



B1

ISSN: 2595-1661

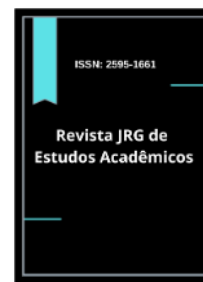
ARTIGO ORIGINAL

Listas de conteúdos disponíveis em [Portal de Periódicos CAPES](https://portaldeperiodicos.capes.gov.br/)

Revista JRG de Estudos Acadêmicos

Página da revista:

<https://revistajrg.com/index.php/jrg>



A eficácia do *ChatGPT* no ensino de medicina por meio de mapeamento sistemático

The effectiveness of *ChatGPT* in teaching medicine through systematic mapping

DOI: 10.55892/jrg.v8i18.2005

ARK: 57118/JRG.v8i18.2005

Recebido: 19/02/2025 | Aceito: 05/04/2025 | Publicado on-line: 07/04/2025

Clodoaldo Lopes da Silva¹

<https://orcid.org/0009-0000-2471-9966>

<https://lattes.cnpq.br/4239471010684142>

Universidade Federal de Alagoas, UFAL, Brasil

E-mail: cloppes.s@gmail.com

Andrea Marques Vanderlei Fregadolli²

<https://orcid.org/0009-0005-6267-9172>

<http://lattes.cnpq.br/5455567894430418>

Universidade Federal de Alagoas, UFAL, Brasil

E-mail: andreaufregadolli@gmail.com

Resumo

Introdução: Com a crescente integração de tecnologias de inteligência artificial (IA) em diversos domínios educacionais, avaliar sua eficácia torna-se primordial. Este estudo investiga a eficácia do *ChatGPT*, um modelo de IA conversacional de última geração, no âmbito do ensino de medicina. **Objetivo:** Analisar a produção científica do ensino de medicina assistido por *ChatGPT*. **Metodologia:** Mapeamento sistemático e descritivo, de natureza qualitativa. Pergunta de pesquisa criada de acordo com a estratégia PICO (*Discentes*, Intervenção, Comparação, Desfecho) Em discentes de Medicina (P), como a integração do ensino assistido por *ChatGPT* (I), como complemento aos métodos convencionais de ensino (C), influencia positivamente no aprimoramento do desempenho acadêmico e na satisfação com o processo de aprendizado (O)? Triagem dos estudos através do instrumento Prisma

¹ Docente voluntário da Faculdade de Medicina da Universidade Federal de Alagoas (FAMED). Psicólogo (CESMAC, 2010), com bacharelado, licenciatura e formação de psicólogo; Turismólogo (CESMAC, 1997). Especialista em Docência no Ensino Superior (CESMAC, 2011); MBA em Gestão de Pessoas (UNIT, 2015); Direitos Humanos, Gênero e Diversidade (UFAL, 2016); Saúde Mental e Atenção Psicossocial (UNIRIO, 2017); e Neuropsicopedagogia (FCE, 2024). Mestre em Ensino na Saúde (FAMED/UFAL, 2024). Graduando em Análise e Desenvolvimento de Sistemas pela UNINASSAU e Letras (Português/Inglês) pela FCE. Exerceu funções como Diretor, Coordenador Pedagógico e docente em cursos de pós-graduação, com ênfase em Saúde Mental, Psicologia, Gerontologia, Educação Física e ABA na UNICORP FACULDADES. Atuou como docente no IFAL, no PRONATEC e em cursos técnicos no Centro de Ensino Santa Juliana. Possui experiência em Psicologia Hospitalar, Jurídica, Escolar, Organizacional e Clínica, com atuação voltada para o público LGBTQIAPN+.

² Professora associada, nível 1, da Universidade Federal de Alagoas da Faculdade de Medicina (FAMED). Nutricionista, Acupunturista, Farmacêutica, Educadora Física, Analista e Desenvolvedora de Sistemas, Perita Grafotécnica, cibernética, judicial, extrajudicial e em Investigação Forense e Criminal. Graduada em Nutrição, Farmácia, Análise e Desenvolvimento de Sistemas, Educação Física e Educação Especial. Graduada em Biblioteconomia e Pedagogia. Especialista em Plantas Medicinais, Farmacologia, Análises Clínicas, Farmácia Clínica com Prescrição de Medicamentos, Psicopedagogia Clínica e Institucional, Nutrição Materno-infantil, Perícia Judicial e Extrajudicial, Investigação Forense e Perícia Criminal, Documentoscopia com Ênfase em Perícia Judicial, Acupuntura e Acupuntura Estética. Pós-graduanda em Ortomolecular, e Farmácia Estética. Mestre em Modelagem Computacional de Conhecimento (UFAL, 2011), Bolsa de Pesquisa no Mestrado: FAPEAL. Doutora em Ciências (UFAL, 2015).

(*Preferred Reporting Items for Systematic and Meta-Analyses*). Para a seleção dos artigos, definição dos descritores: *ChatGPT* (*ChatGPT*), *Teaching* (Ensino), *Medicine* (Medicina), *Learning* (Aprendizagem), *Artificial Intelligence* (Inteligência Artificial), “*Evaluation Studies as Topic*” (Estudos de Avaliação como Assunto), “*Educational Measurement*” (Avaliação Educacional) e “*Education, Medical, Undergraduate*” (Educação de Graduação em Medicina). Varreduras a partir das *strings* de busca, uma longa: *ChatGPT AND Medicine AND (Teaching OR Learning OR Assessment OR Evaluation Studies as Topic OR Educational Measurement OR Education Medical Undergraduate)* e uma outra curta: *ChatGPT AND medicine AND (teaching OR learning)*, nas bibliotecas virtuais (*Pubmed*, Periódicos da CAPES, *Wiley (online library)*, *BVS*, *Scielo* e *Science direct*) a partir do cruzamento de descritores em português e inglês, por meio de operador booleano (*AND*), constituindo as seguintes *strings* de busca: português – *ChatGPT AND medicina AND (ensino OR aprendizagem OR avaliação)*; inglês - *ChatGPT AND medicine AND (teaching OR learning OR assessment)*. As categorias foram criadas por meio da análise de conteúdo de Bardin no *ChatGPT*. **Resultados e Discussão:** Foram selecionados 50 artigos. E a partir da análise de conteúdo de Bardin no *ChatGPT*, foram geradas as seis categorias temáticas e uma subcategoria: Aplicações do *ChatGPT* no Ensino Médico: O *ChatGPT* é uma ferramenta complementar que pode aprimorar a aprendizagem e o desenvolvimento de habilidades críticas, proporcionando experiências personalizadas e interativas. Sua integração no ensino médico exige supervisão para garantir sua eficácia; Considerações éticas na utilização do *ChatGPT* no Ensino Médico: O uso do *ChatGPT* requer supervisão humana e avaliação crítica, sem substituir a expertise médica. A formação ética dos futuros médicos é essencial para garantir uma aplicação segura e eficaz; Subcategoria - Confiabilidade do *ChatGPT* no Ensino Médico: A confiabilidade do *ChatGPT* exige supervisão constante, já que pode gerar informações imprecisas. Sua aplicação no ensino deve ser acompanhada de orientação de educadores qualificados para garantir a precisão do conteúdo; Desafios do *ChatGPT* no Ensino Médico: O *ChatGPT* enfrenta desafios como imprecisão, inconsistência e falta de habilidades interpessoais, limitando sua eficácia como substituto do conhecimento humano. Sua aplicação deve ser monitorada de perto; Contribuições do uso do *ChatGPT* no Ensino Médico: O *ChatGPT* contribui para o ensino médico ao personalizar o aprendizado, fornecer feedback imediato e enriquecer a compreensão de conceitos complexos, sendo uma ferramenta valiosa no processo educacional; Desempenho do *ChatGPT* em Exames de Escolas de Medicina: O *ChatGPT* apresenta bom desempenho em exames médicos, com taxas de aprovação comparáveis a discentes humanos, mas ainda exige supervisão e avaliação rigorosa devido a limitações nos resultados; Direcionamentos futuros e recomendações para a integração do *ChatGPT* no Ensino Médico: Para integrar o *ChatGPT* no ensino médico de forma eficaz e ética, é necessário promover diretrizes claras, alfabetização em IA e avaliação crítica contínua, garantindo que ele complemente o julgamento clínico humano. **Conclusão:** A análise da produção científica sobre o uso do *ChatGPT* no ensino de medicina revela seu potencial como ferramenta complementar, capaz de aprimorar a aprendizagem e o desenvolvimento de habilidades críticas dos discentes. Embora o *ChatGPT* tenha mostrado resultados promissores em exames médicos e ofereça contribuições significativas, como personalização do aprendizado e feedback imediato, sua utilização exige supervisão constante e avaliação crítica. Desafios relacionados à imprecisão e falta de habilidades interpessoais ainda limitam sua eficácia como substituto do conhecimento humano, ressaltando a importância de uma

integração cuidadosa e ética dessa tecnologia. Para uma implementação eficaz, recomenda-se o desenvolvimento de diretrizes claras, além da promoção de alfabetização em IA e monitoramento contínuo. O uso do *ChatGPT* no ensino médico, quando realizado de forma ética e supervisionada, pode enriquecer a experiência educacional, complementando os métodos convencionais e contribuindo para a formação de médicos mais preparados e críticos.

Palavras-chave: *ChatGPT*; inteligência artificial; medicina; aprendizagem; ensino

Abstract

Introduction: With the increasing integration of artificial intelligence (AI) technologies into various educational domains, assessing their effectiveness becomes paramount. This study investigates the efficacy of *ChatGPT*, a state-of-the-art conversational AI model, in the field of medical education. **Objective:** To analyze the scientific production on medical education assisted by *ChatGPT*. **Methodology:** A systematic and descriptive mapping, with a qualitative nature, was employed. The research question was formulated following the PICO strategy (Population, Intervention, Comparison, Outcome): In medical students (P), how does the integration of *ChatGPT*-assisted teaching (I), as a complement to conventional teaching methods (C), positively influence academic performance improvement and learning satisfaction (O)? The screening of studies used the PRISMA tool (Preferred Reporting Items for Systematic and Meta-Analyses). For article selection, descriptors were defined: *ChatGPT*, Teaching, Medicine, Learning, Artificial Intelligence, Evaluation Studies as Topic, Educational Measurement, and Education, Medical, Undergraduate. Searches were conducted using long and short search strings: Long string: *ChatGPT* AND Medicine AND (Teaching OR Learning OR Assessment OR Evaluation Studies as Topic OR Educational Measurement OR Education Medical Undergraduate) end Short string: *ChatGPT* AND medicine AND (teaching OR learning), These searches were performed in virtual libraries (PubMed, CAPES Journals, Wiley Online Library, BVS, SciELO, and ScienceDirect) by combining descriptors in Portuguese and English, using the Boolean operator (AND). The search strings included: Portuguese: *ChatGPT* AND medicina AND (ensino OR aprendizagem OR avaliação) end English: *ChatGPT* AND medicine AND (teaching OR learning OR assessment). Categories were created using Bardin's content analysis method applied to *ChatGPT*. **Results and Discussion:** Fifty articles were selected. Using Bardin's content analysis, six thematic categories and one subcategory emerged: Applications of *ChatGPT* in Medical Education: *ChatGPT* is a complementary tool that enhances learning and critical skill development by offering personalized and interactive experiences. Its integration into medical education requires supervision to ensure efficacy; Ethical Considerations in *ChatGPT* Use in Medical Education: The use of *ChatGPT* demands human oversight and critical evaluation, without replacing medical expertise. Ethical training for future doctors is essential to ensure safe and effective application; Subcategory: Reliability of *ChatGPT* in Medical Education: *ChatGPT*'s reliability requires constant supervision, as it may generate inaccurate information. Its application in education should be guided by qualified educators to ensure content accuracy; Challenges of *ChatGPT* in Medical Education: *ChatGPT* faces challenges such as inaccuracy, inconsistency, and lack of interpersonal skills, limiting its efficacy as a substitute for human knowledge. Close monitoring is essential; Contributions of *ChatGPT* Use in Medical Education: *ChatGPT* enhances medical education by personalizing learning, providing immediate feedback, and enriching the understanding of complex concepts, making it a valuable educational

*tool; ChatGPT's Performance in Medical School Exams: ChatGPT performs well in medical exams, with approval rates comparable to human students, but its results require rigorous supervision and evaluation due to limitations; Future Directions and Recommendations for ChatGPT Integration in Medical Education: Effective and ethical integration of ChatGPT requires clear guidelines, AI literacy, and ongoing critical evaluation, ensuring it complements human clinical judgment. **Conclusion:** The analysis of scientific production on ChatGPT in medical education highlights its potential as a complementary tool capable of enhancing learning and critical skill development among students. While ChatGPT has shown promising results in medical exams and offers significant contributions, such as personalized learning and immediate feedback, its use demands constant supervision and critical evaluation. Challenges related to inaccuracy and the absence of interpersonal skills underscore the importance of careful and ethical integration. For effective implementation, the development of clear guidelines, AI literacy promotion, and continuous monitoring are recommended. When used ethically and under supervision, ChatGPT can enrich the educational experience, complement conventional methods, and contribute to the training of better-prepared and more critical medical professionals.*

Keywords: ChatGPT; artificial intelligence; medicine; learning; teaching

1. Introdução

A introdução de tecnologias de inteligência artificial em artificial (IA) no contexto educacional gerou grande interesse, especialmente na área da educação médica. O uso de ferramentas como o *ChatGPT* oferece novas possibilidades para transformar a maneira como os discentes de Medicina interagem com o conteúdo, facilitando processos de aprendizagem personalizados e oferecendo suporte no desenvolvimento de habilidades críticas e clínicas (ABD-ALRAZAQ et al., 2023). A IA, ao simular interações humanas e gerar conteúdos adaptados às necessidades individuais dos *discentes*, surge como um recurso promissor para métodos complementares tradicionais de ensino, além de fornecer feedback instantâneo e direcionado. Contudo, a implementação dessa tecnologia na educação médica requer uma análise cuidadosa de suas aplicações e limitações, garantindo que seus benefícios sejam maximizados sem comprometer os princípios fundamentais da formação profissional.

Neste contexto, o *ChatGPT* tem sido identificado como uma ferramenta multifacetada capaz de auxiliar em diversas áreas do ensino médico, como a criação de estudos de caso, a formulação de questões de múltipla escolha e até mesmo como tutor virtual, promovendo um ambiente de aprendizagem interativo e acústico (MAYTA-TOVALINO et al., 2023). No entanto, essa integração não é isenta de desafios, especialmente no que se refere à precisão das informações descobertas e ao impacto sobre o pensamento crítico dos discentes (DERGAA; CHAMARI; ZMIJEWSKI, 2023). A supervisão constante e a validação do conteúdo gerado pela IA são essenciais para garantir que a utilização do *ChatGPT* não substitua a reflexão humana e a análise crítica, elementos indispensáveis no campo da Medicina.

A primeira categoria abordada pelos autores, "Aplicações do *ChatGPT* na Educação Médica", destaca como a ferramenta se consolida como uma aliada no processo educacional, fornecendo experiências de aprendizagem personalizadas e suporte ao raciocínio clínico. As simulações clínicas, a geração de perguntas de múltipla escolha, a assistência na elaboração de estudos de caso e a atuação como um tutor virtual reforçam a versatilidade do *ChatGPT*. No entanto, a necessidade de

supervisão e validação das informações geradas é ressaltada para garantir que o uso do *ChatGPT* não prejudique o desenvolvimento do raciocínio clínico e o pensamento crítico dos discentes.

A segunda categoria, "Considerações Éticas na Utilização do *ChatGPT* no Ensino Médico", levanta preocupações sobre a confiabilidade das informações e a potencial disseminação de conteúdo impreciso. A responsabilidade humana, a supervisão constante e a manutenção da integridade acadêmica são indicadas como elementos fundamentais para garantir que a ferramenta seja usada de forma ética e responsável, evitando riscos como plágio, alucinações artificiais e informações desatualizadas.

Em "Confiabilidade do *ChatGPT* no Ensino Médico", os autores reconhecem o potencial do *ChatGPT* para fornecer insights valiosos e respostas precisas, mas também ressaltam a necessidade de mais pesquisas e validações para garantir a padronização e integridade das informações geradas. A IA não deve ser usada como uma fonte isolada de conhecimento, e o papel dos educadores em validar e complementar as respostas da ferramenta permanece crucial.

Na categoria "Desafios (fragilidades) do uso do *ChatGPT* no Ensino Médico", os autores identificam limitações relacionadas à precisão, capacidade de raciocínio crítico e questões éticas. A tendência de gerar informações incorretas ou incompletas, bem como a falta de habilidades interpessoais, destacam a necessidade de supervisão humana rigorosa e a importância de não substituir o conhecimento humano pela IA.

As "Contribuições (fortalezas) do uso do *ChatGPT* no Ensino Médico" mostram que a ferramenta possui diversas aplicações positivas, como o fornecimento de feedback imediato, personalização do aprendizado e auxílio em avaliações médicas. Sua capacidade de criar um ambiente de aprendizado interativo e adaptativo torna-o um recurso valioso para o aprimoramento das habilidades clínicas e teóricas dos discentes.

Por fim, em "Direcionamentos futuros e recomendações para a integração do *ChatGPT* na educação médica", há um consenso sobre a importância de incorporar a IA de forma responsável, ética e colaborativa. A necessidade de desenvolver diretrizes claras, alfabetização em IA e a promoção do uso supervisionado do *ChatGPT* são destacados como passos fundamentais para maximizar seu potencial.

Abd-Alrazaq et al (2023) exploraram as oportunidades e desafios associados ao uso de modelos de linguagem ampla (*LLMs*) na educação médica, com destaque para o potencial transformador dessas tecnologias no desenvolvimento curricular e na metodologia de ensino. No entanto, os autores alertaram para questões críticas, incluindo preconceito algorítmico, confiança excessiva e preocupações com privacidade e direitos autorais, destacando a necessidade de um entendimento aprofundado dos *LLMs* na educação médica.

Por sua vez, o estudo de Dergaa, Chamari e Zmijewski (2023) abordou as implicações éticas do *ChatGPT* e outras tecnologias de Processamento de Linguagem Natural (PNL) na escrita acadêmica e na pesquisa. Eles ressaltaram a importância da transparência e da preservação dos princípios éticos e acadêmicos no uso dessas ferramentas, reconhecendo o papel central do pensamento crítico humano na produção de conhecimento confiável e autêntico.

Além disso, Mayta-Tovalino et al. (2023) conduziram uma análise bibliométrica exploratória sobre o uso de agentes conversacionais inteligentes na educação médica, revelando padrões de colaboração e áreas de foco na pesquisa acadêmica.

Seus resultados destacaram a importância da interconexão entre diferentes conceitos, como "agentes conversacionais" e "Chatbots", na investigação dessas tecnologias.

Esses estudos refletem uma preocupação compartilhada em compreender e orientar o uso responsável e eficaz do *ChatGPT* na educação médica.

Diante disso, esta revisão busca consolidar e ampliar essas perspectivas, oferecendo direcionamentos futuros e recomendações para uma integração ética e produtiva do *ChatGPT* no ensino e na prática médica. Ao examinar os desafios, oportunidades e considerações éticas, emerge o questionamento: em discentes de Medicina, como a integração do ensino assistido por *ChatGPT*, como complemento aos métodos convencionais de ensino, influencia positivamente no aprimoramento do desempenho acadêmico e na satisfação com o processo de aprendizado?

Essa indagação serve como fio condutor para a presente pesquisa, que tem como objetivo avaliar a eficácia do *ChatGPT* no ensino médico por meio do mapeamento sistemático.

2. Metodologia

Diante da vasta produção científica no cenário contemporâneo, a Aprendizagem Baseada em Problemas (ABP) destaca-se como um método essencial para avaliar a relevância de descobertas e abordagens educacionais. O uso da ABP visa garantir que as decisões no campo da educação sejam tomadas com base em pesquisas rigorosas e resultados concretos, conforme apontado por Jaganathan, Bhuminathan e Ramesh (2024). Essa abordagem é cada vez mais relevante à medida que a educação se torna progressivamente complexa e interdisciplinar (RAFIQ; KAMRAN; AFZAL, 2024).

O mapeamento sistemático da literatura, conforme descrito por Petersen et al. (2021), é uma metodologia que facilita a identificação e categorização de estudos primários em áreas específicas, proporcionando uma visão abrangente sobre o estado da arte. A utilização desse tipo de mapeamento em educação permite identificar lacunas, tendências e consolidar as evidências disponíveis para embasar novas intervenções educativas (OTTO et al., 2021).

Neste estudo, o mapeamento sistemático da literatura foi escolhido como o método principal, visto que permite um levantamento robusto e detalhado de publicações relevantes, de forma a fornecer uma visão completa sobre o impacto de novas tecnologias, como o *ChatGPT*, no ensino de Medicina. Essa abordagem segue um processo estruturado de coleta, análise e síntese de dados, visando responder a perguntas específicas de pesquisa e apoiar futuras tomadas de decisão (PONTES JÚNIOR; NAKAYAMA, 2022). Assim, o mapeamento sistemático da literatura emerge como uma ferramenta poderosa para a síntese e análise de evidências, fornecendo uma base sólida para a tomada de decisões informadas e o desenvolvimento futuro da pesquisa.

Trata-se de um mapeamento sistemático e descritivo, de natureza qualitativa. A pergunta de pesquisa foi criada de acordo com a estratégia PICO (*Discentes*, Intervenção, Comparação, Desfecho). Em discentes de Medicina (P), como a integração do ensino assistido por *ChatGPT* (I), como complemento aos métodos convencionais de ensino (C), influencia positivamente no aprimoramento do desempenho acadêmico e na satisfação com o processo de aprendizado (O)?

No contexto de uma pesquisa educacional, a estratégia PICO, originalmente desenvolvida para a pesquisa clínica, pode ser adaptada para se ajustar ao campo da educação. Para essa pesquisa a estratégia PICO ficou organizada da seguinte maneira:

Discentes (P): Discentes de Medicina. Este grupo representa os participantes da pesquisa, ou seja, os discentes que estão sendo estudados para compreender os efeitos da intervenção.

Intervenção (I): Integração do ensino assistido por *ChatGPT*. Isso refere-se à aplicação do *ChatGPT* como uma ferramenta complementar aos métodos de ensino tradicionais utilizados na formação médica.

Comparação (C): Métodos convencionais de ensino. Este é o grupo de comparação que representa a condição sem a intervenção do *ChatGPT*, ou seja, os métodos tradicionais de ensino sem a utilização dessa tecnologia.

Desfecho (O): Aprimoramento do desempenho acadêmico e satisfação com o processo de aprendizado. Esses são os resultados ou efeitos que estão sendo avaliados como consequência da intervenção com o *ChatGPT*. O objetivo é determinar se a integração dessa tecnologia influencia positivamente tanto o desempenho acadêmico dos discentes de medicina quanto a sua satisfação com o processo de aprendizagem.

Assim, a pesquisa avaliou o impacto da integração do ensino assistido por *ChatGPT* como complemento aos métodos convencionais de ensino na formação de discentes de Medicina.

O processo de triagem dos estudos encontrados foi conduzido utilizando o instrumento “PRISMA” (*Preferred Reporting Items for Systematic and Meta-Analyses*), seguindo as etapas descritas por Hakeem (2023). Estas etapas são: A - Identificação: Nesta fase, os artigos foram extraídos, armazenados e quantificados, e os estudos duplicados foram eliminados; B - Triagem: Nesta etapa, os títulos, resumos e descritores dos artigos foram revisados para pré-selecionar os estudos que atendiam aos critérios de inclusão estabelecidos para a pesquisa; C - Elegibilidade: Os estudos pré-selecionados foram então avaliados quanto à sua elegibilidade com base nos critérios de inclusão específicos da pesquisa; D - Inclusão: Finalmente, os estudos que atenderam aos critérios de elegibilidade foram incluídos na análise final. É importante notar que estas etapas foram realizadas por dois autores do projeto de forma independente. As discrepâncias encontradas durante o processo de triagem foram posteriormente analisadas e resolvidas através de discussão e consenso entre os dois autores. Este método de triagem sistemática e independente ajuda a garantir a integridade e a confiabilidade do processo de seleção dos estudos a serem incluídos na pesquisa.

Na etapa da elegibilidade, os artigos pré-selecionados na fase anterior foram lidos e avaliados mais detalhadamente para determinar se atendiam aos critérios de inclusão estabelecidos para a pesquisa. Durante essa etapa, os artigos que não estavam diretamente relacionados à pergunta de pesquisa ou que não preenchiam os critérios de inclusão foram descartados. Para a etapa da inclusão, os artigos que passaram pela fase de elegibilidade e foram considerados relevantes para a pesquisa tiveram seus dados desejados coletados. Isso pode incluir informações como os resultados dos estudos, características da população estudada, metodologia utilizada, entre outros dados relevantes para responder à pergunta de pesquisa e cumprir os objetivos do estudo.

As etapas para facilitar a coleta de dados durante o mapeamento sistemático são detalhadas da seguinte forma: 1ª) Definição do tema, seleção da pergunta norteadora e escolha da estratégia de busca: Nesta fase, o tema da pesquisa é claramente definido, uma pergunta norteadora é estabelecida e uma estratégia – nesta pesquisa foram elaboradas duas estratégias de busca: uma longa e uma curta, a elaboração de duas estratégias de busca, uma longa e outra curta, é uma prática

comum em pesquisas que envolvam a revisão sistemática da literatura. Esse procedimento visa contornar restrições ou limitações impostas por algumas bibliotecas virtuais em relação ao tamanho ou complexidade das estratégias de busca - de busca é selecionada, incluindo descritores e bases de dados mais relevantes para levantar as publicações; 2ª) Escolha dos critérios de inclusão e exclusão: São determinados os critérios que definirão quais estudos serão incluídos e quais serão excluídos da análise; 3ª) Identificação dos estudos pré-selecionados e selecionados: Os estudos são pré-selecionados através da leitura dos agentes indexadores das publicações, como resumos, palavras-chave e títulos. Posteriormente, os estudos pré-selecionados são identificados e selecionados; 4ª) Classificação dos estudos selecionados: Os estudos selecionados são classificados, uma matriz de síntese é elaborada e as informações dos estudos são analisadas criticamente. Também é criada uma biblioteca individual contendo os estudos selecionados; 5ª) Análise, interpretação e discussão dos resultados: Os resultados são analisados, interpretados e discutidos à luz da pergunta norteadora e dos objetivos da pesquisa; 6ª) Apresentação da revisão em formato de artigo: Os resultados e conclusões da revisão são apresentados em formato de artigo científico, incluindo propostas para estudos futuros que possam complementar ou expandir a pesquisa realizada.

Essas etapas fornecem uma estrutura sistemática e organizada para conduzir um mapeamento completo da literatura sobre um determinado tema, garantindo que todas as informações relevantes sejam coletadas, analisadas e interpretadas de forma rigorosa e consistente.

Para a seleção dos artigos utilizados neste mapeamento, foi adotada estratégia que envolveu a seleção de palavras-chave, termos de busca e critérios de inclusão para identificar os estudos para serem posteriormente avaliados e incluídos no mapeamento. Foram definidos os descritores: *ChatGPT* (*ChatGPT*), *Teaching* (Ensino), *Medicine* (Medicina), *Learning* (Aprendizagem), *Assessment* (Avaliação), *Evaluation Studies as Topic* (Estudos de Avaliação como Assunto), *Educational Measurement* (Avaliação Educacional), *Education, Medical, Undergraduate* (Educação de Graduação em Medicina). As varreduras foram realizadas a partir das strings de busca, sendo uma longa: *ChatGPT AND Medicine AND (Teaching OR Learning OR Assessment OR Evaluation Studies as Topic OR Educational Measurement OR Education Medical Undergraduate)* e uma curta: *ChatGPT AND medicine AND (teaching OR learning)*.

As varreduras dos artigos foram feitas nas bibliotecas virtuais (*Pubmed*, *Periódicos da CAPES*, *Wiley (online library)*, *BVS*, *Scielo* e *Science direct*) a partir do cruzamento de descritores em português e inglês, por meio de operador booleano (*AND*), constituindo as seguintes *strings* de busca: (longa) português – *ChatGPT AND Medicina AND (Ensino OR Aprendizagem OR Avaliação OR Estudos de Avaliação OR Medição Educacional OR Educação Médica de Graduação)*; (longa) inglês - *ChatGPT AND Medicine AND (Teaching OR Learning OR Assessment OR Evaluation Studies as Topic OR Educational Measurement OR Education Medical Undergraduate)*; (curta) português - *ChatGPT AND medicina AND (ensino OR aprendizagem)*; (curta) inglês - *ChatGPT AND medicine AND (teaching OR learning)*.

Os critérios de inclusão e exclusão são fundamentais para garantir que os estudos selecionados sejam relevantes e atendam aos objetivos da pesquisa. Os critérios de inclusão aqui utilizados foram: texto completo disponível gratuitamente, do tipo: artigo original e de revisão e publicados no período de 2019 a 2024. Enquanto os de exclusão foram: artigos repetidos e que não contemplavam a temática sobre a eficácia do *ChatGPT* no ensino de medicina.

A aplicação rigorosa desses critérios durante o processo de seleção dos estudos contribuirá para a qualidade e relevância dos trabalhos incluídos na pesquisa.

A análise de conteúdo realizada por meio de um *Chatbot (ChatGPT)*. Durante o processo, o *ChatGPT* foi programado para identificar como a integração do ensino assistido por *ChatGPT*, como complemento aos métodos convencionais de ensino, influencia positivamente no aprimoramento do desempenho acadêmico e na satisfação do aprendizado em medicina e categorizar os possíveis encontros e desencontros dos discentes com base em certos critérios estabelecidos. Os critérios puderam incluir questões relacionadas à experiência dos discentes com o ensino assistido por *ChatGPT*, seus desafios, benefícios percebidos, áreas de melhoria e assim por diante, utilizando os fundamentos da análise de conteúdo de Bardin (2011), é uma abordagem inovadora e interessante que envolve várias etapas, como pré-análise, exploração do material, tratamento dos resultados obtidos e interpretação dos dados. Através dessas etapas, as categorias temáticas podem ser identificadas e os padrões emergentes nos dados podem ser compreendidos e interpretados e assim fornecer uma compreensão mais profunda das percepções e experiências dos discentes de medicina em relação ao ensino assistido por *ChatGPT*. Esta metodologia permite a análise sistemática e a extração de padrões e temas dos 50 artigos selecionados.

O Quadro 1 contém as etapas do mapeamento sistemático*.

*Quadro 1 – Detalhamento das etapas do Mapeamento Sistemático.

TÓPICOS DO MAPEAMENTO SISTEMÁTICO	DETALHAMENTO DE CADA TÓPICO
Tema	A Eficácia do <i>ChatGPT</i> no Ensino de Medicina por Meio de Mapeamento Sistemático
Pergunta norteadora	Em discentes de Medicina (P), como a integração do ensino assistido por <i>ChatGPT</i> (I), como complemento aos métodos convencionais de ensino (C), influencia positivamente no aprimoramento do desempenho acadêmico e na satisfação com o processo de aprendizado (O)?
Objetivo geral	Analisar a utilidade do <i>ChatGPT</i> como uma ferramenta de ensino na área da medicina, além de fornecer orientações para sua aplicação eficaz e sua integração bem-sucedida nos ambientes educacionais médicos.
Estratégias de busca	Cruzamento de descritores por meio do operador booleano <i>AND</i>, <i>OR</i>. Uso de aspas nos politermos (descriptor com mais de um termo) para que a varredura de artigos científicos contemplasse o termo exato; Uso de descritores estruturados (codificação) no DECS ou MESH; Uso de metadados (filtros) nas bibliotecas virtuais; Uso de descritores em inglês para ampliar o número de artigos. Português – <i>ChatGPT AND Medicina AND (Ensino OR Aprendizagem OR Avaliação OR Estudos de Avaliação OR Medição Educacional OR Educação Médica de Graduação) / ChatGPT AND medicina AND (ensino OR aprendizagem),</i>

	Inglês - <i>ChatGPT AND Medicine AND (Teaching OR Learning OR Assessment OR Evaluation Studies as Topic OR Educational Measurement OR Education Medical Undergraduate)/ ChatGPT AND medicine AND (teaching OR learning).</i>		
Bancos de terminologias	Banco	Link	
	DeSC	http://decs.bvs.br/	
	MeSH	https://www.ncbi.nlm.nih.gov/mesh	
Descritores livres e estruturados	Descritor	DeCS (Registro)	MeSH (Identificador Único)
	<i>ChatGPT (ChatGPT)</i>	*	*
	<i>Teaching (Ensino)</i>	14042	D013663
	<i>Medicine (Medicina)</i>	8678	D008511
	<i>Learning (Aprendizagem)</i>	8036	D007858
	<i>Artificial intelligence (Inteligência Artificial)</i>	22729	D001185
	<i>Evaluation Studies as Topic (Estudos de Avaliação como Assunto)</i>	5166	D005069
	<i>Educational Measurement (Avaliação Educacional)</i>	4591	D004521
	<i>Education, Medical, Undergraduate (Educação de Graduação em Medicina)</i>	4572	D004504
String de busca	Longa:		Curta:
	1. Português: <i>ChatGPT AND Medicina AND (Ensino OR Aprendizagem OR Avaliação OR Estudos de Avaliação OR Medição Educacional OR Educação Médica de Graduação)</i> 1. Inglês: <i>ChatGPT AND Medicine AND (Teaching OR Learning OR</i>		2. Português: <i>ChatGPT AND medicina AND (ensino OR aprendizagem)</i> 2. Inglês: <i>ChatGPT AND medicine AND (teaching OR learning)</i>

	Assessment OR Evaluation Studies as Topic OR Educational Measurement OR Education Medical Undergraduate)		
Bibliotecas Virtuais	Biblioteca	Link	
	BVS	http://brasil.bvs.br/	
	Wiley	https://onlinelibrary.wiley.com/	
	ScienceDirect	https://www.sciencedirect.com/	
	Periódicos da CAPES	https://www.periodicos.capes.gov.br/	
	Pubmed	https://pubmed.ncbi.nlm.nih.gov/	
	Scielo	https://www.scielo.br/	
Período de coleta dos dados	Dezembro de 2023 a fevereiro de 2024		
Critérios de inclusão	Texto completo do tipo: artigo científico. Publicação (2019 - 2024).		
Critérios de exclusão	1. Artigos repetidos; 2. Artigos que não contemplam a relação entre interações do <i>ChatGPT</i> com ensino em medicina.		
Número de trabalhos selecionados para mapeamento sistemático a partir da leitura dos agentes indexadores das publicações (resumo, palavras-chave e título) e resultados, os quais devem conter os descritores utilizados neste estudo	50		

Categorias obtidas com a análise dos trabalhos científicos investigados	1. Aplicações do <i>ChatGPT</i> no Ensino Médico 2. Considerações éticas na utilização do <i>ChatGPT</i> no Ensino Médico: subcategoria chamada confiabilidade 3. Desafios do <i>ChatGPT</i> no Ensino Médico 4. Contribuições do uso do <i>ChatGPT</i> 5. Desempenho do <i>ChatGPT</i> em exames de escolas de medicina 6. Direcionamentos futuros e recomendações para a integração do <i>ChatGPT</i> no Ensino Médico		
Análise, interpretação e discussão dos resultados	Ver em “Resultados e Discussão”		
Tecnologias digitais utilizadas	Tecnologia (<i>software</i> ou <i>website</i>)	Link	Utilidade
	<i>ChatGPT</i> 3.5	https://chat.openai.com/	Criação das categorias temáticas.
Apresentação da revisão em formato de artigo, o qual contemple propostas para estudos futuros	Este artigo.		

*Não foi encontrada, porém faz parte como palavra imprescindível para esta pesquisa.

3. Resultados

Ao término da busca, foram identificados 50 artigos, que estão listados na Tabela 1.

Tabela 1 – Varreduras nas bibliotecas virtuais.

STRING BUSCA	DE	<i>biblioteca virtual</i>	<i>Total de publicações sem o filtro “assunto principal”</i>	<i>Textos completos disponíveis após aplicar os filtros</i>	<i>Textos aproveitados no Mapeamento Sistemático</i>	<i>Textos (excluídos ou repetidos)</i>
<i>ChatGPT Medicine (Teaching Learning Assessment Evaluation Studies as Topic</i>	AND	<i>Science Direct</i>	955	278	6	
	OR	<i>Periódicos CAPES</i>	468	270	16	

OR Educational Measurement OR Education Medical Undergraduate)	Pubmed	605	334	27	
ChatGPT medicine (teaching learning) AND AND OR	BVS Brasil Medline)	87	84	11	
	SCIELO	1	1	0	
	Wiley (online Library)	402	142	6	
TOTAL	6	2.518	1.109	66	50

Fonte: Desenvolvida conforme pesquisa realizada (2023).

Realizou-se um fichamento dos artigos utilizados na análise do Mapeamento Sistemático (Quadro 2).

Ver em: <http://educapes.capes.gov.br/handle/capes/972290>

4. Discussão

Durante a condução desta pesquisa qualitativa, utilizou-se o *ChatGPT* como uma ferramenta de análise de conteúdo, fundamentada em inteligência artificial, visando à organização das categorias temáticas. A análise foi realizada com base nos textos correspondentes aos resultados e conclusões dos 50 artigos selecionados (conforme detalhado no Quadro 2). Em resposta à consulta ao *ChatGPT*, o sistema sugeriu a identificação de seis categorias temáticas, acompanhadas de suas respectivas explicações.

Todavia, ao explorar as interações geradas pelo modelo de linguagem, identificou-se a necessidade de descartar uma categoria e realizar o ajuste das seis restantes por meio de inteligência humana, sendo elas: 1- Aplicações do *ChatGPT* na educação médica; 2- Considerações éticas na utilização do *ChatGPT* no ensino médico; 3 – Desafios do *ChatGPT* no ensino médico; 4 – Contribuições do uso do *ChatGPT* no ensino médico; 5- Despenho do *ChatGPT* em exames de escolas médicas; 6- Direcionamentos futuros e recomendações para integração do *ChatGPT* no ensino médico. Segue, abaixo, a discussão das seis categorias.

Categoria 1 - Aplicações do *ChatGPT* na educação médica

O uso de tecnologias avançadas, como o *ChatGPT*, tem se tornado um tema central na educação médica contemporânea. Tamsah et al. (2023) avultam que essa ferramenta pode aprimorar a aprendizagem e a interpretação de informações médicas, fornecendo dados comparáveis aos de discentes de medicina em áreas específicas. Contudo, há uma necessidade de padronização e validação das informações, pois a dependência excessiva do *ChatGPT* pode prejudicar o raciocínio clínico dos discentes, e suas respostas podem faltar contexto ou não se adaptar às necessidades individuais dos *discentes*.

Preiksaitis e Rose (2023) identificam diversas aplicações potenciais para a IA generativa na educação médica, como a aprendizagem autodirigida, cenários de simulação e assistência à escrita, além de promover uma reflexão crítica sobre a precisão das informações em contextos clínicos e educacionais. Corroborando com esse potencial, Alkhaaldi et al. (2023) destacam que o *ChatGPT* oferece experiências de aprendizado personalizadas e feedback imediato, apoiando os discentes na correção de seus erros e contribuindo para um ambiente de aprendizado adaptativo.

Nesse contexto, Ghorashi et al. (2023) ressaltam que *Chatbots* como o *ChatGPT* podem funcionar como tutores interativos, auxiliando os discentes e educadores em treinamentos clínicos desafiadores por meio de simulações de

pacientes, demonstrando sua aplicabilidade em situações práticas. Lee (2023) reforça essa visão ao evidenciar que a ferramenta obteve uma taxa de precisão superior a 60% no Exame de Licenciamento Médico dos Estados Unidos (USMLE), fornecendo insights clínicos abrangentes e coerentes, o que sugere um uso promissor do *ChatGPT* na formação médica.

Complementando essa perspectiva, Wójcik et al. (2023) afirmam que o *ChatGPT* pode atuar como um "auxiliar pedagógico virtual", oferecendo feedback e esclarecimentos sobre conceitos desafiadores, criando um ambiente de aprendizado mais envolvente para os discentes. Hosseini et al. (2023) adicionam que o *ChatGPT* teve um desempenho próximo do nível de aprovação em exames do USMLE, sendo especialmente utilizado por estagiários de medicina e pesquisadores, indicando sua crescente aceitação na área médica.

Apesar dos benefícios, Choi (2023) alerta para a necessidade de adotar estratégias que mitiguem riscos, como o plágio e a disseminação de informações imprecisas, ao invés de simplesmente proibir o uso da tecnologia. Em concordância, Abd-Alrazaq et al. (2023) reforçam que a integração do *ChatGPT* no ensino médico pode enriquecer o processo educacional, inclusive na produção de estudos de casos clínicos e simulações com pacientes virtuais.

Ignjatović e Stevanović (2023) complementam ao destacar o uso do *ChatGPT* na criação de perguntas de múltipla escolha e na oferta de informações precisas para discentes, embora ressaltem a necessidade de pesquisas adicionais para confirmar a padronização e integridade da ferramenta. Bagde et al. (2023) ampliam a discussão ao sugerir que o *ChatGPT* também pode ser um recurso valioso para profissionais da saúde em diagnósticos e planos de tratamento, demonstrando sua versatilidade no campo médico.

De forma semelhante, Totlis et al. (2023) sugerem que o *ChatGPT* pode ser utilizado eficazmente no ensino de anatomia, permitindo que os discentes interajam com a ferramenta para obter respostas detalhadas e personalizadas. Breeding et al. (2023) acrescentam que o *ChatGPT* pode melhorar habilidades de anamnese e criar simulações padronizadas, fortalecendo a prática clínica dos discentes.

Seguindo a mesma linha, Cross et al. (2023) afirmam que o *ChatGPT* tem sido amplamente utilizado para a criação de perguntas de múltipla escolha e vinhetas clínicas, o que contribui para a eficiência dos professores na preparação de avaliações somativas. Complementarmente, Roos et al. (2023) ressaltam que o *ChatGPT*, alimentado por modelos avançados como GPT-3.5-Turbo e GPT-4, é uma ferramenta valiosa que gera respostas de alta qualidade, promovendo o envolvimento e o pensamento crítico dos discentes.

Além disso, Khlaif et al. (2023) destacam que o *ChatGPT* revolucionou a educação e a pesquisa médica ao proporcionar acesso em tempo real a informações médicas atualizadas, facilitando o aprendizado contínuo e autodirigido. Weidener e Fischer (2024) corroboram ao observar que o *ChatGPT* teve sucesso em realizar a parte escrita do USMLE, sinalizando seu potencial para o contexto educacional médico.

Na mesma linha, Knopp et al. (2023) afirmam que a IA, incluindo o *ChatGPT*, oferece um aprendizado mais envolvente e personalizado, apesar dos desafios relacionados ao acesso e aos custos envolvidos. Park (2023) complementa que o *ChatGPT* se diferencia por fornecer informações interativas e personalizadas, contribuindo para um processo de aprendizado mais autônomo e significativo.

Ainda no mesmo contexto, Ghosh e Bir (2023) destacam que o *ChatGPT* pode responder a perguntas que exigem pensamento de nível superior, sendo

especialmente útil em disciplinas como a bioquímica médica. Huang et al. (2023) reforçam essa observação, afirmando que o *ChatGPT* tem capacidades avançadas para criar cenários de exames e fornecer informações precisas sobre serviços comunitários.

Morjaria et al. (2023) ressaltam que o *ChatGPT* foi capaz de auxiliar discentes de medicina com precisão em exames complexos, demonstrando seu potencial de aplicação em diversas formas de avaliação. Em complemento, Ilgaz e Çelik (2023) destacam como plataformas de IA, como *ChatGPT* e Google Bard, podem contribuir para o ensino de anatomia, auxiliando na elaboração de artigos e no questionamento de conceitos-chave.

Ao discutir a elaboração de avaliações, Agarwal, Sharma e Goswami (2023) avaliaram a aplicabilidade do *ChatGPT* na criação de questões de múltipla escolha, indicando que há espaço para aprimorar essas habilidades. Mahajan et al. (2023) concordam e reforçam que a IA ainda precisa de aprimoramento para criar questões MCQs de alta qualidade.

Cheung et al. (2023) sugerem que o *ChatGPT* pode transformar a avaliação médica, enquanto Panthier e Gatinel (2023) confirmam seu potencial ao observar o desempenho da ferramenta em exames de oftalmologia. Wong (2023) também ressalta a competência do *ChatGPT* ao superar outras ferramentas em avaliações como o USMLE.

Weidener e Fischer (2023) destacam a capacidade do *ChatGPT* de auxiliar no ensino da ética da IA, abordando questões como privacidade e transparência. Huynh, Bonebrake e Schultis (2023) reforçam que uma das principais aplicações da IA na educação médica é a adoção de modelos como o *ChatGPT*.

Sonntagbauer, Haar e Kluge (2023) enfatizam que o *ChatGPT* apresenta aplicações significativas na medicina, especialmente no suporte à redação científica. Ismail Dergaa et al. (2023) ressaltam que o *ChatGPT* pode incentivar educadores a reavaliar suas metodologias de ensino, ampliando seu potencial pedagógico.

Sallam (2023) acrescenta que o *ChatGPT* pode melhorar a redação científica e contribuir para a prática e educação em saúde. Cheng et al. (2023) observam que o *ChatGPT* pode ser usado para preencher prontuários médicos, aliviando a burocracia dos profissionais de saúde. Qu et al. (2023) afirmam que, apesar de seu rápido crescimento na medicina, a eficácia do *ChatGPT* ainda precisa ser amplamente avaliada.

No desenvolvimento de vinhetas clínicas, Bakkum et al. (2023) confirmam que o *ChatGPT* foi eficaz no desenvolvimento de casos médicos inclusivos e diversos. Em complemento, Vaira et al. (2023) relatam que o *ChatGPT* tem sido útil na formulação de perguntas e cenários clínicos, apoiando o aprendizado dos discentes.

Preiksaitis e Rose (2023) enfatizam que o *ChatGPT* está sendo utilizado em tarefas acadêmicas, como a redação de cartas de recomendação, demonstrando sua adaptabilidade para contextos educacionais. Tovalino et al. (2023) destacam a utilidade do *ChatGPT* em ambientes acadêmicos, particularmente na geração de linguagem natural e na resolução de problemas complexos.

Amedu e Ohene-Botwe (2024) enfatizam que o *ChatGPT* atua como tutor virtual no ensino de radiografia, facilitando o aprendizado interativo e individualizado. Ngo et al. (2024) destacam que a ferramenta também pode gerar perguntas de múltipla escolha com explicações detalhadas, auxiliando na preparação de exames práticos.

Li (2023) observa que o *ChatGPT* pode fornecer à equipe médica e discentes acesso rápido a informações atualizadas, auxiliando no desenvolvimento profissional

e no aprendizado contínuo. Dave, Athaluri e Sing (2023) acrescentam que o *ChatGPT* pode apoiar atividades de diagnóstico clínico e pesquisa.

Das et al. (2023) destacam que o *ChatGPT* é eficaz em tópicos como microbiologia, promovendo o aprendizado baseado em competências. Huh (2023) afirma que a ferramenta pode atuar como tutor em tarefas de casa, ajudando os discentes de medicina a compreender conceitos complexos. Tsang (2023) observa que o *ChatGPT* pode ser um "assistente de ensino virtual," reforçando a autonomia do aprendizado. Brin et al. (2023) finalizam ao afirmar que o *ChatGPT* pode atender às demandas de habilidades sociais, como empatia e julgamento, essenciais na prática médica.

Por fim, Meo et al. (2023) destacam o potencial do *ChatGPT* para auxiliar discentes e professores em processos educacionais e científicos, enquanto Jeyaraman et al. (2023) ressaltam o imenso potencial da ferramenta na educação, gestão clínica e pesquisa. Chen et al. (2023) finalizam a discussão ao destacar que o *ChatGPT* pode processar informações médicas especializadas, auxiliando em diagnósticos e gerenciamento de casos complexos, demonstrando eficácia em exames radiológicos.

Diante do exposto, as 50 citações demonstram que o *ChatGPT* possui uma ampla gama de aplicações no ensino médico, destacando-se como uma ferramenta complementar para aprimorar a aprendizagem, prática clínica e o desenvolvimento de habilidades críticas. Sua integração no ambiente educacional proporciona oportunidades para experiências de aprendizado personalizadas, interativas e eficazes. No entanto, para maximizar seu potencial, é fundamental garantir a padronização, validação e supervisão do uso do *ChatGPT*, assegurando que ele atenda às necessidades individuais dos discentes e apoie seu raciocínio clínico. Ao ser utilizado de forma complementar e consciente, o *ChatGPT* pode desempenhar um papel transformador no campo da educação médica, promovendo a excelência no processo de ensino e aprendizagem.

Categoria 2 - Considerações éticas na utilização do *ChatGPT* no ensino médico

O uso crescente de inteligência artificial (IA), como o *ChatGPT*, no ensino médico, tem sido alvo de discussões éticas devido a preocupações com a precisão e a confiabilidade das informações geradas. Cheung et al. (2023) apontam que um dos maiores desafios está relacionado ao risco de informações incorretas, o que pode prejudicar o aprendizado e a formação dos discentes. Complementando essa visão, Wong (2023) alerta para o fenômeno das "alucinações artificiais", onde o *ChatGPT* gera informações fabricadas ou inverídicas, impactando negativamente a integridade do processo educacional. Portanto, é essencial explorar como essas preocupações éticas se manifestam e como os diferentes autores debatem essas questões.

Dentro desse contexto, Weidener e Fischer (2023) ressaltam a importância de uma reflexão profunda sobre a responsabilidade associada ao uso da IA no ensino médico, destacando a necessidade de supervisão humana para evitar a disseminação de informações incorretas. Isso reforça a ideia apresentada por Panthier e Gatinel (2023), que argumentam que, apesar do desempenho impressionante do *ChatGPT*, existem preocupações em relação à interpretação de informações médicas, especialmente quando se trata de detalhes e atualizações recentes, que a IA pode não captar com precisão.

A questão da supervisão e da responsabilidade é também enfatizada por Huynh, Bonebrake e Schultis (2023), que apontam que a IA pode fornecer respostas plausíveis, mas que podem induzir à desinformação se não forem avaliadas

criticamente. Assim, eles sugerem que a supervisão humana deve ser um componente essencial na aplicação da IA na educação médica, para garantir a precisão e evitar a perpetuação de erros.

Sonntagbauer, Haar e Kluge (2023) concordam que o uso de *Chatbots* como o *ChatGPT* em ambientes educacionais apresenta riscos de plágio e a dificuldade de controlar a precisão dos conteúdos gerados, o que pode levar a desafios éticos. Isso se relaciona com a preocupação expressa por Ismail Dergaa et al. (2023), que afirmam que, apesar das capacidades do *ChatGPT*, a responsabilidade pela autoria de artigos científicos deve permanecer com os humanos para manter a integridade acadêmica.

Essa necessidade de manter a integridade é reforçada por Sallam (2023), que destaca que o risco de plágio e a potencial imprecisão das informações geradas pelo *ChatGPT* exigem cautela ao incorporar essa tecnologia na educação médica. Isso sugere que, embora o *ChatGPT* possa ser uma ferramenta poderosa, ele deve ser usado de forma complementar e sob cuidadosa supervisão humana.

A limitação do *ChatGPT* em capturar nuances em contextos clínicos é enfatizada por Cheng et al. (2023), que alertam para sua incapacidade de interpretar variações sutis nas descrições de sintomas dos pacientes. De forma semelhante, Qu et al. (2023) destacam que o *ChatGPT* ainda está em uma fase inicial de implementação na medicina e pode produzir respostas inconsistentes, o que levanta preocupações éticas sobre sua aplicação no ensino.

Nesse sentido, Ghosh e Bir (2023) acrescentam que o fato de o *ChatGPT* ser alimentado por uma base de dados limitada até 2021 torna-o potencialmente desatualizado, especialmente em áreas da medicina que evoluem rapidamente. Esse ponto reforça a importância de avaliação constante de suas capacidades e limitações, para garantir que as informações fornecidas sejam relevantes e atualizadas.

Huang et al. (2023) também destacam que, mesmo com as capacidades avançadas do GPT-4, ainda existe o risco de gerar "alucinações", o que pode comprometer a integridade e a precisão das informações apresentadas aos discentes. Isso reforça a necessidade de supervisão e verificação das informações, garantindo que a IA não induza os *discentes* a aceitar respostas incorretas como verdadeiras.

Adicionalmente, Bakkum et al. (2023) sugerem que, embora o *ChatGPT* possa gerar respostas bem estruturadas, ele não possui um entendimento médico real, o que pode resultar em erros, como cálculos incorretos ou atribuições equivocadas de gênero e etnia. Dessa forma, torna-se fundamental que os educadores monitorem e corrijam esses erros para evitar a perpetuação de informações incorretas.

A preocupação com a possibilidade de fornecer informações desatualizadas ou incorretas também é enfatizada por Vaira et al. (2023), que alertam para o fato de o *ChatGPT* se basear em dados disponíveis apenas até setembro de 2021. Preiksaitis et al. (2023) também mencionam que a IA pode gerar textos enviesados ou "alucinações", especialmente ao criar cartas de recomendação ou outras redações acadêmicas.

Morjaria et al. (2023) ressaltam a importância de critérios de avaliação detalhados para identificar respostas geradas pela IA e distinguir entre o trabalho de *discentes* e o produzido pelo *ChatGPT*. Da mesma forma, Ilgaz e Çelik (2023) aconselham que o uso de modelos de linguagem de grande porte (*LLMs*) na educação médica deve ser feito com cautela, uma vez que esses modelos ainda estão em desenvolvimento e podem gerar informações imprecisas.

Essas considerações éticas também são abordadas por Tovalino et al. (2023), que apontam o risco de o *ChatGPT* produzir conteúdo enganoso ou discriminatório,

além das limitações na redação científica. Amedu e Ohene-Botwe (2024) reforçam que o uso do *ChatGPT* na educação em saúde pode comprometer a privacidade dos pacientes e a segurança dos dados, além de gerar uma dependência excessiva de sistemas de IA, afetando a integridade acadêmica.

Ngo et al. (2024) chamam atenção para o risco de violação de privacidade, já que o *ChatGPT* pode analisar dados pessoais dos discentes para personalizar o ensino, o que pode ser invasivo. Essa questão é ecoada por Li (2023), que destaca que o uso de *Chatbots* na saúde pode apresentar desafios éticos em relação à exposição de dados sensíveis dos pacientes.

Ao discutir as diretrizes éticas que o *ChatGPT* segue, Dave, Athaluri e Sing (2023) mencionam que, embora a IA siga orientações éticas da União Europeia, ainda há preocupações sobre autoria e responsabilidade pelo conteúdo gerado. Para aprimorar a precisão e confiabilidade, Das et al. (2023) sugerem que a revisão por múltiplos avaliadores pode ajudar a reduzir vieses nas respostas da IA.

Nesse sentido, Agarwal, Sharma e Goswami (2023) e Mahajan et al. (2023) destacam que estudos realizados sem a participação de sujeitos humanos não necessitam de revisão ética, ressaltando que as considerações éticas no uso da IA devem ser adaptadas ao contexto de cada aplicação.

Além disso, Huh (2023) enfatiza que é crucial que os discentes de medicina desenvolvam a habilidade de avaliar criticamente as informações fornecidas pela IA. Tsang (2023) alerta para o risco de má conduta acadêmica, com discentes usando o *ChatGPT* para produzir trabalhos escritos que apresentam como seus, desafiando a integridade acadêmica.

Brin et al. (2023) observam que o GPT-4 superou o *ChatGPT* em lidar com dilemas éticos de forma empática, mas Meo et al. (2023) ressaltam que ainda há ambiguidade sobre o uso ideal do *ChatGPT* na educação médica. Jeyaraman et al. (2023) concordam que, apesar do potencial do *ChatGPT*, ele não pode substituir a experiência e o julgamento humano, especialmente no contexto da educação médica e pesquisa.

Chen et al. (2023) reforçam a necessidade de verificação humana das respostas do *ChatGPT*, destacando que ele pode "alucinar" ao gerar respostas persuasivas, mas incorretas. Temsah et al. (2023) complementam ao sugerir que o treinamento do *ChatGPT* deve ser supervisionado para evitar vieses e desinformação, enquanto Alkhaaldi et al. (2023) alertam sobre os riscos do preconceito, violação de direitos autorais e desumanização no processo de aprendizagem.

Abd-Alrazaq et al. (2023) e Ignjatović e Stevanović (2023) destacam os riscos de vieses e informações enganosas geradas pela IA, enquanto Bagde et al. (2023) e Totlis et al. (2023) sublinham a necessidade de que o *ChatGPT* complemente, e não substitua, a orientação de educadores humanos. Breeding et al. (2023) e Cross et al. (2023) enfatizam a importância de avaliar cuidadosamente o uso do *ChatGPT* antes de adotá-lo amplamente como ferramenta educacional.

Roos et al. (2023) lembram que, apesar das melhorias do GPT-4, a supervisão humana permanece necessária para evitar "alucinações," enquanto Khlaif et al. (2023) reforçam a importância do julgamento humano. Weidener e Fischer (2024) alertam que o uso crescente de IA traz desafios éticos que precisam ser enfrentados pelos futuros profissionais médicos. Knopp et al. (2023) concluem que a desconfiança dos pacientes em relação à tecnologia pode impactar negativamente sua adesão a tratamentos, ressaltando a necessidade de ética e precisão no uso do *ChatGPT*.

Por fim, autores como Ghorashi et al. (2023) defendem que os *Chatbots* devem citar claramente fontes confiáveis para evitar a disseminação de referências falsas.

Lee (2023) observa que a dependência do *ChatGPT* em entradas de texto limita sua capacidade de considerar outros sinais clínicos. Wójcik et al. (2023) e Hosseini et al. (2023) reforçam a necessidade de uso responsável da IA, recomendando verificações constantes para evitar o uso inadequado em exames.

As considerações éticas em torno do uso do *ChatGPT* na educação médica apontam para a necessidade de uma abordagem cuidadosa e responsável. A supervisão humana, a avaliação crítica e a formação ética dos futuros profissionais médicos são essenciais para garantir que o *ChatGPT* seja utilizado de forma segura e eficaz, servindo como uma ferramenta complementar que enriquece o aprendizado, mas que não substitui o papel indispensável da expertise e julgamento humanos.

Subcategoria: Confiabilidade do *ChatGPT* no ensino médico

A confiabilidade do *ChatGPT* no ensino médico tem sido um tema amplamente debatido, com vários autores destacando seus potenciais e limitações. Temsah et al. (2023) iniciam essa discussão ao alertar que, embora o *ChatGPT* apresente um grande potencial, ele tem uma tendência preocupante de gerar conteúdo que pode parecer plausível, mas que, na verdade, é incorreto. Isso compromete seriamente a integridade científica e a precisão das informações médicas, algo fundamental na educação médica, onde a exatidão é crucial para a formação de futuros profissionais.

Corroborando com essa visão, Preiksaitis e Rose (2023) ressaltam que a precisão e a confiabilidade do *ChatGPT* estão intrinsecamente ligadas à sua base de dados de treinamento. Eles chamam a atenção para o fato de que a IA pode apresentar referências inexistentes e dados enganosos, o que mina a confiança dos usuários em seus resultados. Essa questão é particularmente importante no ensino médico, pois os discentes dependem de informações precisas e verificáveis para desenvolver suas competências clínicas.

Em consonância, Alkhaaldi et al. (2023) enfatizam a tendência do *ChatGPT* de "alucinar", ou seja, gerar respostas que parecem lógicas, mas que são incorretas ou ilógicas. Essa característica pode induzir os discentes ao erro e comprometer o aprendizado, reforçando a necessidade de uma supervisão crítica ao utilizar o *ChatGPT* como ferramenta educacional. Choi (2023) reforça essa preocupação ao apontar que, embora o *ChatGPT* tenha atingido uma precisão geral de 76%, seu desempenho foi inconsistente em questões de resolução de problemas. Ele frequentemente gerou informações incorretas, afetando sua confiabilidade e demonstrando que, apesar de parecer convincente, o *ChatGPT* pode induzir a interpretações errôneas.

Na mesma linha, Ignjatović e Stevanović (2023) reconhecem que o *ChatGPT* tem potencial para fornecer informações precisas, mas alertam que mais pesquisas são necessárias para garantir sua padronização e confiabilidade. A confiabilidade da IA não pode ser assumida sem uma validação rigorosa, especialmente em um campo tão crítico como a medicina.

Expansivamente, Bagde et al. (2023) destacam que, embora o *ChatGPT* demonstre precisão promissora, sua confiabilidade é variável, e a heterogeneidade dos resultados em diferentes estudos aponta para a necessidade de mais investigação. Isso sugere que a ferramenta, apesar de útil, ainda precisa ser melhor compreendida e aprimorada para garantir sua eficácia e confiabilidade no contexto educacional.

Totlis et al. (2023) também contribuem para o debate ao afirmar que a utilização do *ChatGPT* deve ser acompanhada por uma avaliação crítica dos educadores. Eles destacam que, apesar de o *ChatGPT* oferecer respostas detalhadas e insights

valiosos, a IA pode apresentar limitações em termos de precisão e, portanto, não deve ser utilizada como uma fonte isolada de conhecimento. O papel dos educadores é, portanto, essencial para validar e complementar as informações geradas pela IA.

Breeding et al. (2023) expandem essa discussão ao sugerir que a confiabilidade do *ChatGPT* é ainda mais questionável no contexto da produção acadêmica. Eles observam que a ferramenta tem sido vista como um atalho para substituir a escrita acadêmica, mas a falta de mecanismos eficazes para detectar plágio gerado por IA é uma preocupação significativa. Isso levanta questões sobre a originalidade e a autenticidade do conteúdo produzido com a assistência do *ChatGPT*, especialmente em um ambiente que valoriza a ética e a precisão.

Enfim, Khlaif et al. (2023) trazem à tona um aspecto crítico ao apontar que o *ChatGPT* pode fabricar citações e informações, comprometendo seriamente a integridade da pesquisa acadêmica. Eles destacam que essa característica limita a confiabilidade do *ChatGPT* como ferramenta de pesquisa e reforçam a necessidade de validação e supervisão humana para garantir que as informações utilizadas pelos discentes sejam verdadeiras e úteis.

Em síntese, os autores apresentam uma visão consensual de que, apesar do potencial do *ChatGPT* para apoiar o ensino médico, sua confiabilidade é um aspecto que requer atenção e supervisão constantes. O *ChatGPT* pode ser uma ferramenta poderosa para auxiliar no processo de aprendizado, mas seu uso deve ser sempre acompanhado por uma avaliação crítica e pela orientação de educadores qualificados, garantindo que o conteúdo seja preciso, relevante e que os discentes desenvolvam habilidades de pensamento crítico e análise. A integração responsável do *ChatGPT* no ensino médico, portanto, depende de um equilíbrio cuidadoso entre o uso da tecnologia e a expertise humana.

Categoria 3 – Desafios do uso do *ChatGPT* no ensino médico

A introdução do *ChatGPT* na educação médica tem gerado um intenso debate sobre suas fragilidades e os desafios que ele enfrenta para ser uma ferramenta realmente eficaz. Sallam (2023) inicia essa discussão ao apontar que, apesar do potencial do *ChatGPT*, ele é suscetível a introduzir informações falsas, vieses e problemas de plágio, o que pode impactar negativamente o ensino médico. Essa questão é reforçada por Cheng et al. (2023), que destacam a tendência do *ChatGPT* em "alucinar" fatos, criando respostas que, embora pareçam estatisticamente plausíveis, podem ser incorretas e enganosas. O desafio de distinguir entre informações precisas e fabricadas é, portanto, uma preocupação central na aplicação desta tecnologia.

Ampliando essa discussão, Qu et al. (2023) ressaltam que a limitação do *ChatGPT* em processar informações além de 2021 agrava a questão de fornecer dados precisos e atualizados, uma limitação que é ainda mais preocupante no campo da medicina, que requer informações em constante evolução. Bakkum et al. (2023) também destacam que o *ChatGPT* tende a refletir preconceitos e falta de diversidade, frequentemente fornecendo respostas que carecem de rigor técnico e precisão biológica, o que pode limitar seu uso em situações clínicas diversificadas.

Essa falta de precisão é particularmente preocupante quando aplicada ao diagnóstico clínico. Vaira et al. (2023) relatam que o *ChatGPT* apresentou uma taxa de erro significativa em diagnósticos clínicos, além de fornecer referências bibliográficas inexistentes. Em consonância, Preiksaitis et al. (2023) apontam que a dificuldade de distinguir entre cartas de recomendação geradas por IA e aquelas

escritas por humanos compromete a integridade acadêmica, sugerindo que o *ChatGPT* não é confiável como fonte de conteúdo original e verídico.

Além dos desafios de precisão, há preocupações éticas levantadas por Tovalino et al. (2023), que observam questões de privacidade e propriedade intelectual associadas ao uso do *ChatGPT* na medicina. Amedu e Ohene-Botwe (2024) complementam ao afirmar que a incapacidade do *ChatGPT* de compreender contextos complexos e a variabilidade de suas respostas podem induzir ao erro, destacando a necessidade de supervisão humana rigorosa ao utilizar a ferramenta em um ambiente educacional.

Ngo et al. (2024) reforçam a ideia de que o *ChatGPT* pode não ser eficiente em ambientes de ensino que exigem precisão e supervisão, pois uma parte considerável das respostas fornecidas pela IA foi incorreta ou exigiu ajustes significativos. Li (2023) também observa que, por serem artificiais, *Chatbots* como o *ChatGPT* não têm a capacidade de expressar empatia, um componente essencial na prática médica e no processo de ensino.

A capacidade do *ChatGPT* de substituir a expertise humana é contestada por Dave, Athaluri e Sing (2023), que destacam a incapacidade da IA em identificar informações relevantes e diferenciar fontes confiáveis de não confiáveis. Em apoio a essa perspectiva, Das et al. (2023) identificam que a inconsistência do *ChatGPT* em fornecer respostas precisas em tópicos como microbiologia limita sua eficácia como ferramenta de aprendizagem.

Enquanto isso, Temsah et al. (2023) e Preiksaitis e Rose (2023) argumentam que, mesmo com seus potenciais, o *ChatGPT* pode prejudicar o desenvolvimento do pensamento crítico, especialmente quando há uma dependência excessiva da tecnologia, pois a IA pode não fornecer informações atualizadas ou completas. Essa fragilidade é reforçada por Alkhaaldi et al. (2023), que aponta preocupações relacionadas à ética, como preconceito e violações de direitos autorais.

Outro ponto de discussão refere-se à confiabilidade das perguntas padronizadas geradas pelo *ChatGPT*. Ghorashi et al. (2023) observam que os *Chatbots*, ao dependerem de fontes abertas, podem não ser suficientemente confiáveis, e Lee (2023) ressalta que, embora o *ChatGPT* possa promover o pensamento crítico, há um risco de vieses embutidos, como destacado por Hosseini et al. (2023).

No tocante à eficácia do *ChatGPT* como ferramenta de aprendizagem autônoma, Choi (2023) e Abd-Alrazaq et al. (2023) observam que a IA frequentemente apresenta inconsistências e erros, comprometendo a aprendizagem. Essa perspectiva é ampliada por Ignjatović e Stevanović (2023), que enfatizam a necessidade de suporte humano para superar a falta de pensamento crítico e a tendência do *ChatGPT* a cometer erros em cálculos matemáticos complexos.

A variabilidade da precisão do *ChatGPT* é mais evidente quando analisada por Bagde et al. (2023), que sugere que a precisão pode variar significativamente entre os estudos. Totlis et al. (2023) também observam que o *ChatGPT* não oferece imagens ou pesquisas em tempo real, limitando sua utilidade em áreas como anatomia. Esses problemas são enfatizados por Breeding et al. (2023) e Cross et al. (2023), que destacam a necessidade de filtragem de informações incorretas devido à tendência do *ChatGPT* de apresentar referências "alucinantes."

Em relação à aplicação da IA na prática clínica, Roos et al. (2023) destacam que o *ChatGPT* ainda enfrenta problemas de alucinação e carece de acesso a bases de conhecimento externas, o que compromete sua capacidade de fornecer informações precisas. Khlaif et al. (2023) também destacam que a ausência de

interação humana limita o desenvolvimento de habilidades essenciais, como comunicação e empatia, na formação médica.

Para que a IA seja efetivamente integrada à educação médica, Weidener e Fischer (2024) e Knopp et al. (2023) reforçam a necessidade de abordar questões éticas e garantir que os currículos de ensino médico incluam discussões sobre a utilização da IA de maneira responsável e crítica. Park (2023) reforça essa necessidade ao identificar que a produção de respostas sem evidências de suporte é uma das desvantagens mais significativas do *ChatGPT*.

Na prática médica, Ghosh e Bir (2023) e Huang et al. (2023) identificam que o *ChatGPT* ainda não é capaz de lidar com inferências complexas e raciocínio clínico, sugerindo que a IA não pode substituir especialistas humanos em consultas que exigem interpretação detalhada. Morjaria et al. (2023) observam que as respostas do *ChatGPT* foram confundidas com respostas humanas, levantando preocupações sobre a validade de avaliações clínicas.

Cheung et al. (2023) e Panthier e Gatinel (2023) enfatizam que a dependência excessiva do *ChatGPT* pode prejudicar o desenvolvimento do pensamento crítico. Wong (2023) destaca os desafios relacionados à precisão e autenticidade das informações geradas pela IA. Enquanto Weidener e Fischer (2023) apontam que a complexidade e a falta de corpo docente especializado em ética da IA dificultam sua integração no ensino médico, Huynh, Bonebrake e Schultis (2023) destacam a baixa precisão do *ChatGPT* em exames específicos.

A fragilidade do *ChatGPT* em fornecer informações confiáveis é destacada por Sonntagbauer, Haar e Kluge (2023), que identificam o risco das "alucinações" na geração de informações. Essa perspectiva é corroborada por Ismail Dergaa et al. (2023), que destacam a potencial integração de informações falsas em pesquisas. Tsang (2023) também reforça as limitações do *ChatGPT* devido ao seu treinamento com dados até 2021.

Brin et al. (2023) e Meo et al. (2023) ressaltam a necessidade de não substituir a expertise humana pela IA, pois o *ChatGPT* ainda demonstra falta de habilidades sociais e consistência. Jeyaraman et al. (2023) reforça que a IA deve ser utilizada com cautela, sem suplantando o conhecimento humano. Finalmente, Chen et al. (2023) concluem que o *ChatGPT* enfrenta dificuldades em fornecer decisões clínicas precisas e confiáveis.

Huh (2023) completa esse debate ao destacar que o *ChatGPT* foi inferior aos discentes de medicina na interpretação de perguntas de múltipla escolha, não conseguindo processar informações contidas em gráficos, figuras e tabelas, o que ressalta ainda mais suas limitações.

Finalmente, Ilgaz e Çelik (2023) e Agarwal, Sharma e Goswami (2023) destacam que a geração de artigos e questões de múltipla escolha pelo *ChatGPT* ainda é insuficiente e menos eficaz em comparação com outras ferramentas, sendo necessário um uso cauteloso. Mahajan et al. (2023) também identificam que a validade das questões geradas pela IA foi significativamente inferior, comprometendo a eficácia da avaliação dos discentes.

Os autores analisados concordam que o *ChatGPT* enfrenta desafios significativos em sua aplicação no ensino médico, principalmente relacionados à precisão, consistência, capacidade de raciocínio crítico e questões éticas. Embora apresente potencial como uma ferramenta de apoio, sua tendência a gerar informações incorretas, "alucinações" e falta de habilidades interpessoais limita sua eficácia como substituto do conhecimento humano.

Categoria 4 – Contribuições do uso do *ChatGPT* no ensino médico

A integração do *ChatGPT* na educação médica tem sido amplamente discutida pelos autores, que exploram suas contribuições e potencial para aprimorar o processo de ensino e aprendizado. Amedu e Ohene-Botwe (2024) reconhecem o papel do *ChatGPT* como um tutor virtual, especialmente na área de radiografia, destacando seu potencial para fornecer feedback e reforçar conceitos clínicos básicos. Ignjatović e Stevanović (2023) reforçam essa perspectiva ao afirmar que o *ChatGPT* mantém o interesse dos discentes por meio de respostas claras e detalhadas, incentivando o desenvolvimento de habilidades críticas. A partir dessas primeiras impressões, os autores expandem a discussão sobre como o *ChatGPT* pode ser uma ferramenta valiosa para a formação de discentes e profissionais de medicina.

A capacidade do *ChatGPT* de personalizar e adaptar o aprendizado às necessidades dos discentes é um dos aspectos mais destacados por diversos autores. Preiksaitis e Rose (2023) enfatizam que a IA tem mostrado potencial em oferecer experiências de aprendizado personalizadas, auxiliando especialmente falantes não nativos de inglês na escrita e pesquisa. Da mesma forma, Alkhaaldi et al. (2023) destacam que o *ChatGPT* proporciona feedback imediato e acesso a informações baseadas em evidências, o que enriquece o treinamento dos discentes e melhora seu atendimento ao paciente. Além disso, Ghorashi et al. (2023) apontam que os *Chatbots* têm o potencial de aprimorar a compreensão e a retenção do conhecimento médico, oferecendo apoio ao desenvolvimento de siglas e mnemônicos para diagnósticos.

Temsah et al. (2023) também ressaltam o papel do *ChatGPT* no desenvolvimento de habilidades clínicas, mencionando como ele pode simular cenários clínicos e fornecer acesso rápido a informações médicas. Essa perspectiva é complementada por Hosseini et al. (2023), que mencionam a eficiência do *ChatGPT* na documentação médica como um benefício importante para o sistema de saúde. Choi (2023) acrescenta que, apesar das limitações, o *ChatGPT* mostrou uma precisão geral superior ao desempenho médio de discentes em exames, sugerindo seu valor como ferramenta de aprendizado.

O impacto positivo do *ChatGPT* na educação médica é confirmado por Abd-Alrazaq et al. (2023), que afirmam que os modelos de linguagem como o GPT-4 têm um papel significativo no planejamento curricular e no fornecimento de feedback personalizado. De forma semelhante, Bagde et al. (2023) ressaltam que o *ChatGPT* tem sido preciso em áreas como exames médicos e diagnóstico por imagem, evidenciando sua utilidade na prática clínica. Totlis et al. (2023) também destacam a capacidade do *ChatGPT* de gerar questionários de múltipla escolha e facilitar o aprendizado de terminologia anatômica complexa, proporcionando um ambiente de aprendizagem interativo e enriquecedor.

A adoção do *ChatGPT* por professores e discentes demonstra sua aceitação e eficácia na educação médica. Cross et al. (2023) mencionam que a ferramenta foi rapidamente adotada por educadores, sugerindo que seu uso pode se expandir ainda mais no futuro. Khlaif et al. (2023) também ressaltam o potencial do *ChatGPT* para aumentar a produtividade na pesquisa e na redação acadêmica, integrando-se ao currículo médico para apoiar o aprendizado autodirigido e a tomada de decisões informadas.

Além disso, Weidener e Fischer (2024) e Knopp et al. (2023) enfatizam a necessidade de incorporar a IA, como o *ChatGPT*, nos currículos médicos, destacando sua capacidade de preparar os futuros profissionais para os desafios da prática médica moderna. A efetividade do *ChatGPT* em avaliações é corroborada por

Roos et al. (2023), que apontam seu desempenho superior em exames, e Park (2023), que revela que os discentes valorizam a capacidade do *ChatGPT* de fornecer respostas e resumir informações de forma eficiente.

A aplicação do *ChatGPT* na resolução de problemas complexos também é amplamente reconhecida. Ghosh e Bir (2023) e Huang et al. (2023) ressaltam sua eficácia em áreas como patologia e microbiologia, enquanto Morjaria et al. (2023) destacam que seu sucesso em avaliações de resposta curta representa uma nova forma de medir competências no ensino médico. Ilgaz e Çelik (2023) reforçam essa ideia ao observar que o *ChatGPT* e outras ferramentas de IA têm o potencial de criar ambientes de aprendizagem mais interativos, contribuindo para uma educação médica mais eficaz.

No âmbito da avaliação, Agarwal, Sharma e Goswami (2023) e Mahajan et al. (2023) reconhecem que o *ChatGPT* desempenhou bem na geração e resolução de MCQs (questões de múltipla escolha), demonstrando seu potencial para aumentar a eficiência nas avaliações médicas. Cheung et al. (2023) reforçam esse ponto ao afirmar que o *ChatGPT* pode reduzir a carga de trabalho da equipe médica, fornecendo rapidamente questões de alta qualidade.

Na prática clínica, Panthier e Gatinel (2023) e Wong (2023) ressaltam a eficácia do *ChatGPT* em avaliações oftalmológicas e sua capacidade de aprimorar o processo de tomada de decisões médicas, respectivamente. Weidener e Fischer (2023) e Huynh, Bonebrake e Schultis (2023) também destacam a importância de integrar a ética da IA na formação médica, enfatizando como o *ChatGPT* pode contribuir para um melhor entendimento dessas questões.

A aplicação prática do *ChatGPT* em diferentes áreas da medicina é ainda mais destacada por Sonntagbauer, Haar e Kluge (2023), que reconhecem seu potencial na redação de documentação médica, e Ismail Dergaa et al. (2023), que enfatizam seu valor na pesquisa acadêmica. Sallam (2023) também reconhece sua eficiência na escrita e pesquisa científica, enquanto Cheng et al. (2023) mencionam sua eficácia na geração de resumos clínicos.

No contexto clínico, Qu et al. (2023) e Bakkum et al. (2023) observam que o *ChatGPT* pode fornecer diagnósticos diferenciais precisos e criar casos clínicos diversificados, destacando seu papel como uma ferramenta analítica valiosa. Vaira et al. (2023) também ressaltam o potencial do *ChatGPT* ao resolver cenários clínicos complexos com precisão. Preiksaitis et al. (2023) e Tovalino et al. (2023) destacam sua eficácia na geração de cartas de recomendação e o papel crescente da IA na medicina e na saúde.

Apesar de algumas limitações, como observado por Ngo et al. (2024), que mencionam a necessidade de revisões cuidadosas, a aceitação e o uso do *ChatGPT* por discentes sugerem seu papel cada vez mais importante na educação médica. Li (2023) e Dave, Athaluri e Sing (2023) também reconhecem os benefícios dos *Chatbots* no ensino e na prática médica, enquanto Das et al. (2023) e Huh (2023) destacam sua capacidade de responder a perguntas complexas e fornecer explicações detalhadas.

Brin et al. (2023) e Tsang (2023) destacam a eficácia do *ChatGPT* em enfrentar dilemas éticos e fornecer resumos clínicos, enquanto Meo et al. (2023) reconhecem sua habilidade de tradução e interpretação de informações complexas. Chen et al. (2023) e Jeyaraman et al. (2023) finalizam a discussão, destacando o potencial do *ChatGPT* para auxiliar na educação médica e na gestão clínica, desde que utilizado em conjunto com a expertise humana.

A discussão entre os autores demonstra que o *ChatGPT* possui uma série de contribuições significativas para o ensino médico, destacando sua capacidade de

personalizar o aprendizado, fornecer feedback imediato e aprimorar a compreensão de conceitos complexos. Sua aplicação em diferentes áreas da medicina, desde a elaboração de questionários até o fornecimento de diagnósticos diferenciais, revela seu potencial para enriquecer o processo educacional e clínico. Embora ainda existam desafios e a necessidade de supervisão humana, o *ChatGPT* surge como uma ferramenta inovadora e valiosa que pode complementar e melhorar a formação de profissionais de saúde, oferecendo oportunidades para um aprendizado mais eficiente e adaptativo no contexto da educação médica.

Categoria 5 – Desempenho do *ChatGPT* em exames de escolas médicas

A crescente incorporação do *ChatGPT* em exames de escolas de medicina tem gerado discussões sobre seu desempenho e seu potencial para contribuir no processo de avaliação médica. Tensah et al. (2023) destacam que o *ChatGPT* demonstrou um desempenho notável em exames padronizados como o USMLE, alcançando pontuações próximas ou superiores às necessárias para aprovação, embora ainda enfrente dificuldades em lidar com casos médicos complexos. Essa percepção é corroborada por Preiksaitis e Rose (2023), que, apesar de reconhecerem o desempenho impressionante do *ChatGPT*, apontam a presença de erros significativos, sugerindo a necessidade de reavaliar os métodos de avaliação médica tradicionais.

O desempenho do *ChatGPT* em exames de medicina é um aspecto que tem impressionado vários pesquisadores. Por exemplo, Alkhaaldi et al. (2023) relatam que o modelo demonstrou competência equivalente à de discentes de terceiro ano de medicina nas etapas 1 e 2 do USMLE, evidenciando sua capacidade de processar e interpretar grandes volumes de informação de maneira precisa. Wójcik et al. (2023) também destacam que o *ChatGPT* conseguiu passar no USMLE com pontuações que variaram entre 52,4% e 75%, sem treinamento especial, o que sugere sua eficácia em exames complexos.

Essa eficácia é reforçada por Hosseini et al. (2023), que relatam que o *ChatGPT* atingiu um desempenho próximo ao nível de aprovação no USMLE, indicando que o modelo pode ser uma ferramenta eficaz para discentes em preparação para exames. Choi (2023) amplia essa discussão ao observar que o *ChatGPT* alcançou uma precisão geral de 76% no K-CBMSE, superando a média dos discentes coreanos, destacando-se especialmente em questões de recordação e interpretação.

No entanto, questões éticas e de integridade acadêmica são levantadas por Abd-Alrazaq et al. (2023), que destacam os riscos de plágio e trapaça, pois o *ChatGPT* pode produzir respostas semelhantes às de textos humanos. Esta preocupação é validada por Ignjatović e Stevanović (2023), que observam que, embora o *ChatGPT* tenha conseguido passar em todas as três etapas do USMLE, ainda há questões sobre como ele pode ser utilizado de forma justa e ética em contextos de avaliação.

A precisão do *ChatGPT* apresenta variações, como observado por Bagde et al. (2023), que relatam taxas de precisão entre 42% e 64,4% em diferentes conjuntos de dados do USMLE. Apesar dessas variações, Cheung et al. (2023) indicam que o *ChatGPT* demonstrou conhecimento suficiente para passar no Exame de Licenciamento Médico dos Estados Unidos, sugerindo que pode ter um desempenho comparável ao de discentes em certos contextos de avaliação. Breeding et al. (2023) acrescentam que o *ChatGPT* foi capaz de passar na etapa 1 do USMLE e apresentou clareza e organização na explicação de patologias cirúrgicas, reforçando seu potencial como ferramenta de apoio educacional.

A evolução do *ChatGPT* é evidenciada por estudos como o de Roos et al. (2023), que relatam que o GPT-4 obteve o melhor desempenho geral em exames, respondendo corretamente a 88,1% das perguntas e superando versões anteriores, como o GPT-3.5-Turbo. Cross et al. (2023) destacam que o *ChatGPT* versão 3.5 conseguiu atingir a marca de aprovação de 60% em testes padronizados para o Exame de Licenciamento Médico dos EUA (USMLE), e a versão 4 alcançou uma média superior a 80% em testes similares, demonstrando uma melhora significativa no desempenho. Khlaif et al. (2023) confirmam essa melhoria, observando que o *ChatGPT* atingiu uma pontuação geral de 88,1%, destacando-se na qualidade de suas respostas em comparação com os discentes.

Estudos de Weidener e Fischer (2024) e Knopp et al. (2023) enfatizam que o *ChatGPT*, assim como outros modelos de linguagem, tem mostrado capacidade de concluir exames de licenciamento médico com sucesso, demonstrando seu potencial para apoiar a educação e a prática médica. Park (2023) também observa que a maioria dos discentes avaliou o feedback do *ChatGPT* como altamente preciso, o que sugere que ele pode ser uma ferramenta eficaz de estudo e revisão.

Ghosh e Bir (2023) reconhecem que o *ChatGPT* tem a capacidade de passar no USMLE e fornecer insights clínicos comparáveis aos de um discente de medicina de terceiro ano. Já Huang et al. (2023) destacam que o GPT-4 apresentou um desempenho superior em diversas áreas, como cuidados geriátricos e maternos. Morjaria et al. (2023) apontam que a consistência das respostas do *ChatGPT* é comparável à dos tutores, sugerindo que ele pode ser uma ferramenta confiável para avaliação de desempenho médico.

Em relação ao uso do *ChatGPT* em diferentes idiomas, Ilgaz e Çelik (2023) descobriram que a precisão do modelo variava conforme a língua, obtendo o melhor desempenho na versão em turco. Por outro lado, Agarwal, Sharma e Goswami (2023) e Mahajan et al. (2023) destacam que o *ChatGPT* obteve distinção em exames universitários de fisiologia, alcançando mais de 75% das notas, o que demonstra sua adaptabilidade em diferentes disciplinas médicas.

O desempenho do *ChatGPT* em exames específicos também é destacado por diversos autores. Panthier e Gatinel (2023) relataram que o modelo obteve 91% de sucesso no exame do Conselho Europeu de Oftalmologia, enquanto Wong (2023) observa que o *ChatGPT* conseguiu passar no USMLE, embora tenha falhado em exames de suporte cardiovascular. Weidener e Fischer (2023) também destacam que o *ChatGPT* demonstrou competência ao passar na parte escrita do USMLE, indicando seu potencial impacto no ensino médico.

Além disso, Huynh, Bonebrake e Schultis (2023) relatam que o *ChatGPT* teve um desempenho próximo ao limite de aprovação nos três exames do USMLE, e Sonntagbauer, Haar e Kluge (2023) destacam que o GPT-4 respondeu corretamente a mais de 90% das perguntas do exame. Ismail Dergaa et al. (2023) e Sallam (2023) reforçam que o *ChatGPT* pode ser comparável a residentes médicos em certas avaliações.

Em termos de habilidades diagnósticas, Cheng et al. (2023) destacam a precisão do *ChatGPT* em detectar distorções cognitivas em mensagens de texto, evidenciando habilidades avançadas de análise. Já Qu et al. (2023) e Vaira et al. (2023) afirmam que o *ChatGPT* forneceu diagnósticos corretos em até 81,7% dos casos, reforçando seu potencial em exames clínicos.

Finalmente, Amedu e Ohene-Botwe (2024) sugerem que o *ChatGPT* pode ser útil na elaboração de perguntas de avaliação, enquanto Ngo et al. (2024) alertam para inconsistências no desempenho da IA em exames médicos. Outros autores como Li

(2023), Das et al. (2023), Huh (2023), Tsang (2023), Brin et al. (2023), Meo et al. (2023) e Chen et al. (2023) reforçam o desempenho variado do *ChatGPT* em exames médicos, indicando que há espaço para melhorias, mas também reconhecendo seu crescente potencial como uma ferramenta de apoio no aprendizado médico.

A análise dos diferentes estudos revela que o *ChatGPT* tem um desempenho notável em exames médicos, frequentemente alcançando taxas de aprovação comparáveis às de discentes de medicina e até mesmo superando-os em algumas áreas. No entanto, as limitações do modelo, como inconsistências e variabilidade no desempenho, indicam que ele ainda não é um substituto completo para a expertise humana. Embora o *ChatGPT* ofereça um grande potencial como ferramenta complementar na educação médica, seu uso deve ser acompanhado de supervisão e critérios de avaliação rigorosos para garantir que a aprendizagem e o desenvolvimento das habilidades clínicas sejam realizados de maneira eficaz e ética. O avanço da tecnologia, como evidenciado pela evolução do GPT-3.5 para o GPT-4, sugere que o desempenho do *ChatGPT* em exames médicos pode continuar a melhorar, mas a integração cuidadosa de sua utilização será fundamental para maximizar seus benefícios na formação de profissionais médicos competentes e éticos.

Categoria 6 - Direcionamentos futuros e recomendações para a integração do *ChatGPT* na educação médica

A integração do *ChatGPT* no ensino médico tem despertado discussões sobre seu papel e impacto no processo de aprendizagem e avaliação. Cross et al. (2023) destacam a necessidade de monitorar a aceitação e o uso contínuo do *ChatGPT*, considerando seus efeitos nos hábitos de estudo dos discentes e os possíveis desafios emergentes. Temsah et al. (2023) sugerem que o *ChatGPT* deve ser incorporado como uma ferramenta complementar ao currículo de medicina, com a devida supervisão e ênfase em diretrizes éticas. Esses autores abrem o caminho para um debate mais amplo sobre como o *ChatGPT* pode se integrar de forma segura e eficaz na educação médica.

Muitos autores concordam que a integração do *ChatGPT* deve ser realizada com cuidado, com foco em seu potencial e limitações. Roos et al. (2023) sugerem que futuros estudos explorem a detecção de respostas incorretas e a preparação para exames com o uso de *LLMs*, como o *ChatGPT*, incluindo a possibilidade de incorporar reconhecimento de imagem para suporte diagnóstico. Breeding et al. (2023) complementam ao ressaltar que o *ChatGPT* deve ser avaliado em sua capacidade de fornecer informações sobre uma gama maior de condições médicas, utilizando amostras maiores para testar a consistência e generalização dos resultados.

A integração cuidadosa é destacada por Khlaif et al. (2023), que enfatizam a importância de preservar o toque humano na educação e na pesquisa, utilizando o *ChatGPT* em conjunto com insights de profissionais. Preiksaitis e Rose (2023) concordam, recomendando currículos que melhorem a alfabetização em IA e incentivem a avaliação crítica. Alkhaaldi et al. (2023) acrescentam que a IA deve ser introduzida no início da formação médica para preparar futuros médicos para liderar o uso dessas tecnologias.

O desenvolvimento de estruturas éticas robustas é outro ponto de convergência. Ghorashi et al. (2023) e Lee (2023) enfatizam a necessidade de estabelecer diretrizes claras e segurança no uso do *ChatGPT*, com foco na precisão das informações. Choi (2023) reforça a importância de treinar os discentes para avaliar criticamente as respostas do *ChatGPT* e rastrear evidências até suas fontes,

garantindo que o uso da ferramenta complemente, mas não substitua, o pensamento independente.

Enquanto alguns autores alertam sobre os desafios éticos e de plágio, Abd-Alrazaq et al. (2023) apontam que a análise das oportunidades e desafios do uso do *ChatGPT* deve orientar práticas futuras. Ignjatović e Stevanović (2023) recomendam que educadores apresentem aos discentes as vantagens e desvantagens da ferramenta, enquanto Bagde et al. (2023) destacam a necessidade de estudos para entender os fatores que influenciam a precisão do *ChatGPT* e como aprimorá-lo.

Há consenso sobre o papel complementar do *ChatGPT* na educação médica. Totlis et al. (2023) e Weidener e Fischer (2024) sugerem que o *ChatGPT* pode ser uma ferramenta valiosa, mas não deve substituir o papel dos educadores, reforçando a importância da integração da IA e da ética nos currículos. Knopp et al. (2023) sugerem que uma estrutura ética e a colaboração interdisciplinar são fundamentais para o uso responsável da IA.

Park (2023) aponta que a colaboração entre educadores, pesquisadores e profissionais é crucial para a integração do *ChatGPT*, destacando a necessidade de diretrizes claras e práticas eficazes. Ghosh e Bir (2023) reforçam a importância de monitorar e atualizar regularmente os sistemas de IA para garantir relevância e alinhamento com avanços na medicina.

Quando se trata da aplicação clínica do *ChatGPT*, Huang et al. (2023) destacam seu desempenho superior em comparação com residentes médicos em certas áreas, validando sua credibilidade como ferramenta educacional. Morjaria et al. (2023) sugerem que estudos futuros devem combinar as respostas do *ChatGPT* com as de discentes para aprimorar o uso da IA na educação médica.

Além disso, a adaptação da IA a diferentes contextos e linguagens é reforçada por Ilgaz e Çelik (2023), que sugerem a exploração de figuras 2D e 3D para aprimorar o aprendizado em anatomia. Na área de avaliação, Agarwal, Sharma e Goswami (2023) e Mahajan et al. (2023) recomendam o uso de questões de múltipla escolha geradas pela IA para revolucionar a educação médica.

A moral e a necessidade de regulamentação são destacadas por Cheung et al. (2023), que recomendam que o *ChatGPT* seja usado para complementar o ensino tradicional com treinamento adequado dos educadores. Panthier e Gatinel (2023) sugerem que o *ChatGPT* seja integrado aos currículos como complemento à expertise humana, com atualizações contínuas.

Wong (2023) reforça a necessidade de regulamentações que garantam a conformidade e transparência no uso do *ChatGPT*, enquanto Weidener e Fischer (2023) recomendam mais pesquisas sobre a ética da IA na educação médica. Huynh, Bonebrake e Schultis (2023) sugerem pesquisas adicionais para entender melhor as capacidades e limitações dos LLMs, reforçando a necessidade de considerar as implicações éticas no uso da IA.

A aplicação do *ChatGPT* em ambientes clínicos é uma tendência inevitável, segundo Sonntagbauer, Haar e Kluge (2023), e Ismail Dergaa et al. (2023) avultam a importância de educadores experimentarem a ferramenta para entender seu potencial. Sallam (2023) recomenda a implementação de um código de ética para o uso do *ChatGPT*.

Explorando áreas mais específicas, Cheng et al. (2023) identificam o potencial do *ChatGPT* em diagnóstico psiquiátrico, enquanto Qu et al. (2023) sugerem a colaboração de profissionais de saúde no desenvolvimento da IA. Bakkum et al. (2023) defendem o compartilhamento de casos educacionais criados com o *ChatGPT* para ampliar seu uso.

Vaira et al. (2023) recomendam o desenvolvimento de protocolos rigorosos para garantir a precisão do *ChatGPT*, enquanto Preiksaitis et al. (2023) sugerem que a IA pode reformular práticas acadêmicas. Tovalino et al. (2023) reforçam a necessidade de soluções baseadas em IA com colaboração multidisciplinar. Amedu e Ohene-Botwe (2024) e Ngo et al. (2024) concordam na necessidade de uso cuidadoso e aprimoramento contínuo do *ChatGPT*.

No futuro, Li (2023) prevê a expansão do uso de *Chatbots* na educação médica, enquanto Dave, Athaluri e Sing (2023) destacam a importância de uma integração supervisionada. Das et al. (2023) enfatizam a necessidade de melhorias contínuas no desenvolvimento da IA para seu uso acadêmico. Huh (2023) prevê que as capacidades interpretativas do *ChatGPT* evoluirão rapidamente e sugere sua incorporação nos currículos médicos. Tsang (2023) defende políticas explícitas sobre o uso da IA nas escolas de medicina.

Afinal, Brin et al. (2023) recomendam pesquisas que incluam cenários éticos mais diversificados para avaliar as habilidades sociais dos *LLMs*, enquanto Meo et al. (2023) sugerem o desenvolvimento de métodos educacionais que aproveitem ao máximo o *ChatGPT*. Jeyaraman et al. (2023) destacam a necessidade de abordar a integração da IA com entusiasmo, mas também com cautela. Chen et al. (2023) e Hosseini et al. (2023) reforçam a importância de estudos futuros que explorem a capacidade do *ChatGPT* de se integrar à tomada de decisões médicas e melhorar a centralização de informações.

Os autores concordam que o *ChatGPT* tem um potencial significativo para enriquecer a educação médica, desde que sua integração seja feita de forma cuidadosa e ética. A colaboração interdisciplinar, o desenvolvimento de diretrizes claras e a promoção da alfabetização em IA são elementos cruciais para garantir que o *ChatGPT* seja uma ferramenta eficaz e segura. Embora o *ChatGPT* ofereça oportunidades únicas para apoiar a aprendizagem e prática médica, sua aplicação deve complementar, e não substituir, a experiência humana e o julgamento clínico. À medida que a tecnologia evolui, é fundamental que sua implementação seja acompanhada de uma avaliação crítica contínua, garantindo que o *ChatGPT* contribua para a formação de profissionais de saúde competentes, éticos e bem preparados para os desafios do futuro.

5. Conclusão

A crescente integração de tecnologias de inteligência artificial no ensino médico revela-se promissora, mas também desafiadora. Este estudo analisou a eficácia do *ChatGPT* como ferramenta complementar na formação de discentes de medicina, destacando tanto seu potencial quanto suas limitações. Por meio de um mapeamento sistemático e descritivo, sustentado pela metodologia qualitativa e pela estratégia PICO, foram examinadas contribuições, desafios e considerações éticas dessa tecnologia no ambiente educacional.

Os resultados evidenciam que o *ChatGPT* pode desempenhar um papel significativo no aprendizado médico, proporcionando experiências personalizadas, feedback imediato e apoio no desenvolvimento de habilidades críticas. No entanto, sua eficácia depende de supervisão humana constante, dado que informações imprecisas podem comprometer o processo educacional. A confiabilidade do modelo reforça a necessidade de orientação pedagógica e avaliação contínua para garantir a qualidade do conteúdo oferecido.

Do ponto de vista ético, a utilização do *ChatGPT* requer atenção especial. Sua implementação deve respeitar os limites da tecnologia e priorizar a formação ética dos

futuros médicos, garantindo que a IA não substitua a expertise humana, mas sim atue como uma aliada estratégica no ensino e na prática médica. A supervisão constante é fundamental para equilibrar os benefícios da ferramenta com os desafios inerentes à sua aplicação.

Entre os desafios identificados, estão a falta de habilidades interpessoais e as inconsistências ocasionais do *ChatGPT*, que o tornam inadequado como substituto completo do conhecimento humano. Esses fatores reforçam a importância de uma integração cautelosa e ética, considerando as particularidades do ensino médico e os riscos associados ao uso indiscriminado da tecnologia.

Por outro lado, o desempenho do *ChatGPT* em exames de medicina demonstra seu potencial como um recurso complementar. Apesar de alcançar taxas de aprovação similares às de discentes humanos, ele ainda exige validação rigorosa de suas respostas e acompanhamento por parte de educadores qualificados. Essas medidas são essenciais para consolidar sua utilidade como ferramenta educacional.

O futuro da integração do *ChatGPT* no ensino médico depende do estabelecimento de diretrizes claras, da promoção de alfabetização em IA e de avaliações críticas contínuas. Essas iniciativas são indispensáveis para que a tecnologia complemente o julgamento clínico humano, enriqueça a experiência de aprendizagem e contribua para a formação de profissionais de saúde mais bem preparados.

Assim, conclui-se que o *ChatGPT*, quando utilizado de maneira ética, supervisionada e integrada aos métodos tradicionais de ensino, tem o potencial de transformar a educação médica. Contudo, seu uso exige cautela, planejamento e compromisso com a qualidade, a ética e a formação de médicos críticos e capacitados para os desafios da prática profissional.

Referências

ABD-ALRAZAQ, A.; ALSAAD, R.; ALHUWAIL, D.; et al. Large language models in medical education: opportunities, challenges, and future directions. *JMIR Medical Education*, [s.l.], v. 9, e48291, 2023. DOI: 10.2196/48291. Disponível em: <https://preprints.jmir.org/preprint/48291>. Acesso em: 3 jul. 2024.

AGARWAL, M.; SHARMA, P.; GOSWAMI, A. Analysing the applicability of *ChatGPT*, Bard, and Bing to generate reasoning-based multiple-choice questions in medical physiology. *Cureus*, [s.l.], v. 15, n. 6, e40977, 2023. DOI: 10.7759/cureus.40977. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37519497/>. Acesso em: 8 out. 2024.

ALKHAALDI, S. M. I.; KASSAB, C. H.; DIMASSI, Z.; et al. Medical student experiences and perceptions of *ChatGPT* and artificial intelligence: cross-sectional study. *JMIR Medical Education*, [s.l.], v. 9, e51302, 2023. DOI: 10.2196/51302. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/38133911/>. Acesso em: 26 set. 2024.

AMEDU, C.; OHENE-BOTWE, B. Harnessing the benefits of *ChatGPT* for radiography education: a discussion paper. *Radiography*, v. 30, p. 209-216, 2024. DOI: 10.1016/j.radi.2023.11.009. Disponível em: [https://www.radiographyonline.com/article/S1078-8174\(23\)00225-0/fulltext](https://www.radiographyonline.com/article/S1078-8174(23)00225-0/fulltext). Acesso em: 10 out. 2024.

BAGDE, H.; DHOPTE, A.; KHURSHEED, M.; et al. A systematic review and meta-analysis on *ChatGPT* and its utilization in medical and dental research. *Heliyon*, v. 9, e23050, 2023. Disponível em: <https://doi.org/10.1016/j.heliyon.2023.e23050>. Acesso em: 14 mai. 2024.

BAKKUM, M. J.; HARTJES, M. G.; PIËT, J. D.; et al. Using artificial intelligence to create diverse and inclusive medical case vignettes for education. *British Journal of Clinical Pharmacology*, 2024, p. 1-9. DOI: 10.1111/bcp.15977. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/38016816/>. Acesso em: 26 set. 2024.

BASKARA, R.; MUKARTO, M. Exploring the implications of *ChatGPT* for language learning in higher education. *Indonesian Journal of English Language Teaching and Applied Linguistics*, v. 7, n. 2, p. 343-358, 2023. DOI: 10.21093/ijeltal.v7i2.1387. Disponível em: <https://ijeltal.org/index.php/ijeltal/article/view/1387>. Acesso em: 3 abr. 2024.

BORGES, V. A.; SOUSA, S. O. Aprendizagem baseada em problemas e as tecnologias digitais de informação e comunicação. 2024. Disponível em: <https://doi.org/10.23925/1809-3876.2024v22e61490>. Acesso em: 11 jul. 2024.

BREEDING, T.; MARTINEZ, B.; PATEL, H.; et al. The utilization of *ChatGPT* in reshaping future medical education and learning perspectives: a curse or a blessing?. *The American Surgeon*, v. 0, p. 1-7, 2023. DOI: 10.1177/00031348231180950. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37309705/>. Acesso em: 18 jun. 2024.

BRIN, D.; SORIN, V.; VAID, A.; et al. Comparing *ChatGPT* and GPT-4 performance in USMLE soft skill assessments. *Scientific Reports*, v. 13, 16492, 2023. Disponível em: <https://doi.org/10.1038/s41598-023-43436-9>. Acesso em: 18 jun. 2024.

CARVALHO, M. W. S.; ANDRADE, W. Z. N.; CANSANÇÃO, I. F.; et al. Aprendizagem baseada em problemas como método de ensino na formação médica. *Revista Docência do Ensino Superior*, Belo Horizonte, v. 10, p. 1–21, 2020. DOI: 10.35699/2237-5864.2020.19801. Disponível em: <https://periodicos.ufmg.br/index.php/rdes/article/view/19801>. Acesso em: 2 ago. 2024.

CHEN, T. C.; MULTALA, E.; KEARNS, P.; et al. Assessment of *ChatGPT*'s performance on neurology written board examination questions. *BMJ Neurology Open*, v. 5, e000530, 2023. DOI: 10.1136/bmjno-2023-000530. Disponível em: <https://neurologyopen.bmj.com/content/5/2/e000530>. Acesso em: 19 jul. 2024.

CHENG, S. W.; CHANG, C. W.; CHANG, W. J.; et al. The now and future of *ChatGPT* and GPT in psychiatry. DOI: 10.1111/pcn.13588. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37612880/>. Acesso em: 22 abr. 2024.

CHEUNG, B. H. H.; LAU, G. K. K.; WONG, G. T. C.; et al. *ChatGPT* versus human in generating medical graduate exam multiple choice questions—A multinational prospective study (Hong Kong S.A.R., Singapore, Ireland, and the United Kingdom). *PLOS ONE*, v. 18, n. 8, e0290691, 2023. DOI: 10.1371/journal.pone.0290691.

Disponível em:

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0290691>. Acesso em: 13 jun. 2024.

CHOI, W. Assessment of the capacity of *ChatGPT* as a self-learning tool in medical pharmacology: a study using MCQs. *BMC Medical Education*, v. 23, 864, 2023. DOI: 10.1186/s12909-023-04832-x. Disponível em:

<https://bmcmmededuc.biomedcentral.com/articles/10.1186/s12909-023-04832-x>. Acesso em: 24 jun. 2024.

CROSS, J.; ROBINSON, R.; DEVARAJU, S.; et al. Transforming medical education: assessing the integration of *ChatGPT* into faculty workflows at a Caribbean medical school. *Cureus*, v. 15, n. 7, e41399, 2023. DOI: 10.7759/cureus.41399. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37426402/>. Acesso em: 19 mai. 2024.

DAS, D.; KUMAR, N.; LONGJAM, L. A.; et al. Assessing the capability of *ChatGPT* in answering first- and second-order knowledge questions on microbiology as per competency-based medical education curriculum. *Cureus*, v. 15, n. 3, e36034, 2023. DOI: 10.7759/cureus.36034. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37056538/>. Acesso em: 20 mai. 2024.

DAVE, T.; ATHALURI S. A.; SINGH S. *ChatGPT* in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Front. Artif. Intell.* 6:1169595. Doi: 10.3389/frai.2023.1169595. Disponível em: <https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2023.1169595/full>. Acesso em: 08 jun. 2024.

DERGAA, I.; CHAMARI.; ZMIJEWSKI, P.; et al. From human writing to artificial intelligence generated text: examining the prospects and potential threats of *ChatGPT* in academic writing. *Biol Sport.* 2023; 40(2):615–622. Received: 2023-02-26; Reviewed: 2023-03-04; Re-submitted: 2023-03-06; Accepted: 2023-03-09; Published: 2023-03-15. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37077800/>. Acesso em: 14 mai. 2024.

DIAS, C. M.; SASAKI, D. G. G. APRENDIZAGEM BASEADA EM PROBLEMAS E AS HABILIDADES DO SÉCULO XXI: REVISÃO SISTEMÁTICA. *SciELO Preprints*, 2023. DOI: 10.1590/SciELOPreprints.7179. Disponível em: <https://preprints.scielo.org/index.php/scielo/preprint/view/7179>. Acesso em: 10 jul. 2024.

GHORASHI, N.; ISMAIL, A.; GHOSH P, et al. (August 10, 2023) AI-Powered Chatbots in Medical Education: Potential Applications and Implications. *Cureus* 15(8): e43271. DOI 10.7759/cureus.43271. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37692629/>. Acesso em: 15 mai. 2024.

GHOSH, A.; BIR, A. Evaluating *ChatGPT*'s Ability to Solve HigherOrder Questions on the Competency-Based Medical Education Curriculum in Medical Biochemistry. (April 02, 2023) Evaluating *ChatGPT*'s Ability to Solve Higher-Order Questions on the Competency-Based Medical Education Curriculum in Medical Biochemistry. *Cureus*

15(4): e37023. DOI 10.7759/cureus.37023. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37143631/>. Acesso em: 26 abr. 2024.

HOSSEINI, M.; GAO, C. A.; LIEBOVITZ, D.; et al. An exploratory survey about using *ChatGPT* in education, healthcare, and research. medRxiv preprint. Doi: 10.1101/2023.03.31.23287979; this version posted April 3, 2023. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37066228/>. Acesso em: 12 jul. 2024.

HUANG, R. S. T.; LU, K. J.; MEANEY, C.; et al. Assessment of Resident and AI Chatbot Performance on the University of Toronto Family Medicine Residency Progress Test: Comparative Study. (JMIR Med Educ 2023;9:e50514) doi: 10.2196/50514. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37725411/>. Acesso em: 17 jun. 2024.

HUH, S. Are *ChatGPT*'s knowledge and interpretation ability comparable to those of medical students in Korea for taking a parasitology examination?: a descriptive study. J Educ Eval Health Prof 2023;20:1 . Doi: 10.3352/jeehp.2023.20.1 e ISSN: 1975-5937. Received: January 3, 2023; Accepted: January 11, 2023; Published: January 11, 2023. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/36627845/>. Acesso em: 19 ago. 2024.

HUYNH, L.; BONEBRAKE, B. T.; SCHULTIS, K.; et al. New Artificial Intelligence *ChatGPT* Performs Poorly on the 2022 Self-assessment Study Program for Urology. Disponível em: <https://www.auajournals.org/doi/10.1097/UPJ.0000000000000406>. Acesso em: 20 set. 2024.

IGNJATOVIC, A.; STEVANOVIC, L. Efficacy and limitations of *ChatGPT* as a biostatistical problem-solving tool in medical education in Serbia: a descriptive study. J Educ Eval Health Prof 2023;20:28. Doi: 10.3352/jeehp.2023.20.28 e ISSN: 1975-5937. Disponível em: <https://jeehp.org/journal/view.php?doi=10.3352/jeehp.2023.20.28>. Acesso em: 16 set. 2024.

ILGAZ, H. B.; ÇELİK, Z. (September 15, 2023) The Significance of Artificial Intelligence Platforms in Anatomy Education: An Experience With *ChatGPT* and Google Bard. Cureus 15(9): e45301. DOI 10.7759/cureus.45301. Disponível em: <https://www.researchgate.net/publication/373972430>. Acesso em: 22 ago. 2024.

JAGANATHAN, S.; BHUMINATHAN, S.; RAMESH, M. Problem-Based Learning – An Overview. *Journal of Pharmacy and Bioallied Sciences*. 2024. DOI: 10.4103/jpbs.jpbs_820_23. Disponível em: <https://scispace.com/papers/problem-based-learning-an-overview-nw3wshgn6b>. Acesso em: 17 ago. 2024.

JEYARAMAN, M.; PRIVA, K.; JEYARAMAN, N.; et al. *ChatGPT* in Medical Education and Research: A Boon or a Bane?. *Cureus*, v. 15, n. 8, e44316, 2023. DOI: 10.7759/cureus.44316. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37779749/>. Acesso em: 03 set. 2024.

KHLAIF, Z. N.; MOUSA, A.; HATTAB, M. K.; et al. The Potential and Concerns of Using AI in Scientific Research: *ChatGPT* Performance Evaluation. *JMIR Medical*

Education, v. 9, e47049, 2023. DOI: 10.2196/47049. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37707884/>. Acesso em: 06 set. 2024.

KNOPP, M.; WARM, E. J.; WEBER, D.; et al. AI-Enabled Medical Education: Threads of Change, Promising Futures, and Risky Realities Across Four Potential Future Worlds. *JMIR Medical Education*, v. 9, e50373, 2023. DOI: 10.2196/50373. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/38145471/>. Acesso em: 27 abr. 2024.

LEE, H. Using *ChatGPT* as a Learning Tool in Acupuncture Education: Comparative Study. *JMIR Medical Education*, v. 9, e47427, 2023. DOI: 10.2196/47427. Disponível em: <https://mededu.jmir.org/2023/1/e47427/>. Acesso em: 07 mai. 2024.

LI, L. Role of Chatbots on Gastroenterology: Let's Chat About the Future. *Gastroenterology & Endoscopy*, v. 1, p. 144–149, 2023. DOI: 10.1016/j.gande.2023.06.002. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2949752323000341>. Acesso em: 11 abr. 2024.

MAHAJAN, A. P.; SHABET, C. L.; SMITH, J.; et al. Assessment of Artificial Intelligence Performance on the Otolaryngology Residency In-Service Exam. *OTO Open*, v. 7, n. 4, e98, 2023. DOI: 10.1002/oto2.98. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/38034065/>. Acesso em: 12 jul. 2024.

MAYTA-TOVALINO, A. F.; ESPINOZA-CARHUANCHOB, F.; ALVITEZ-TEMOCHE, D.; et al. Scientometric Analysis on the Use of *ChatGPT*, Artificial Intelligence, or Intelligent Conversational Agent in the Role of Medical Training. *Educación Médica*, v. 25, p. 100873, 2024. DOI: 10.1016/j.edumed.2023.100873. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1575181323000839?via%3Dihub>. Acesso em: 23 jun. 2024.

MEO, S. A.; AL-MASRI, A. A.; ALOTAIBI, M.; et al. *ChatGPT* Knowledge Evaluation in Basic and Clinical Medical Sciences: Multiple Choice Question Examination-Based Performance. *Healthcare*, v. 11, p. 2046, 2023. DOI: 10.3390/healthcare11142046. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37510487/>. Acesso em: 26 jun. 2024.

MORJARIA, L.; BURNS, L.; BRACKEN, K.; et al. Examining the Threat of *ChatGPT* to the Validity of Short Answer Assessments in an Undergraduate Medical Program. *Journal of Medical Education and Curricular Development*, v. 10, p. 1–7, 2023. DOI: 10.1177/23821205231204178. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37780034/>. Acesso em: 05 mai. 2024.

NGO, A.; GRUPTA, S.; PERRINE, O.; et al. *ChatGPT* 3.5 fails to write appropriate multiple choice practice exam questions. *Academic Pathology*, v. 11, n. 1, p. 100099, 2024. DOI:10.1016/j.acpath.2023.100099. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/38162414/>. Acesso em: 18 fev. 2024.

OTTO, D.; SCHRÖDER, N.; DIEKMANN, D.; SANDER, P. Trends and Gaps in Empirical Research on Open Educational Resources (OER): A Systematic Mapping

of the Literature from 2015 to 2019. *Contemporary Educational Technology*, v. 13, p. ep325, 2021. DOI: 10.30935/cedtech/11145. Disponível em: <https://www.researchgate.net/publication/353911347>. Acesso em: 10 out. 2024.

PANTHIER, C.; GATINEL, D. Success of *ChatGPT*, an AI language model, in taking the French language version of the European Board of Ophthalmology examination: A novel approach to medical knowledge assessment. *Journal français d'ophtalmologie*, v. 46, p. 706-711, 2023. DOI: 10.1016/j.jfo.2023.05.006. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0181551223003054?via%3Dihub>. Acesso em: 17 abr. 2024.

PARK, J. Medical students' patterns of using *ChatGPT* as a feedback tool and perceptions of *ChatGPT* in a Leadership and Communication course in Korea: a cross-sectional study. *J Educ Eval Health Prof*, v. 20, p. 29, 2023. DOI: 10.3352/jeehp.2023.20.29. Disponível em: <https://jeehp.org/journal/view.php?doi=10.3352/jeehp.2023.20.29>. Acesso em: 04 abr. 2024.

PETERSEN, K.; FELDMANN, B.; MUFF, S. Systematic mapping of primary studies in software engineering. *IEEE Transactions on Education*, v. 64, n. 4, p. 356-372, 2021.

PREIKSAITIS, C.; NASH, C.; GOTTLIEB, M.; et al. Brain versus bot: Distinguishing letters of recommendation authored by humans compared with artificial intelligence. *AEM Educ Train.*, v. 7, p. 1-9, 2023. DOI: 10.1002/aet2.10924. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/38046089/>. Acesso em: 19 jun. 2024.

PONTES JUNIOR, A. C. F.; NAKAYAMA, B. C. M. S. O Mapeamento Sistemático e a pesquisa qualitativa: (Inter)conexões com os paradigmas da educação. *Revemop*, v. 4, p. e202219, 2022. Disponível em: <https://periodicos.ufop.br/revemop/article/view/5135>. Acesso em: 12 abr. 2024.

QU, R. W.; QURESHI, U.; PETERSEN, G.; et al. Diagnostic and Management Applications of *ChatGPT* in Structured Otolaryngology Clinical Scenarios. *OTO Open*, v. 7, n. 3, p. e67, 2023. DOI: 10.1002/oto2.67. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37614494/>. Acesso em: 20 abr. 2024.

RAFIQ, S.; KAMRAN, F.; AFZAL, A. Investigating the Benefits and Challenges of Interdisciplinary Education in Higher Education Settings. *Journal of Social Research Development*, v. 5, p. 87-100, 2024. DOI: 10.53664/JSRD/05-01-2024-08-87-100. Disponível em: <https://www.researchgate.net/publication/380074464>. Acesso em: 02 jun. 2025.

REZENDE, K. T. A.; COSTA, M. C. G.; RODRIGUES, A.; et al. Aprendizagem Baseada em Problemas em um Curso de Medicina: Desafios na sua Implementação. *Revista Brasileira de Educação Médica*, v. 44, n. 4, p. e119, 2020. DOI: 10.1590/1981-5271v44.4-20190222.ING. Disponível em: <https://doi.org/10.1590/1981-5271v44.4-20190222.ING>. Acesso em: 01 set. 2024.

ROSS, J.; KASAPOVIC, A.; JANSEN, T.; et al. Artificial Intelligence in Medical Education: Comparative Analysis of *ChatGPT*, Bing, and Medical Students in Germany. *JMIR Medical Education*, v. 9, e46482, 2023. DOI: 10.2196/46482. Disponível em: <https://mededu.jmir.org/2023/1/e46482/>. Acesso em: 27 mai. 2024.

SALLAM, M. *ChatGPT* Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. *Healthcare*, v. 11, p. 887, 2023. DOI: 10.3390/healthcare11060887. Disponível em: <https://www.mdpi.com/2227-9032/11/6/887>. Acesso em: 11 abr. 2024.

SONNTAGBAUER, M.; HAAR, M.; KLUGE, S. Künstliche Intelligenz: Wie werden *ChatGPT* und andere KI-Anwendungen unseren ärztlichen Alltag verändern? *Medizinische Klinik - Intensivmedizin und Notfallmedizin*, v. 118, p. 366–371, 2023. DOI: 10.1007/s00063-023-01019-6. Disponível em: <https://link.springer.com/article/10.1007/s00063-023-01019-6>. Acesso em: 13 jun. 2025.

TEMSAH, O.; KHAN, S. A.; CHAIAH, Y.; et al. Overview of Early *ChatGPT*'s Presence in Medical Literature: Insights From a Hybrid Literature Review by *ChatGPT* and Human Experts. *Cureus*, v. 15, n. 4, e37281, 2023. DOI: 10.7759/cureus.37281. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37038381/>. Acesso em: 02 mai. 2024.

TOTLIS, T.; NATSIS, K.; FILOS, D.; et al. The Potential Role of *ChatGPT* and Artificial Intelligence in Anatomy Education: A Conversation with *ChatGPT*. *Surgical and Radiologic Anatomy*, v. 45, p. 1321–1329, 2023. DOI: 10.1007/s00276-023-03229-1. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37584720/>. Acesso em: 18 abr. 2024.

TSANG, R. Practical Applications of *ChatGPT* in Undergraduate Medical Education. *Journal of Medical Education and Curricular Development*, v. 10, p. 1–3, 2023. DOI: 10.1177/23821205231178449. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37255525/>. Acesso em: 19 mar. 2024.

VAIRA, L. A.; LECHIEN, J. R.; ABBATE, V.; et al. Accuracy of *ChatGPT*-Generated Information on Head and Neck and Oromaxillofacial Surgery: A Multicenter Collaborative Analysis. *Otolaryngology–Head and Neck Surgery*, v. 00, n. 00, p. 1–12, 2023. DOI: 10.1002/ohn. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37595113/>. Acesso em: 18 mar. 2024.

WEIDENER, L.; FISCHER, M. Teaching AI Ethics in Medical Education: A Scoping Review of Current Literature and Practices. *Perspectives on Medical Education*, v. 12, n. 1, p. 399–410, 2023. DOI: 10.5334/pme.954. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37868075/>. Acesso em: 08 abr. 2024.

_____. Artificial Intelligence in Medicine: Cross-Sectional Study Among Medical Students on Application, Education, and Ethical Aspects. *JMIR Medical Education*, v. 10, e51247, 2024. DOI: 10.2196/51247. Disponível em: <https://mededu.jmir.org/2024/1/e51247/>. Acesso em: 9 abr. 2024.

WÓJCIK, S.; RULKIEWICZ, A.; PRUSCZYK, P.; et al. Reshaping Medical Education: Performance of *ChatGPT* on a PES Medical Examination. *Cardiology Journal*, v. 30, n. 2, p. 335–336, 2023. DOI: 10.5603/cj.97517. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37830257/>. Acesso em: 28 jun. 2024.

WONG, R. S. Y.; MING, L. C.; RAJA, A. L. A. The Intersection of *ChatGPT*, Clinical Medicine, and Medical Education. *JMIR Medical Education*, v. 9, e47274, 2023. DOI: 10.2196/47274. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37988149/>. Acesso em: 26 mar. 2024.