualidade metodológica dos trabalhos incluídos nesta RSL.

5 CONCLUSÕES

A inteligência artificial tem propiciado a criação de ferramentas de suporte ao diagnóstico clínico do TDAH. Este trabalho apresentou uma RSL guiada pelo protocolo PRISMA, cujo corpus final é composto por 9 estudos. Os resultados confirmam que a confiança e adesão clínica dos modelos de aprendizagem de máquina propostos vão além de atributos técnicos e metodológicos tradicionais, como acurácia, capacidade de generalização, explicabilidade, emprego de biomarcadores ou adoção de equipe multidisciplinar. São necessárias abordagens e soluções adicionais para transpor os desafios da translação clínica e desenvolver ferramentas efetivamente aplicáveis no cotidiano médico.

Contribuição dos autores. A concepção e delineamento do estudo foram conduzidos por Mayra Pereira e Clauirton Siebra. A coleta, análise e interpretação dos dados foram realizadas por Mayra Pereira e Viviane Borges. A redação do manuscrito foi realizada por Mayra Pereira. A revisão crítica foi realizada por Mayra Pereira, Viviane Borges e Clauirton Siebra. A supervisão do projeto foi exercida por Clauirton Siebra. Todos os autores participaram da revisão e aprovaram a versão final do manuscrito.

Agradecimentos. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES), código de Financiamento 001.

Financiamento. Os autores não declararam fonte de financiamento.

Disponibilidade de dados e materiais. Estão contidos neste artigo publicado a estratégia de busca completa para todas as bases de dados consultadas, o fluxograma PRISMA detalhado e as colunas da matriz de extração de dados. Os registros brutos das etapas de triagem e elegibilidade podem ser disponibilizados pelo autor correspondente mediante solicitação razoável.

Conflitos de interesse. Os autores declaram a ausência de conflitos de interesse de ordem financeira, pessoal ou profissional na condução deste estudo.

Declaração de uso de IA. No desenvolvimento deste trabalho foi utilizada a ferramenta de IA Gemini. A ferramenta foi utilizada para: apoio à redação e revisão de texto (clareza, gramática e estilo); organização e estruturação do manuscrito; auxílio na síntese dos artigos estudados na RSL; tradução do inglês para o português e do português para o inglês.

REFERÊNCIAS

1. American Psychiatric Association. Manual diagnóstico e estatístico de transtornos mentais: DSM-5. 5ª ed. Porto Alegre: Artmed; 2014.
2. Fernández-Delgado M, Cruz S, Cernadas E, Alateyat H, Tubío-Fungueiriño M, Sampaio A, et al. Population-based detection of children ASD/ADHD comorbidity from atypical sensory processing. Appl Intell. 2024;54(20):9906-23. doi: https://doi.org/10.1007/s10489-024-05655-z
3. Barkley RA, Benton CM. Vencendo o TDAH adulto: transtorno de déficit de atenção/hiperatividade. 2. ed. Porto Alegre: Artmed; 2022.
4. Lima Teles da Hora AF, Rodrigues Louzã Neto M. ASD and ADHD in children: a preliminary study. Psico. 2024:e41923. doi: https://doi.org/10.15448/1980-8623.2023.2.41923
5. Ul Ain Q, Jawed S, Rauf Subhani A, Haider Butt W, Usman Akram M. Examining AI-Powered ADHD Diagnosis: Current Trends, Key Challenges, and Future Directions in the Field. IEEE Access. 2025;13:93148-77. doi: https://doi.org/10.1109/access.2025.3567427
6. Ribeiro ADB, Silva GTR, Rodrigues GHC, Belatto MFB, Carvalho LBA, Pelisão T, et al. Mapeando o TDAH no Brasil: prevalência e desigualdades por região, faixa etária e raça. Rev. Contemp. 2024;4(7):e5267. doi: https://doi.org/10.56083/rcv4n7-210
7. Tomazelli J, Girianelli VR, Fernandes CS. Incidência de transtorno global do desenvolvimento em crianças: características e análise a partir dos CAPSi. Psicol. USP. 2023;34:e210002. doi: https://doi.org/10.1590/0103-6564e210002
8. Kim TM, Kim YH, Song SH, Choi IY, Kim DJ, Ko T. Explainability Enhanced Machine Learning Model for Classifying Intellectual Disability and Attention-Deficit/Hyperactivity Disorder With Psychological Test Reports. J Korean Med Sci. 2025;40(11):e26. doi: https://doi.org/10.3346/jkms.2025.40.e26
9. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. Wiley Interdiscip Rev Data Min Knowl Discov. 2019;9(4):e1312. doi: https://doi.org/10.1002/widm.1312

10. Stock S, Hausberg J, Armengol-Urpi A, Kaufmann T, Schinle M, Gerdes M, et al. Towards EEG-based objective ADHD diagnosis support using convolutional neural networks. In: 2023 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB). [s.l.]: IEEE; 2023. p. 1-8. doi: https://doi.org/10.1109/cibcb56990.2023.10264876

11. Loh HW, Ooi CP, Oh SL, Barua PD, Tan YR, Molinari F, et al. Deep neural network technique for automated detection of ADHD and CD using ECG signal. Comput Methods Programs Biomed. 2023;241:107775. doi: https://doi.org/10.1016/j.cmpb.2023.107775

12. Grazioli S, Crippa A, Buo N, Busti Ceccarelli S, Molteni M, Nobile M, et al. Use of Machine Learning Models to Differentiate Neurodevelopment Conditions Through Digitally Collected Data: Cross-Sectional Questionnaire Study. JMIR Form Res. 2024;8:e54577. doi: https://doi.org/10.2196/54577

13. Jia J, Liang R, Ding Y, Zhang C, Li Y, Wang C, et al. MMAFN: Multi-Modal Attention Fusion Network for Attention-Deficit Hyperactivity Disorder Classification. In: 2025 IEEE 2nd International Conference on Deep Learning and Computer Vision (DLCV). [s.l.]: IEEE; 2025. p. 1-5. doi: https://doi.org/10.1109/dlcv65218.2025.11088700

14. Ramamoorthy B, Sivasankaran HB, P J. NeuroScope:A Transformer-Based Explainable EEG Diagnostic Framework for ADHD Detection in Children. In: 2025 International Conference on Intelligent Computing, Information and Control Systems (ICOIICS). [s.l.]: IEEE; 2025. p. 874-9. doi: https://doi.org/10.1109/icoiics67115.2025.11390130

15. Kim JW, Kim BN, Kim JI, Yang CM, Kwon J. Electroencephalogram (EEG) Based Prediction of Attention Deficit Hyperactivity Disorder (ADHD) Using Machine Learning. Neuropsychiatr Dis Treat. 2025;21:271-9. doi: https://doi.org/10.2147/NDT.S509094

16. Shakeel HM, Antoniou G, Adamou M. Integrating explainable AI with clinical features to enhance ADHD diagnostic understanding. Front Psychiatry. 2025;16:1706216. doi: https://doi.org/10.3389/fpsyt.2025.1706216

17. Zhu Z, Liu Y, Liu Q, Yue X, Pan M, Gao Y, et al. Multimodal-based deep learning detected disrupted precuneus connectivity and its related genetic profiles for predicting adults with ADHD. J Affect Disord. 2026;396:120821. doi: https://doi.org/10.1016/j.jad.2025.120821

NOTAS EDITORIAIS

O Journal of Health Informatics é uma publicação científica mantida pela Sociedade Brasileira de Informática em Saúde, associação científica sem fins lucrativos dedicada ao desenvolvimento da Informática em Saúde e da Saúde Digital no Brasil. O processo editorial é conduzido com independência científica, sem interferência institucional, comercial ou de patrocinadores, e segue as boas práticas internacionais de ética em publicação científica.

Este artigo foi avaliado por pares, em sistema duplo-cego, por especialistas da área. A equipe editorial usou ferramentas de inteligência artificial, com supervisão humana, na triagem inicial, na conferência do texto, das referências e dos números e, quando necessário, no redesenho de figuras, quadros e tabelas, aprovado pelos autores na prova. As ferramentas e as salvaguardas estão na Política de uso de inteligência artificial da revista. As decisões editoriais foram tomadas por pessoas.

SBIS Sociedade Brasileira de Informática em Saúde